Laboratory for Information and Decision Systems

The Laboratory for Information and Decision Systems (LIDS) is an interdepartmental laboratory for research and education in systems, control, optimization, communication, networks, and statistical inference. These disciplines, which span the domain of the analytical information and decision sciences, play a critical and pervasive role in science, engineering, and society more broadly. LIDS continues to be at the forefront of research in traditional core disciplines in this area. Over the past few years, the scope of LIDS research has broadened to include new emerging areas at the intersection of several fields such as modeling, analysis, and optimization of multi-agent networked systems, including infrastructure systems (communication networks, traffic networks, and power grids) and social, economic, and financial systems; efficient large-scale inference (using optimization, statistics, and algorithms); and autonomous systems. LIDS provides an intellectually cohesive and interactive environment that fosters the research needed for the future and instills in our students the disciplinary depth and interdisciplinary understanding required of research and engineering leaders of today and tomorrow.

The faculty members within LIDS are principally drawn from the Department of Electrical Engineering and Computer Science (EECS) and the Department of Aeronautics and Astronautics. As a result of the interdisciplinary nature of recent focal work and the pervasiveness of analytical methodology utilized by LIDS researchers, the laboratory has also built strong collaborations and interactions with many other entities at MIT, including the Computer Science and Artificial Intelligence Laboratory, the Research Laboratory of Electronics, the Operations Research Center (ORC), the Media Lab, the Harvard-MIT Division of Health Sciences and Technology, the Department of Civil and Environmental Engineering, the Department of Mechanical Engineering, the Department of Economics, and the Sloan School of Management. LIDS faculty members play leadership roles in major new initiatives involving applications of critical importance to society. These initiatives include projects focusing on transportation systems (e.g., the Future Urban Mobility initiative of the Singapore-MIT Alliance for Research and Technology [SMART]), energy systems (e.g., programs in collaboration with the MIT Energy Initiative), and social, economic, and financial networks in close collaboration with Economics and Sloan.

LIDS is now part of the newly launched Institute for Data, Systems, and Society (IDSS), which brings in statistics, social science, and application domain expertise. LIDS faculty members are playing a pivotal role in defining the new institute's intellectual agenda, leading the statistics effort and flagship projects (in finance, autonomy, and smart cities), designing new academic programs in statistics and complex systems, and leading the search for new faculty in networks and statistics.

LIDS has continued on its growth trajectory this year. We are excited to welcome Professors Guy Bresler and Caroline Uhler, both from the Department of Electrical Engineering and Computer Science. Bresler, whose expertise includes high-dimensional inference in the context of graphical models, information theory, and biology, joined LIDS in September 2015. Uhler joined in October 2015 and brings in expertise in algebraic statistics, combining algebraic geometry and convex optimization to develop new paradigms and algorithms for data analysis. Dr. Kalyan Veeramachaneni joined LIDS as a principal research scientist in March 2016 and brings expertise in machine learning and systems, focusing on developing data-driven predictive analytics platforms for a variety of data science projects. Another excellent addition to LIDS is Professor Ali Jadbabaie, a former faculty member at the University of Pennsylvania who is a well-known researcher in control theory and networked systems, focusing on control and optimization of multi-agent systems and analysis of learning and information dynamics in social and economic networks. Our final addition is Professor Marija Ilic (on leave from Carnegie Mellon University), who joined IDSS as a visiting professor and LIDS as a principal investigator; Ilic brings expertise in modeling and control of large-scale electric power systems. These additions have strengthened our position as a center of excellence in research on networked systems, control and optimization, and inference and learning.

LIDS researchers continue to have great success in obtaining funding for our broad and deep research agenda, and we continue to develop our relationships with industrial organizations and national laboratories including Draper Laboratory, Lincoln Laboratory, the Los Alamos National Laboratory, the Argonne National Laboratory, Siemens, Honeywell, the Ford Motor Company, Aurora Flight Sciences, the Coca-Cola Company, Accenture, and Microsoft Research. Also, thanks to a rich history of research excellence and leadership, LIDS remains a magnet for the very best, attracting not only outstanding students but a continual stream of world-leading researchers as visitors and collaborators.

Intellectual Vision and Research Areas

The domain of research in LIDS can be described along several different dimensions:

- A set of core analytical disciplines, including probability and statistics, dynamical systems, and optimization and decision theory
- A set of core engineering disciplines, including inference and statistical data processing, transmission of information, networks, and systems and control
- A set of broad challenges in traditional and emerging applications of critical importance to society

Research at LIDS involves activities within and across all of these dimensions: advancing the state of the art in analytical and engineering disciplines within information and decision sciences and addressing fundamental challenges that emerge in applications and problems of relevance to societal needs. The simultaneous existence of these different efforts within the same lab leads to strong synergies among them: work in each of the mathematical disciplines leads to new methodologies that enable advances in core disciplines and in interdisciplinary investigations, and work in attacking those emerging interdisciplinary challenges provides direction and drivers for fundamental disciplinary activities and has led to the charting out of emerging new disciplines.

The availability of increasingly capable sensing, communication, and computation systems enables the collection and transfer of large amounts of data pertaining to complex and heterogeneous interconnected systems. The need for an intellectual platform to simultaneously address questions of data fusion, distributed learning, information transfer, and distributed decision making is stronger than ever, as existing techniques fall short in addressing issues of scalability, robustness, and performance limits. Examples of areas in which LIDS research has and will continue to contribute include the following:

- Coordination of unmanned autonomous systems
- Energy services and information systems
- Large-scale data assimilation for the geosciences
- Network scheduling and routing
- Transportation network analysis, control, and design
- Machine learning for recommendation systems and social media
- Social network analysis and characterization
- Ultra-wideband and other emerging communications technologies
- Intelligence, surveillance, and reconnaissance systems
- Biological systems and biomedical data analysis

Furthermore, the recognition that research within traditional boundaries in information and decision sciences is not adequate to address many of the emerging societal challenges that are ahead of us has motivated LIDS researchers to branch out to areas at the intersection of several disciplines. As a result, we have initiated a set of fundamental research themes that cut across disciplinary boundaries and involve considerable interaction and collaboration with colleagues in other units at MIT and in other disciplines:

- Foundations of network science, including network dynamics, control, and efficient algorithms
- Foundational research in game theory and mechanism design involving study of new equilibrium notions and dynamics in games, as well as design of incentives for large-scale networked, dynamic environments that admit efficient computation
- A unifying foundational framework for modeling and understanding systemic risk
- Foundations of cyber-physical systems, including architectural design, security and privacy, cross-layer algorithms, and tools for analysis, verification, and performance guarantees
- Foundational theory for multi-scale/multi-granularity modeling, including methods for describing complex phenomena at multiple granularities, learning of such models from complex and heterogeneous data, and reduction/simplification of models to levels appropriate for particular questions of analysis or design
- Development of scalable and efficient inference algorithms for problems involving "big data" and foundational research on graphical models, providing a general and rich framework for high-dimensional inference

Faculty Activities

Much of the major research activity of LIDS faculty not only cuts across the disciplines, applications, and emerging areas mentioned above but also is collaborative with others within LIDS and elsewhere at MIT.

Dimitri Bertsekas

Professor Bertsekas (McAfee Professor of Electrical Engineering) is interested in deterministic optimization problems and the role of convexity in solving them, possibly through the use of duality. In May 2016, he completed a major revision of his textbook on nonlinear programming. It incorporates some of his recent research on incremental methods for problems whose cost function involves a sum of a large number of component functions, including aggregated versions of incremental proximal methods and augmented Lagrangian methods.

Professor Bertsekas also performs research on problems of sequential decision making under uncertainty, which are pervasive in communication networks, manufacturing systems, logistics, and in the control of nonlinear dynamical systems. In theory, such problems can be addressed with dynamic programming techniques. In practice, only problems with a moderately sized state space can be handled. This research effort deals with the application of neural networks and other approximation and interpolation methodologies to overcome the curse of dimensionality of real-world stochastic control problems. Recent efforts have focused on analysis of properties of value and policy iteration methods and their applications in deterministic optimal control. In addition, extensions of the fundamental concept of regular policies in abstract dynamic programming were developed as a follow-up to a research monograph on the subject published in 2013. A second edition of this monograph, which will incorporate Professor Bertsekas's latest research, is currently being written and is planned for publication in early 2017.

Robert Berwick

Professor Berwick's book with Professor Noam Chomsky, *Why Only Us*, was published by the MIT Press in January and is already in its fourth printing. It has led to several speaking engagements internationally, most prominently before Pope Francis at the Vatican. The pope was presented with an autographed copy of the book.

In addition, Professor Berwick and his students have made progress on how children acquire the particular syntax of their language through "parameter setting"—for example, the fact that in English and French verbs and prepositions most often are at the front of phrases (eat ice cream, with a spoon), while in Japanese the reverse is true. Thus, children must set their language's structure from hearing particular examples. They have designed the (currently) largest set of parameters covering the broadest range of several hundred natural languages and have demonstrated, via a computer program, that it is possible to set these parameters for different languages within a feasible amount of time. They have shown that the parameters implemented, drawn from extensive comparative research, appear to capture the range of syntactic variation in a compact and easily navigable way. In one sense, this should not be surprising:

children do acquire an enormously complex grammatical system in a few short years, and the hypothesis space in which language learning operates must be favorable to learning. At the same time, the results can be regarded as a vindication of the parameter-based theory and its empirical reach: there now is a plausible answer as to what such a favorable hypothesis space looks like.

Guy Bresler

Professor Bresler (Bonnie and Marty Tenenbaum Assistant Professor of Electrical Engineering and Computer Science) works in information theory, statistics, and applied probability. Specifically, his research aims to understand the relationship between the combinatorial structure and computational tractability of high-dimensional inference in graphical models and other statistical models.

With PhD student Mina Karzand, Professor Bresler has been working on the problem of learning graphical models from data with the guarantee that predictions from partial observations are accurate. Prediction accuracy turns out to be captured by a certain local notion of total variation distance. This sort of criterion is natural in many applications of graphical models and stands in contrast to the more commonly studied goal of trying to recover the graph underlying such a model. Results so far address the problem of learning a high-dimensional tree Ising model with a guarantee on local total variation. One of the significant findings from the work is that the necessary number of samples is dramatically lower than the number required to learn the exact underlying graph. On the technical side, this requires reasoning about predictions computed using an incorrect tree and entails showing several new combinatorial properties of tree Markov random fields.

With master's student Sam Park, Professor Bresler has investigated computational relationships between statistical problems. The specific goal is to understand whether fast algorithms for one problem imply fast algorithms for other problems or, conversely, whether computational hardness for one problem implies hardness for another. The focus is on sparse principal component analysis (SPCA) and sparse linear regression (SLR), which are used in a wide range of applications and have attracted a tremendous amount of attention in the last two decades as canonical examples of statistical problems in high dimension. A variety of algorithms have been proposed for both SPCA and SLR over the years. The team has shown how to efficiently transform a blackbox solver for SLR into an algorithm for SPCA. Under the assumption that the SLR solver satisfies prediction error guarantees achieved by existing efficient algorithms such as those based on the lasso, they have shown that the SPCA algorithm derived from the solver achieves state-of-the-art performance, matching guarantees for testing and for support recovery under the spiked covariance model. This reduction not only highlights the inherent similarity between the two problems but, from a practical viewpoint, allows one to obtain a collection of algorithms for SPCA directly from known algorithms for SLR. Experimental results show that these algorithms work well in practice.

Munther Dahleh

Systemic Risk in Finance

Professor Dahleh (IDSS director and William A. Coolidge Professor of Electrical Engineering and Computer Science) is studying fire sales and their negative systemic spillovers, which are the main theoretical rationale to prevent the excessive accumulation of short-term funding. Distressed sales by financial intermediaries to meet immediate liquidity demands often cause deep price dislocations transmitted to other market participants through common exposures of their balance sheets, forcing them to also liquidate assets. The result is a self-reinforcing downward spiral in asset prices that impinges on the entire financial system.

In this work, Dahleh and his team develop a general equilibrium model of investment with financial frictions that links the severity of fire sales with the likelihood of banking runs, identifying a channel that connects measures of systemic risk with macroprudential regulatory policies. This connection is based on a different kind of fire sale externality, labeled the credit risk externality. This externality operates through endogenous credit risk constraints and captures how runs on individual financial institutions can disrupt overall financial stability. It arises because banks do not internalize that, by increasing interest rates, an additional unit of fire-sold assets worsens the terms at which others can raise funding and thus increases their probability of default.

The main normative implication is that constrained efficiency can be restored with a Pigouvian (corrective) tax levied on each unit of short-term debt raised by banks. The optimal tax is decomposed in two terms: one that is proportional to the social shadow value of borrowing short term against collateral and another one that corresponds to the marginal credit risk externality at the social planner's solution. The marginal credit risk externality captures private agents' overvaluation of short-term debt and is equal to the difference between the marginal change in the probability of a crisis weighted by the real economic costs of such an event and the marginal change in the credit risk of a representative bank weighted by the cost of a run on its short-term liabilities. In this respect, the model illustrates how measures of aggregate systemic risk should discount individual contributions to set the tax at the appropriate level.

Finally, this work offers a distinctive account of how better economic times may lead to higher inefficiencies, thus calling for tighter macroprudential regulation. As investment prospects improve in a second-order stochastic dominance sense, the probability of financial crises decreases. However, because the distribution of returns becomes more "concentrated" around its mean, banks grow less concerned about tail risks and the possibility of damaging runs, and hence they increase their reliance on short-term funding. Interestingly, then, even though the incidence of runs is lower, their severity can actually be greater because of more significant market-wide liquidations.

How Networked Agents Make Decisions: Coordination with Local Information

Professor Dahleh, along with his student Spyros Zoumpoulis and Professor John Tsitsiklis, studies decision making by networked entities by examining coordination among strategic agents that share local information. In particular, they study the role of local information channels in enabling coordination among strategic agents. Building on the standard nite-player global games framework, they show that the set of equilibria of a coordination game is highly sensitive to how information is locally shared among agents, specifically that the coordination game has multiple equilibria if there exists a collection of agents (1) that do not share a common signal with any agent outside of that collection and (2) whose information sets form an increasing sequence of nested sets. The characterization thus extends the results on the uniqueness and multiplicity of equilibria in global games beyond the well-known case in which agents have access to purely private or public signals. They then provide a characterization of how the extent of equilibrium multiplicity is determined by the extent to which subsets of agents have access to common information, showing that the size of the set of equilibrium strategies increases with the extent of variability in the size of the subsets of agents observing the same signal. They study the set of equilibria in large coordination games, showing that as the number of agents grows, the game exhibits multiple equilibria if and only if a non-vanishing fraction of the agents have access to the same signal.

Robustness Scaling in Large Networks

We propose new asymptotic notions that capture the robustness of a large interconnected network as the dimension of the network grows. Specifically, we address the question of how network robustness changes with network dimension. Motivated by applications in transportation and economic networks, we study network topologies that are robust to scaling in network dimension and investigate characteristics of networks with poor scaling properties. Furthermore, we analyze how perturbation of network edge links affects our notions of robustness, allowing us to identify critical links that can lead to more robust network design.

Jonathan How

Professor How leads research efforts focused on the control of multiple autonomous agents, with an emphasis on distributed decision making with uncertainty; path planning, activity, and task assignment; mission planning for unmanned aerial vehicles and unmanned ground vehicles; sensor network design; and robust, adaptive, and nonlinear control. Professor How is also the principal investigator for the Aerospace Controls Laboratory. Recent research includes the following.

Truncated Random Measures

This work provides practical representations for a very broad class of models within Bayesian nonparametrics: all completely random measures (CRMs) and all normalized completely random measures. Since a complete infinite-dimensional CRM cannot be used explicitly for computation, sequential representations are often truncated for tractability. We provide truncation error analyses for each type of sequential representation, as well as normalized versions, thereby generalizing and improving upon existing truncation error bounds in the literature. We analyze the computational complexity of the sequential representations, which in conjunction with our error bounds allows us to study which representations are the most efficient. We provide a suite of tools allowing for straightforward representation and analysis of CRMs that have not previously been used in a Bayesian nonparametric context.

Streaming Distributed Variational Inference for Bayesian Nonparametrics

This work provides a methodology for creating streaming distributed inference algorithms for Bayesian nonparametric (BNP) models. In the proposed framework, processing nodes receive a sequence of data mini-batches, compute a variational posterior for each, and make asynchronous streaming updates to a central model. In contrast to previous algorithms, the framework is truly streaming, distributed, asynchronous, learning rate free, and truncation free. The key challenge in developing the framework, arising from the fact that BNP models do not impose an inherent ordering on their components, was finding the correspondence between mini-batch and central BNP posterior components before performing each update. To address this, a combinatorial optimization problem over component correspondences was developed, along with an efficient solution technique. The methodology has been applied to the Dirichlet process mixture model, with experimental results demonstrating its practical scalability and performance.

Real Time Multi-Agent Decision Making

This area of research addresses the problem of multi-agent task allocation in communication-contested environments. The general multi-agent task allocation problem is a well-studied topic. However, the majority of research has focused on environments that utilize fairly rigid communication and information state assumptions. This work merges three major planning paradigms (offline planners, implicit coordination planners, and plan consensus planners) into a single algorithm, the Hybrid Information and Plan Consensus (HIPC) algorithm, capable of performing in environments that do not fit into the rigid requirements of this set of paradigms. In practice, attempting to merge the behavior of these three paradigms changes the planning problem in a major way. The insight of Professor How's work is that these methods can have complementary advantages if care is taken to properly stitch them together. Offline solutions can provide powerful insights into the aspects that can be modeled and can determine how these modeled effects will shape the overall solution. Even if all of the required planning information is not available a priori, offline methods can reduce the size of the search space that online methods must cover. This can involve everything from scheduling actions such as refueling ahead of time to completely shaping objective functions and completely changing the outcomes of the algorithm at run time. Implicit coordination provides a powerful tool for online coordination. Agents can plan in what are effectively locally centralized teams to provide highly coupled allocations. Finally, plan consensus can be used to enforce assignment constraints that may have been violated due to mismatches in situational awareness across the fleet. Combining these algorithms is a non-trivial endeavor because the stitching together requires that each paradigm know exactly what the others provide and what assumptions they are built on. A naïve combination of these paradigms may lead to inefficient use of quadcopters.

Decoupled Multi-Agent Path Planning in Non-convex Domains

Together with students Steven Chen and Mark Cutler, Professor How has developed a multi-agent path planning algorithm based on sequential convex programming (SCP) that finds locally optimal trajectories. Previous work using SCP has efficiently computed motion plans in convex spaces with no static obstacles. In many scenarios where the spaces are non-convex, previous SCP-based algorithms failed to find feasible solutions because the convex approximation of collision constraints leads to the formation of a sequence of infeasible optimization problems. The new work of Professor How and his group addresses this problem by tightening collision constraints incrementally, thus forming a sequence of more relaxed, feasible intermediate optimization problems. They have shown that the proposed algorithm increases the probability of finding feasible trajectories by 33% for teams of more than three vehicles in non-convex environments. Furthermore, they have shown that decoupling the multiagent optimization problem to a number of single-agent optimization problems leads to significant improvement in computational tractability. They developed a decoupled implementation of the proposed algorithm and showed that it runs 14% faster and finds feasible trajectories with higher probability than a decoupled implementation of previous SCP-based algorithms. The algorithm is implementable in real time and has been validated through hardware experiments on a team of quadrotors.

Measurable Augmented Reality for Prototyping Cyber-Physical Systems

Cyber-physical systems (CPSs) refer to engineering platforms that rely on the integration of physical systems with control, computation, and communication technologies. Autonomous vehicles are instances of CPSs that are rapidly growing with applications in many domains. Due to the integration of physical systems with computational sensing, planning, and learning in CPSs, hardware-in-the-loop experiments are an essential step for transitioning from simulations to real-world experiments. We have proposed an architecture for rapid prototyping of CPSs that has been developed in the Aerospace Controls Laboratory. This system, referred to as MAR-CPS (Measurable Augmented Reality for Prototyping Cyber-Physical Systems), includes physical vehicles and sensors, motion capture technology, a projection system, and a communication network. The role of the projection system is to augment a physical laboratory space with autonomous vehicles' beliefs and a simulated mission environment, which in turn will be measured by physical sensors on the vehicles. The main focus of this method is on rapid design of planning, perception, and learning algorithms for autonomous single-agent or multi-agent systems. Moreover, the proposed architecture allows researchers to project a simulated counterpart of outdoor environments in a controlled indoor space, which can be crucial when testing in outdoor environments is disfavored due to safety, regulatory, or monetary concerns. We have demonstrated the real-time capabilities of MAR-CPS in a variety of problems in autonomy, such as motion planning, multi-robot coordination, closed-loop vision-based perception and planning, and learning of spatio-temporal fields.

Patrick Jaillet

The research of Professor Jaillet (Dugald C. Jackson Professor in Electrical Engineering and Computer Science and ORC co-director) has primarily focused on formulating and analyzing online and dynamic versions of classical optimization problems such as the shortest path problem, the traveling salesman problem, the assignment/matching problem, and the matroid secretary problem, as well as some of their generalizations. The research deals with provable results (algorithmic design and analysis) on how to solve such problems under uncertainty, with or without explicit stochastic modeling of uncertainty. Methodological tools include those from online optimization (competitive analysis), stochastic optimization (robust analysis), online learning (min-max regret analysis, Bayesian updates), game theoretic concepts (price of anarchy), and their integrations.

Motivating applications include routing and location problems that arise from transportation and logistics networks, data communication and sensor networks, and/or autonomous multiagent systems, as well as dynamic resource allocation problems in various applications arising from the digital economy (search engines and online auctions), health care (kidney exchange programs), and social interactions (job search and house exchanges).

Professor Jaillet's research group at MIT this past academic year included one postdoc (Le Hoang, EECS/LIDS), eight graduate students from ORC (Maximilien Burq, Arthur Flajolet, Virgile Galle, Chong Yang Goh, Swati Gupta, Nikita Korolko, Sebastien Martin, and Konstantina Mellou), one graduate student from EECS (Dawsen Hwang), and two graduate students from Leaders for Global Operations/EECS (Andrew Johnston and Kfir Yeshayahu). The research group in Singapore included two postdocs from SMART (Jie Chen and Sanjay Jena) and several graduate students (from the National University of Singapore, Nanyang Technological University, and Singapore Management University).

Current funded research programs originate from the Mathematical Optimization and Operations Research Division of the Office of Naval Research ("Online Optimization and Learning under Uncertainty" and "Decentralized Online Optimization in Multi-Agent Systems in Dynamic and Uncertain Environments") and SMART ("Future Mobility," a large project involving several other MIT principal investigators).

Sertac Karaman

Professor Karaman (Class of '48 Career Development Professor), carries out research in the areas of control theory, optimization, formal methods, stochastic processes, and applied probability, with applications to robotics, mobile sensor networks, cyber-physical systems, and data-driven application systems. His current work focuses on joint sensing and control for agile robotic vehicles, high-performance robotic networks, and information-theoretic methods in robot motion.

Autonomous Vehicle Systems for Transportation and Logistics

Autonomous vehicles hold the potential to revolutionize transportation, particularly in urban centers. For example, driverless cars may enable taxi services and drones may enable the next generation of high-throughput, low-delay urban logistics services. This research area aims to better understand the potential impact of autonomous vehicles by studying new autonomous vehicle systems. For instance, one potential technology allows trucks to follow each other closely, much closer than what human drivers can safely achieve. Thanks to this new technology, trucks could form platoons on highways, allowing them to save energy. However, forming long platoons requires some vehicles to wait for others, resulting in excessive transportation delays. In our most recent work, we have developed a queuing model to understand this energy-delay tradeoff. In particular, we have analyzed both open-loop (timetable) policies and feedback policies (that can observe when trucks arrive), and we have shown that the addition of feedback benefits the system very little. These ideas are not limited to truck platooning. They can be applied to ride sharing to reduce transportation costs, formation flight for commercial airlines to save energy, and many other areas.

Computational Methods for Inference, Estimation, and Control

Consider the navigation problem involved with high-performance mobile robotic vehicles, such as rotary-wing aerial vehicles tasked with flying inside a building as fast as possible. This navigation issue requires solving (in real time) a control problem in which the 12-degrees-of-freedom state of the robot must be controlled effectively while the robot is navigating at high speeds in a complex environment. The adverse effect of the robot's non-holonomic differential constraints is amplified at high speeds, limiting the robot's maneuvering capability. The geometric constraints impose hazards that must be reliably avoided. Unfortunately, computing such control laws is extremely challenging. The problem is NP (nondeterministic polynomial-time) hard, with an increasing number of degrees of freedom. Hence, it is unlikely that there exists an efficient algorithm with running time that scales polynomially with increasing degrees of freedom. This phenomenon is often called the curse of dimensionality.

In our recent work, we propose a novel algorithm that runs in polynomial time for all problem instances that admit a certain low-rank structure. The key insight involves compressing a certain cost-to-go function (as in data compression) and working on the compressed version of the function iteratively, in order to improve it toward an optimal one. We rigorously prove that the new compression-based algorithm runs in polynomial time with increasing degrees of freedom. The algorithm also scales polynomially with the "rank" of the value function. Hence, if the value function has a low-rank structure due to, for example, geometry and physics (in other words, if the value function is "compressible"), the new algorithm provides substantial computational savings. Effectively, it breaks the curse of dimensionality for problems that exhibit a low-rank, compressible cost-to-go function. Our recent work applies similar ideas to correct-by-design control, differential games, and agile drone flight, among many other areas.

Sanjoy Mitter

Professor Mitter's research, conducted jointly with Mukul Agarwal PhD '12 and Takashi Tanaka (former LIDS postdoctoral associate), has focused on information theory and stochastic control. The group has proved a source channel separation theorem for a general class of channels. This requires a high-level view of a communication model. A variety of stochastic control problems involving minimization of directed information, subject to distortion constraints, have been considered. Interestingly, these problems can be reduced to semi-definite programming problems.

Professor Mitter participated in a panel discussion, "Systemic Risk in Non-Financial Systems," as part of a conference on financial stability held at the University of Michigan Law School in October 2015.

Eytan Modiano

Professor Modiano (associate head of the Department of Aeronautics and Astronautics) leads the Communications and Networking Research Group (CNRG), which consists of eight graduate students and two postdoctoral researchers. The primary goal of CNRG is the design of architectures for aerospace networks that are cost effective and scalable and meet emerging needs for high-datarate and reliable communications. In recent years the group has focused on robust network designs, wireless networks, and interdependent cyber-physical networks.

In the area of wireless networks, the group has been pursuing a number of avenues, including robust designs, dynamic resource allocations, and fundamental performance limits. Wireless networks are notoriously unreliable due to mobility issues and wireless interference, which leads to time-varying connectivity. Researchers at CNRG have pursued a number of approaches toward making wireless networks robust. This past year, Professor Modiano and graduate student Qingkai Liang developed a new survivability framework for wireless networks based on the notion of time-varying graphs and studied the survivability properties of such graphs. They developed new metrics for graph robustness and algorithms for evaluating these metrics. In related work, Modiano and Liang studied the use of spectrum sharing to improve network robustness. Spectrum sharing enables efficient use of an idle spectrum by allowing secondary users to utilize the spectrum when it is not occupied by primary users. However, such users must relinquish the spectrum when it is needed by its primary owner. This makes it difficult to design a dependable service based on secondary spectrum access, limiting the potential use of such a spectrum. In their work, Modiano and Liang show that high levels of reliability can indeed be achieved using a novel combination of dedicated and shared spectra.

Over the past year, the group has been studying the problem of efficient information dissemination in wireless networks. In particular, they studied the problem of wireless broadcast and the network functionality of delivering data from a source node to all other nodes in a network. Broadcast is important for mission-critical military communications, live video streaming, and data dissemination in sensor networks. Despite its importance, however, the problem of throughput optimal broadcast in wireless networks has remained unsolved. This past year, Professor Modiano and his student Abhishek Sinha developed a throughput optimal algorithm for broadcasting in wireless networks that form a directed acyclic graph (DAG). In addition, they proposed an extension of the algorithm that can effectively be used for broadcasting in any network with arbitrary topology.

The group continued work on its project addressing the robustness of interdependent networks. Many engineering systems involve interactions between two or more networked systems. Cyber-physical systems, for example, consist of networked computer systems that are used to control physical systems such as power grids, water or gas distribution systems, and transportation networks. While this cyber-physical interaction is critical for the functionality of the overall system, it also introduces vulnerabilities in the form of interdependence failure cascades, with failures in the cyber network leading to failures in the physical network and vice versa. Over the past year, Professor Modiano and his student Marzieh Parandehgheibi PhD '16 studied the interdependence between the power grid and the communication network used to control the grid. In particular, they studied the problem of distributed control under communication constraints and showed that distributed control algorithms fail when communication links fail. Also, they developed new control schemes that achieve optimal performance, even when communication links fail, by leveraging power flow information in making control decisions.

In recent years, Professor Modiano and his group have been pursuing industrial collaborations to increase the impact of their work on practical systems. To that end, the group is conducting a joint project with BBN Technologies on resilient overlay networks and a project with Qualcomm on mission-critical communications. Also, the group is involved in a close collaboration with researchers at Lincoln Laboratory on the design of network architectures and protocols for military communications.

CNRG's research crosses disciplinary boundaries by combining techniques from network optimization, queuing theory, graph theory, network protocols and algorithms, hardware design, and physical layer communications.

Asuman Ozdaglar

Professor Ozdaglar (director of LIDS, associate director of IDSS, and Keithley Professor of Electrical Engineering and Computer Science) and her research group focus on modeling, analysis, and optimization of large-scale, dynamic multi-agent networked systems. Their research draws on advances in game theory, optimization theory, dynamical systems, and stochastic network analysis. It focuses on both investigating substantive problems in these areas and developing new mathematical tools and algorithms for the analysis of these systems.

A major current research area in Professor Ozdaglar's group is socio-technological networks, which include both infrastructure systems (communication networks, traffic networks, and power grids) and social, economic, and financial systems. The operation and failures of such networks depend not only on their technological constraints but also on the human element—the intentions, information, and strategic behavior of the users and participants. Professor Ozdaglar studies the interplay of the human element and technological constraints in the context of networked systems.

Her recent work focuses on design of incentives in infrastructure networks, such as transportation and power networks, to improve the efficiency of the equilibrium emerging from strategic interactions of users. The increasing availability and use of real-time GPS-based traffic apps, such as Waze and Google Maps, have already begun to transform the daily commutes of millions of drivers. Although these technologies are hailed as veritable breakthroughs and harbingers of improved traffic experiences for all users, their implications for equilibrium traffic flows have not yet been studied. In recent work conducted jointly with graduate student Ali Makhdoumi, Professor Daron Acemoglu from the Department of Economics, and Azarakhsh Malekian (former LIDS postdoc and now an assistant professor at the University of Toronto), Professor Ozdaglar has developed a framework for systematically analyzing how changes in the information sets of users in a traffic network (e.g., due to route guidance systems) affect traffic equilibrium. This work builds on a classic traffic flow model relying on the notion of Wardrop equilibrium. They first generalized their framework to allow for different types of users with different sets of information about available network links and then showed how this expansion of information can increase travel times.

In another project, conducted jointly with postdoc Insoon Yang, Professor Ozdaglar studied efficient allocation of power in electricity markets. Volatility in wholesale electricity market prices is a significant concern for suppliers and generation companies, which cannot forecast their revenues (resulting in decreasing investments and innovations in electricity generation), as well as risk-averse consumers. In this project, they developed an alternative approach to alleviating risk concerns in electricity markets by using storage as an instrument to absorb real-time wholesale price fluctuations. They proposed an optimal storage charging/discharging control method to penalize price volatility while further reducing the total generation costs of economic dispatch. The proposed stochastic dynamic programming approach optimally shifts energy over time and space (via transmission lines) by fully utilizing the distributional information of net demand. A distinctive feature of the proposed approach is that energy storage does not have a direct role in wholesale markets but only transparently affects the market parameters in real-time economic dispatch.

In another collaboration with Insoon Yang and Daron Acemoglu, Professor Ozdaglar is studying the policy implications of adopting clean energy technologies in the energy-producing sector. Their model generalizes an industry equilibrium model of entry, exit, and firm dynamics proposed in the economics literature by adding a technology upgrade component. Given a time-varying carbon tax sequence, they first characterize the nonstationary dynamic equilibrium. They show that the optimal upgrading, entry, and exit strategies are characterized as thresholding rules on the firm’s generation efficiency. They then determine a time-varying carbon tax that maximizes the utility of consumers subject to the equilibrium, which leads to a bilevel dynamic optimization problem. Their model will provide qualitative insight on the role of policy instruments in sustainable development of clean generating technologies to reduce greenhouse gases as well as an optimal strategy for firms to upgrade technologies and enter and exit the energy-producing sector. This work is supported by a National Science Foundation (NSF) Cyber-Physical Systems grant (with Berkeley, Michigan, and Vanderbilt) and a joint Masdar Institute/MIT project.

In other recent work, the group considers how privacy concerns affect individual choices in the context of a network formation game (where links can be interpreted as friendships in a social network, connections over a social media platform, or trading activity on an online platform). In the model, each individual decides which other agents to "befriend" (i.e., form links with). Such links bring direct (heterogeneous) benefits from friendship and also lead to the sharing of information. But such information can travel over the equilibrium network via other linkages (e.g., the party acquiring the information shares it with others) through a percolation process. Privacy concerns are modeled as a disutility that individuals suffer as a result of their private information being acquired by others. The team specifies conditions under which pure-strategy equilibria exist and characterize both pure-strategy and mixed-strategy equilibria. The findings show that, as in many real-life examples, the resulting equilibrium networks feature clustered connections and homophily. Clustering emerges because if player A is a friend with B and B is a friend with C, then A's information is likely to be shared indirectly with C anyway, thus making it less costly for A to befriend C. Homophily emerges because even an infinitesimal advantage in terms of direct benefits of friendship within a group makes linkages within that group more likely, and travel of information within that group reduces the costs (and thus increases the likelihood) of further within-group links. This project is supported by an Army Research Office joint multidisciplinary university research initiative (with the University of Pennsylvania, Stanford, and Cornell) and Draper Laboratory's directed research and development program.

Professor Ozdaglar's group also works on developing novel algorithms for processing large-scale data and distributed optimization over networks. In recent work, motivated by machine learning applications, they focused on additive convex cost optimization problems, wherein the objective function is the sum of m convex component functions. In many applications, the number of component functions m is large (because the number of data points or the number of agents making up the system is large), and therefore solving this problem using traditional gradient methods is impractical. Incremental algorithms, which process component functions sequentially one at a time, can be used as an alternative. Recent work conducted jointly with postdoctoral researcher Mert Gurbuzbalaban and Professor Pablo Parrilo has made significant contributions to this area, filling in important gaps in theory and leading to improved algorithms. In one paper, the team focused on deterministic cyclic incremental gradient and incremental Newton methods and considered the case in which the component functions are smooth and the sum function is strongly convex. Under these assumptions, they showed faster convergence rates than the currently available results in the literature.

In another paper, the team addressed a long-standing issue in the analysis of randomized versions of incremental methods, the random reshuffling (RR) algorithm. RR is a randomized first-order incremental algorithm that selects a uniformly random order/permutation and processes the component functions one at a time according to this order. Although RR has been numerically observed to outperform its with-replacement counterpart, stochastic gradient descent (SGD), characterization of its convergence rate is notoriously difficult and has been a long-standing open question.

The group's recent work answered this question by showing that RR with iterate averaging and a diminishing step size αk = Θ(1/ks) for s is an element in the set of (1/2,1) converges at rate Θ(1/k2s) with probability one in the suboptimality of the objective value, thus improving upon the Ω(1/k) rate of SGD.

In other research, the group is studying the incremental aggregated gradient (IAG) method, which processes a single component function at a time (as with incremental methods) but keeps a memory of the most recent gradients of all component functions so that an approximate gradient descent direction of the sum function can be constructed at each iteration. They proved that a general form of deterministic IAG (wherein the component function sampling times can be arbitrary as long as they occur within a bounded interval) converges linearly for the general case with strongly convex component functions. This research effort is supported by a joint Office of Naval Research project with the University of Pennsylvania.

Pablo A. Parrilo

Professor Parrilo (associate director of LIDS) and his research group are focused on mathematical optimization, systems theory, and control, with an emphasis on development and application of computational tools based on convex optimization and algorithmic algebra.

Professor Parrilo and PhD student Diego Cifuentes are investigating how structural combinatorial parameters such as treewidth can help efficiently solve basic questions in computational algebraic geometry (e.g., solving systems of polynomial equations). In recent work, they introduced chordal networks, a novel representation of structured polynomial ideals. The sparsity structure of a polynomial ideal is often described through a graph that captures the interactions among variables. Chordal networks provide a computationally convenient decomposition of the polynomial ideal into simpler (triangular) polynomial sets while preserving the underlying graphical structure. They showed that many interesting families of polynomial ideals involve compact chordal network representations, even though the number of components might be exponentially large. Chordal networks can be computed for arbitrary polynomial systems using a refinement of an earlier chordal elimination algorithm. They applied their methods to examples from algebraic statistics and vector addition systems and showed that algorithms based on chordal networks can outperform existing techniques by orders of magnitude.

In collaboration with PhD students Hamza Fawzi and James Saunderson, the group developed a general methodology to study lifts of convex sets and, in particular, concise certificates of the nonnegativity of a function on a finite abelian group. The setup focuses on nonnegative functions on a finite abelian group G that are sparse with respect to the Fourier basis. They established combinatorial conditions on subsets S and T of Fourier basis elements under which nonnegative functions with Fourier support S are sums of squares of functions with Fourier support T. Their combinatorial condition involves constructing a chordal cover of a Cayley graph related to G and S, with maximal cliques related to T. This result relies on two main ingredients: the decomposition of sparse positive semidefinite matrices with a chordal sparsity pattern and a simple but key observation exploiting the structure of the Fourier basis elements of G. This methodology has very interesting consequences. In the case of the Boolean hypercube, by constructing a particular chordal cover of the half-cube graph, they proved that any nonnegative quadratic form in n binary variables is a sum of squares of functions of degree at most [n/2], establishing a conjecture of Laurent. For the case of nonnegative functions of degree d on ZN (the discrete cycle), by constructing a particular chordal cover of the *d*th power of the N cycle, they proved that any such function is a sum of squares of functions with at most 3dlog(N/d) nonzero Fourier coefficients. This shows that a certain cyclic polytope in R^{2d} with N vertices can be expressed as a projection of a section of the cone of psd matrices of size 3dlog(N/d). This construction provides the first explicit family of polytopes in increasing dimensions wherein the semidefinite extension complexity is asymptotically smaller than the linear programming extension complexity.

Professor Parrilo and PhD student Frank Permenter have developed a new dimension reduction method for semidefinite programming through the use of Jordan algebras. Specifically, they consider orthogonal projections satisfying certain invariance conditions (the optimal set should be contained in its range). Adding this constraint to the original optimization problem yields an equivalent primal-dual pair over a lower-dimensional symmetric cone—namely, the cone of squares of a Jordan subalgebra of symmetric matrices. They also propose a simple lattice-theoretic algorithm for minimizing the rank of this projection and, hence, the dimension of the cone. Through the theory of Jordan algebras, the proposed method easily extends to linear programming, second-order cone programming, and, more generally, symmetric cone optimization. The method has been shown to be very effective with several examples from the literature.

Yury Polyanskiy

Associate Professor Polyanskiy conducts research in the areas of mathematics of information (information theory), coding theory, and the theory of random processes. His current work focuses on non-asymptotic characterization of the performance limits of communication systems, non-Shannon information measures, redundant circuits, and probabilistic methods in combinatorics.

Bounds for Codes on Pentagon and Other Cycles

The capacity of a graph is defined as the rate of exponential growth of independent sets in the strong powers of the graph. In strong power, an edge connects two sequences if at each position letters are equal or adjacent. We considered a variation of the problem wherein power graph edges are removed among sequences differing by more than a fraction δ of coordinates. For odd cycles, we derived an upper bound on the corresponding rate that combined Lovsz's bound on the capacity with Delsarte's linear programming bounds on the minimum distance of codes in Hamming spaces. In the case of the pentagon, this showed that for $\delta \geq 1 - 1/\sqrt{5}$, the Lovsz rate was the best possible option, while we proved using a Gilbert-Varshamov-type bound that a higher rate was achievable for $\delta < 2/5$.

Wasserstein Continuity of Entropy and Outer Bounds for Interference Channels

This work shows that under suitable regularity conditions, differential entropy is -Lipschitz as a function of probability distributions on with respect to the quadratic Wasserstein distance. Under similar conditions, (discrete) Shannon entropy is shown to be $O(n)$-Lipschitz in distributions over the product space with respect to Ornstein's -distance (Wasserstein distance corresponding to the Hamming distance). These results, together with Talagrand's and Marton's transportation-information inequalities, allow one to replace unknown multi-user interference with its independent and identically distributed approximations. As an application, a new outer bound for the two-user Gaussian interference channel was proved that solves Costa's "missing corner point" problem.

Short-Packet Communications with Multiple Antennas

We described finite-blocklength upper and lower bounds on the maximum coding rate achievable over a multiple-antenna Rayleigh block-fading channel under the assumption that neither the transmitter nor the receiver has a priori knowledge of the channel realizations. Numerical evidence suggests that the bounds delimit tightly the maximum coding rate for short blocklengths. The bounds allowed us to estimate the number of transmit antennas and the degree of time-frequency diversity that optimally trade off the rate gains resulting from an increase in the number of independent time-frequency/spatial diversity branches against the corresponding increase in channel estimation.

Optimum Power Control at Finite Blocklengths

This research investigates the maximal channel coding rate achievable at a given blocklength n and error probability when the codewords are subject to a long-term (i.e., averaged-over-all-codeword) power constraint. The second-order term in the large-n expansion of the maximal channel coding rate is characterized both for additive white Gaussian noise (AWGN) channels and for quasi-static fading channels with perfect channel state information available at both the transmitter and the receiver. It is shown that, in both cases, the second-order term is proportional to . For the quasi-static fading case, this second-order term is achieved through truncated channel inversion, namely through concatenating a dispersion-optimal code for an AWGN channel subject to a short-term power constraint with a power controller that inverts the channel whenever the fading gain is above a certain threshold. Easy-to-evaluate approximations of the maximal channel coding rate have been developed for both the AWGN and the quasi-static fading case.

Dissipation of Information in Channels with Input Constraints

One of the basic tenets of information theory, the data processing inequality refers to instances in which output divergence does not exceed input divergence for any channel. For channels without input constraints, there are various estimates of the amount of such contraction, Dobrushin's coefficient for total variation being perhaps the most well known. Professor Polyanskiy's work in this area investigates channels with average input cost constraints. It is found that while the contraction coefficient typically equals one (no contraction), the information nevertheless dissipates. A nonlinear function, the Dobrushin curve of the channel, is proposed to quantify the amount of dissipation. Tools for evaluating the Dobrushin curve of additive-noise channels have been developed based on coupling arguments. Some basic applications in stochastic control, uniqueness of Gibbs measures, and fundamental limits of noisy circuits have been discussed.

As an application, it is shown that in the chain of n power-constrained relays and Gaussian channels, the end-to-end mutual information and maximal squared correlation decay as log log n/log n), which is in stark contrast with the exponential decay in chains of discrete channels. Similarly, the behavior of noisy circuits and the broadcasting of information on trees are not at threshold in the signal-to-noise ratio (SNR). Namely, unlike the case of discrete channels, the probability of bit error remains bounded away from 1/2 regardless of the SNR.

Variable-Length Compression Allowing Errors

This work investigates the fundamental limits of the minimum average length of lossless and lossy variable-length compression, allowing a nonzero error probability for lossless compression. We provide non-asymptotic bounds on minimum average length in terms of Erokhin's rate-distortion function, and we use those bounds to obtain a Gaussian approximation on the speed of approach to the limit that is quite accurate for all but small blocklengths. A nonzero error probability not only reduces the asymptotically achievable rate by a factor of , but this asymptotic limit is approached from below (i.e., a larger source dispersion and shorter blocklengths are beneficial). Variable-length lossy compression under excess distortion constraint is shown to exhibit similar properties.

Mardavij Roozbehani

Principal research scientist Mardavij Roozbehani led several research efforts related to mathematical modeling, optimization, and control for networked systems, with a particular emphasis on robustness analysis and architecture design for energy systems. The long-term goal of his research is to understand the sources of robustness, fragility, and systemic risk in energy, financial, and transportation networks. Funding for his research comes from NSF, the MIT-Masdar Institute Cooperative Program, and the MIT-Skoltech Initiative.

Systemic Risk, Robustness and Fragility in Networks

In collaboration with graduate students Georgia Katsargyri and Tuhin Sarkar, Dr. Roozbehani researched several problems related to systemic risk in financial networks, propagation of shocks, and distributed estimation of rare events using abstract models of networks. In financial networks, the role of financial instruments and hedging in patterns of interconnection and systemic risk was studied. This work resulted in new insights on how hedging can sometimes increase risk for the system as a whole while it reduces individual risk. The findings of this research have been disseminated in peer-reviewed publications as well as Katsargyri's PhD thesis.

Dr. Roozbehani's work with Tuhin Sarkar resulted in a unifying framework for analysis and quantification of resilience and fragility in networks across several application domains (transportation, economics, and social). This work extends several results that existed in the literature, albeit sparsely and scattered across different fields.

Dr. Roozbehani's work with postdoctoral associate Yasin Yazicioglu on robustness of transportation networks reveals new sources of fragility in these networks resulting from limited information exchange and decision making based on local observations. This research has led to novel and unexpected results wherein locally optimal decisions or local infrastructure improvements have negative consequences and degrade global performance. Methodologies for identifying such fragile links and bottlenecks were introduced, and the findings have been reported. They hope to extend and generalize these results in more abstract settings to better understand the effects of local decision making on global performance in general networks.

Additional work includes efforts focusing on electricity market design for renewables (with MIT student Ian Schneider), methodologies for aggregation of distributed demand response resources and creation of virtual storage (with MIT postdoctoral associate Daria Madjidian), demand response estimation with limited information from aggregate data (with Donatello Materassi from the University of Tennessee), and the stability of linear switched systems (with Raphael Jungers from the Université catholique de Louvain).

Devavrat Shah

Professor Shah and his research group are involved in designing practical algorithmic solutions with a theoretical understanding of problems arising in large networks and large-scale statistical inference. This interdisciplinary research builds on advances in applied probability, stochastic networks, information theory, artificial intelligence, and algorithms. The primary application areas for these efforts are communication networks such as Internet routers, wireless networks, and data centers; large-scale statistical problems such as those arising in designing algorithms for inference in graphical models; social data processing such as that in recommendation systems; and data-driven algorithm design for applications in operations management. This work spans the areas of computer science, electrical engineering, operations research, and statistics.

Non-parametric Framework for Recommendation Systems

Recommendation systems have become essential in many modern data processing systems wherein the goal is to help users "discover" what they are looking for even when a priori they may not precisely know what they are looking for. Classical applications include recommending products on Amazon, media shows on Netflix and YouTube, and songs on Pandora or Spotify. In modern applications, the task of recommendation extends way beyond these settings, for example matching buyers and sellers on a social marketplace or suggesting content on content-hosting sites.

Despite such importance, collectively very little progress has been made toward understanding the correct framework in which to design algorithms for solving the generic recommendation problem. Over the past year, Professor Shah and his group have proposed a novel non-parametric framework to address this intellectual shortcoming. In the process, using this framework, they have managed to explain why the algorithms (or heuristics) widely employed in practice are successful. More generally, this has led to better algorithms for solving such problems that extend to modern settings.

Suvrit Sra

Principal research scientist Suvrit Sra has been with LIDS since January 2015. His research interests span two broad categories: machine learning, statistics, and optimization and pure and applied mathematics.

Machine Learning, Statistics, and Optimization

Dr. Sra's primary research focuses on large-scale optimization for problems in machine learning, statistics, and related areas. Over the past year, with his collaborators, he has continued to expand the collection of fast optimization algorithms ("training procedures" in machine learning parlance) that use theoretical analysis to guide their implementation. In particular, his work has addressed theory and methods for large-scale nonconvex optimization problems, the kinds of problems that underlie the massively popular subarea of machine learning called "deep learning." His work in this area, for the first time, provides an analysis of when and how fast we can expect certain large-scale optimization algorithms to reach a solution. These results are currently being explored and have already garnered interest from industry given the tremendous importance of "big-data" nonconvex optimization problems.

Of greater relevance to statistics, Dr. Sra's continued work in "geometric optimization" (i.e., optimization of cost functions by using geometry, for example differential geometry, to guide the design and analysis of algorithms) has led to progress on a classical and fundamentally important problem in statistics: fitting a mixture of Gaussians. For this problem, the 40-plus-year-old solution is to use the so-called expectation maximization algorithm, a method thought to be unsurpassable using techniques from nonlinear optimization. Surprisingly, Dr. Sra (with his collaborator) used ideas from differential geometric optimization to reformulate the mixture fitting problem and to obtain a solution often substantially faster than expectation maximization. This result marks the beginning of a long stream of work in geometric optimization that Dr. Sra is pursuing (together with his students and other collaborators).

Pure and Applied Mathematics

Some of the results of Dr. Sra's work on optimization algorithms have turned out to be of broad interest to the mathematical community, including new algorithms for problems in linear algebra (e.g., matrix analysis). The most broadly applicable of these topics is his work on "determinantal point processes," which have ramifications throughout statistical physics, combinatorics, probability theory, and even machine learning. In collaboration with students and other faculty members, Dr. Sra has been working on obtaining algorithms that allow scientists to work with these important models of probability and help scale them up to become relevant for machine learning applications. For example, he recently collaborated on provably (and empirically) fast sampling from determinantal processes, a result missing in the literature despite the study of these processes dating back to the 1970s.

John Tsitsiklis

Professor Tsitsiklis (Clarence J. LeBel Professor of Electrical Engineering) and his students work on system modeling, analysis, optimization, and control in possibly stochastic and dynamic environments and in the possible presence of multiple agents with conflicting interests. Their research activities have focused on developing methodologies, mathematical tools of broad applicability, and computational methods. Motivating applications for recent work have come from domains such as computer networks and social networks.

Resource-Constrained Computer Networks

This research concerns analysis and optimization of resource allocation methods (routing and scheduling) in queuing networks. During the past year, we identified a class of policies that delivered better delay performance than previous alternatives while utilizing comparable amounts of messaging and memory. More interestingly, we have shown some fundamental impossibility findings: if the message rate and the available memory are below certain levels, no policy can deliver results in negligible delays. On the technical side, this result is unlike other findings in this field.

Social Networks

Following up on earlier work on closed-loop control of epidemics and contagion processes, we are now investigating closed-loop control problems associated with other types of network phenomena such as the voter model.

Education

Professor Tsitsiklis oversaw a third offering, through EdX, of the online 6.041x Introduction to Probability undergraduate class. Furthermore, he continued to integrate the available online materials into the corresponding residential class.

Caroline Uhler

Henry L. and Grace Doherty Assistant Professor Uhler joined LIDS in October 2015. She carries out research in the areas of mathematical statistics and optimization with applications to genomics and cell biology. Her current work focuses on parameter estimation in Gaussian models with linear constraints, total positivity and its relation to conditional independence, development of methods for causal inference, and development of geometric models of the spatial arrangement of chromosomes in the cell nucleus.

Maximum Likelihood Estimation for Gaussian Models

In various statistical applications, the covariance matrix carries a certain structure and has to satisfy certain constraints. This work investigates parameter estimation for Gaussian models with linear constraints on the covariance matrix. Examples of such models are correlation matrices and Brownian motion tree models, which are standardly used to reconstruct phylogenetic trees from continuous characters. Maximum likelihood estimation for these models leads to a non-convex optimization problem that typically has many local maxima. This considerably complicates inference for such models, and various heuristics have been employed for this task. However, the likelihood function has a good structure: it is concave over a large convex region of the positive definite cone. Using recent results on the asymptotic distribution of extreme eigenvalues of the Wishart distribution, we proved that the global maximum and the least squares estimator reside in this region with high probability. Hence, any hill-climbing method initiated in the least squares estimator will remain inside the convex region and converge to the global maximum. A direct consequence of this result is a simple and efficient method with provable guarantees for maximum likelihood estimation in Brownian motion models when the underlying tree structure is known. Similar findings might explain why algorithms for deep learning perform surprisingly well and often converge to a point that is sufficiently close to the global optimum. This work was conducted jointly with Donald Richards and Piotr Zwiernik.

Total Positivity in Markov Structures

This work focuses on properties of distributions that are multivariate totally positive of order two (MTP_2). MTP_2, introduced in the 1970s, is one of the strongest forms of positive dependence. This work builds on two important properties of MTP_2 distributions, namely that they are closed with respect to marginalization as well as with respect to conditioning. We described various conditional independence implications for MTP_2 distributions; for example, any conditional independence relation remains valid when increasing the conditioning set. We also analyzed factorization properties of MTP_2 distributions and discussed ways of constructing such distributions. These results suggest the need for studying graphical models under MTP_2 constraints, since their intriguing properties with respect to conditional independence implications could imply model sparsity without the requirement of a tuning parameter. This project was performed in collaboration with Shaun Fallat, Kayvan Sadeghi, Steffen Lauritzen, Nanny Wermuth, and Piotr Zwiernik.

Generalized Permutohedra for Causal Inference

Causal inference is a cornerstone of scientific discovery because it seeks to determine causative relationships between variables. It is of particular interest to determine causal structures among variables based on observational data, since conducting randomized controlled trials is often impractical or prohibitively expensive. Directed graphical models have been used extensively to model causal phenomena. It is desirable to develop methodologies to learn directed graphical models from observational data. Although the faithfulness assumption is a common assumption that allows for efficient algorithms, it is extremely restrictive. A method that does not require this assumption is the sparsest permutation algorithm, which seeks to find the permutation of variables that explains the data with the least number of edges. This algorithm requires searching through all permutations of variables, which is infeasible for large problems. To overcome this computational challenge, we introduced the DAG associahedron, a subpolytope of the permutohedron wherein the vertices represent directed graphs. A simplex-type algorithm for this polytope can be used to determine the sparsest permutation and output the corresponding directed graphical model. This work was conducted jointly with former LIDS postdoc Fatemeh Mohammadi, Charles Wang, and Josephine Yu.)

Geometric Models of Chromosome Organization

Although the genetic information in each cell within an organism is identical, gene expression varies widely between different cell types. An emerging hypothesis is that this is caused by different constraints on cell shape that induce different spatial arrangements of DNA, in turn inducing different co-regulation patterns. In collaboration with G.V. Shivashankar's lab at the National University of Singapore (and with Yejun Wang and Mallika Nagarajan), we probed this hypothesis using a combination of experiments (by forcing cells into specific geometries, imaging the chromosomes, and taking whole-genome microarray measurements) and modeling. We viewed the spatial organization of chromosomes under different geometric constraints on the cell nucleus as a shape-packing problem; we were interested in packing approximately ellipsoidal-shaped non-uniform domains, the chromosomes, into a container, the cell nucleus. By solving a constrained optimization problem, we determined minimal overlap arrangements of the ellipsoids (i.e., the chromosomes). Using this model, we were able to predict the new chromosome neighborhoods that were formed when the shape of the cell nucleus was changed. Furthermore, we found that the spatial neighborhoods of chromosomes are formed in a transcription-dependent manner. These results suggest that cell geometry modulates 3D chromosome arrangement and gene expression patterns in a predictable way. This is central to understanding geometric control of genetic programs involved in cellular homeostasis and associated diseases.

Kalyan Veeramachaneni

Dr. Veeramachaneni joined LIDS this year as a principal research scientist. His research group develops automation technologies for data science, a burgeoning field that focuses on deriving insights from the huge amount of information produced by contemporary systems. To enable an unsupervised computer program to perform the same tasks as a data scientist, Dr. Veeramachaneni and his team spent years observing how data scientists derive insights from data in different domains, from education to medicine. They are now attempting to build algorithms that emulate these human thought processes. While this level of automation is necessary to keep up with the ever-increasing demand for data scientists, the group is also pursuing a concurrent effort, the "human-data interaction project." The goal of this project is to allow humans to interact seamlessly with these technologies, thus enhancing what each can achieve individually.

To explore data science automation at the highest possible level, Dr. Veeramachaneni and his student James Max Kanter developed an end-to-end system called the Data Science Machine (DSM). The DSM can transform raw data into a predictive model by generating variables, selecting a modeling method, and auto-tuning the model's hyperparameters.

Currently, one of the biggest bottlenecks inhibiting the development of data-driven solutions is feature engineering, a process that includes ideating, writing software, and extracting variables from data. This step is typically human driven and takes up the majority of data scientists' time. In their paper "Deep Feature Synthesis: Towards Automating Data Science Endeavors," published in the proceedings of the annual Institute of Electrical and Electronics Engineers (IEEE)/Association for Computing Machinery (ACM) International Conference on Data Science and Advanced Analytics, Veeramachaneni and Kanter presented an algorithm called deep feature synthesis that achieves these three steps automatically. The algorithm follows relationships in a data set back to a base field and then creates a final variable by sequentially applying mathematical functions to those data slices that pertain to an entity in the data along that path. In this way, the algorithm is able to automatically generate variables for relational data sets.

The next step involves transforming these features into a predictive model, a process known as machine learning. Many stages of the machine learning pipeline have hyperparameters that can be tuned, and this tuning may have a noticeable impact on a model's performance. As an example, in the pipeline being considered for one particular problem, a naïve grid search including all possible parameter value combinations would have required going through more than 2 trillion alternatives. The team uses a Gaussian copula process (GCP) to aid in the exploration of this space. GCP models the relationship f between particular hyperparameter choices and the predictive performance of the model generated by executing the entire pipeline with those choices. The team then samples new hyperparameters and uses f to predict how they would change the model's performance. Finally, they apply selection strategies to choose which hyperparameters will best create a better pipeline. Simply put, this process uses machine learning to improve the development of machine learning.

To demonstrate the efficacy of their system, the team entered the DSM in publicly held data science competitions in which hundreds of data scientists tackled the same problems. Over the course of three different competitions, the DSM went up against 906 other teams and beat 615 of them. In two of the three competitions, the automated approach beat a majority of competitors, and in the third it achieved 94% of the best competitor's score. In its best showing, the DSM beat 85.6% of the human teams and achieved 95.7% of the top submission's score.

Another major focus of Dr. Veeramachaneni's research this year involved developing an active learning-based solution for common cybersecurity problems. Security analysts tasked with monitoring cyber activity and detecting threats are often overwhelmed by the sheer amount of data involved, as well as by the number of alarms produced by unsupervised outlier detection systems and the constant evolution of the patterns in data that correspond to cyber attacks. Together with Dr. Igancio Arnaldo, a former postdoc at MIT, and several colleagues from the artificial intelligence company PatternEx, Dr. Veeramachaneni and his team developed an analyst-in-the-loop security system that they presented at the IEEE Big Data Security Conference this year. In this system, three different unsupervised learning methods each detect rare and anomalous events, and the ones they all agree on are combined into a smaller subset. Analysts investigate this subset and determine whether or not each event is an attack, a process called labeling. Using this new information, they then build supervised learning models that can predict these known attacks for the next day's data and once again show the combined predictions to the analyst for labeling. This process occurs daily, repeatedly refining and executing the models in an adaptive fashion. The group validated this system with a real-world data set consisting of 3.6 billion log lines. When the multiple components are run in conjunction on a daily basis and the results are compared with those of an unsupervised outlier detection method, the detection rate improves by an average of 3.41x and false positives are reduced fivefold.

Alan Willsky

Edwin Sibley Webster Professor (retired) Alan Willsky continues to lead the Stochastic Systems Group, whose research focuses on developing statistical algorithms and methodologies for complex problems of information extraction and analysis from signals, images, and other sources of data. The work extends from basic mathematical theory to specific areas of application. Recent funding for this research has come from the Air Force Office of Scientific Research and the Office of Naval Research.

Our most recent research results have involved three components, two of which are aimed at learning statistical models for phenomena with graphical structures that can be exploited for efficient computation and inference. In the past, we have had considerable success in developing modeling and inference methods for phenomena well modeled by placing the variables representing a given phenomenon on nodes of a tree. This has included results on learning models with hidden or latent variables. Our previous results required placing one of two types of constraints on either the tree structure or the dimensionality of the latent variables. That is, we developed methods that could learn models on specified trees in which the dimensions of the latent variables were learned (as well as the statistical relationships among these variables and the other variables in the tree). We also developed models that could learn tree structure, including identifying latent nodes, but in the case of these methods the latent variables had to be scalar. In our most recent work, we have taken advantage of advances in convex and semi-definite programming to develop an optimization-based approach to what appears to be a very ill-posed problem, namely learning both the structure of the tree and the dimensions of the latent variables.

The second part of our work on learning models has focused on a richer set of graphs, specifically those with small feedback vertex sets (i.e., a set of nodes that, when removed, results in a tree). In previous work, we had shown how statistical models of such graphs involve very efficient inference computations. In our most recent work, we have developed methods for learning models of this structure in two cases: the case in which the variables in the feedback vertex set are observed data and the case in which latent variables are introduced to form the feedback vertex set. The algorithms we have developed have very appealing and computationally efficient structures.

Finally, in the case of statistical models that fall on more general graphs, inference computations can be quite complex. For example, one important computation required in many applications is generation of samples from the overall model. Generating such samples is challenging with models on graphs that have many nodes. One approach to doing so is Gibbs sampling, in which individual variables are updated conditional on the current values of variables at neighboring nodes in the graph. While Gibbs sampling is guaranteed to converge to produce samples from the overall model, convergence can be slow, as strict adherence to the sampling protocol requires that computations in conflict be “blocked.” Some rather remarkable empirical results have been reported showing impressive speed ups if these blocking requirements are ignored. While there is anecdotal evidence in the literature that such methods can work, what is clear is that the resulting samples will not derive from the exact joint distribution of all of the variables in the graph. In addition, there have been scant analyses of when such methods will converge and how close the samples are to the desired exact samples. In our work, we provided a thorough analysis of these questions for Gaussian models on regular 2D lattice graphs (i.e., so-called 2D Markov random fields).

Moe Win

The Wireless Communication and Network Sciences Laboratory, led by Professor Win, is involved in multidisciplinary research that encompasses developing fundamental theories, designing cooperative algorithms, and conducting network experiments for a broad range of real-world problems.

To advocate outreach and diversity, the group is committed to attracting graduate and undergraduate students from underrepresented and minority groups and to giving them exposure to theoretical and experimental research at all levels. The group has a strong track record for hosting students from both the Undergraduate Research Opportunities Program (UROP) and the MIT Summer Research Program. The group hosted a visiting female graduate student last year and two undergraduate students as summer research interns. Professor Win maintains dynamic collaborations and partnerships with academia and industry, including the University of Southern California; the University of California, Santa Barbara; Arizona State University; the University of Bologna and the University of Ferrara in Italy; the University of Lund in Sweden; the University of Oulu in Finland; the Singapore University of Technology and Design and Nanyang Technological University in Singapore; Tsinghua University in China; Kyung Hee University in Korea; Draper Laboratory; the Jet Propulsion Laboratory; and Mitsubishi Electric Research Laboratories.

Current research topics include network localization and navigation, network interference exploitation, intrinsic wireless secrecy, adaptive diversity techniques, and ultra-wide-bandwidth systems. Details of a few specific projects are provided below.

Network Localization and Navigation

The group has continued its research on network localization and navigation, in particular developing power management strategies for asynchronous localization networks and designing and implementing network localization and navigation algorithms. The group proposes novel power management strategies wherein each node individually minimizes the square position error bound penalized by its power cost. These strategies are obtained as solutions to two power management games formulated under knowledge of location information and global information. Analytical and numerical results show that the proposed strategies significantly reduce energy consumption with only marginal performance loss in position accuracy. The group is also developing a system that realizes the key ideas of network localization and navigation, including the exploitation of spatio-temporal cooperation and the use of environmental knowledge. The group is designing a real-time belief propagation algorithm to fuse map information with inter- and intra-user measurements, including range estimates among different users as well as acceleration and angular velocity estimates. This algorithm is implemented using only smartphones. Experimental results show that the system provides reliable location information and that spatial cooperation remarkably reduces the location uncertainty of users.

Intrinsic Network Secrecy

The group has also devoted effort to studying how physical-layer security can strengthen secrecy of communications by exploiting intrinsic properties of communication channels and by engineering network interference. Based on the theoretical framework developed over the past few years, the group has created a cooperative interference engineering technique—generalized interference alignment—to enhance wireless network secrecy. The feasibility of this technique has been characterized, and insights have been gained on how the technique can benefit wireless network secrecy. To determine how network topology affects wireless network secrecy, the group has also analyzed the secrecy performance of stochastic networks with heterogonous density. Insights obtained from this analysis will guide the design of wireless networks for a new level of communication confidentiality.

Highlights, Awards, and Events

The laboratory continues to organize the broadly attended LIDS seminar series and the LIDS student conference, which is organized solely by LIDS students and provides an interactive forum to discuss their research. LIDS produces a community-oriented magazine, *LIDS|ALL*, which includes articles on important events related to LIDS and profiles of individuals whose lives have been shaped by LIDS in significant ways. LIDS has recently launched a new website with enhanced content and design.

LIDS faculty continue to be involved in the organization of major workshops and conferences, including the IDSS Finance Workshop (co-organized by Professors Dahleh, Lo, and Ozdaglar), an NSF workshop focusing on smart cities (co-organized by Professor Karaman), the Marconi Society Technology and Entrepreneurship Symposium (organized by Professor Win), the Fifth International Conference on Continuous Optimization (co-organized by Professor Ozdaglar), and the Graphical Models, Statistical Inference, and Algorithms Workshop (co-organized by Professor Shah).

In March, LIDS hosted a symposium in honor of Professor Alan Willsky called "Statistical Inference under the Willskyan Lens." Professor Mitter led the organization of the event and brought in a set of exceptional speakers. The symposium was a great success and a wonderful way to celebrate Professor Willsky's broad and deep contributions to the field of information and decision systems over the course of more than 40 years.

Finally, LIDS faculty, students, and alumni continue to receive substantial recognition for their contributions, with numerous national and international awards and honors. Some notable examples are listed below.

Awards

Daron Acemoglu and LIDS Professors Munther Dahleh and Emilio Frazzoli (together with collaborators Giacomo Como from Lund University and Ketan Savla from the University of Southern California) won the IEEE Control Systems Society's 2015 George S. Axelby Outstanding Paper Award for "Robust Distributed Routing in Dynamical Networks–Part II: Strong Resilience, Equilibrium Selection and Cascaded Failures."

Dimitri Bertsekas was the 2015 winner of the George B. Dantzig Prize from the Society for Industrial and Applied Mathematics (SIAM) and the Mathematical Programming Society. This prize, one of the most prestigious awards in optimization, is bestowed once every three years for original research having a major impact on mathematical programming.

Professor G. David Forney, Jr., was awarded the 2016 Institute of Electrical and Electronics Engineers Medal of Honor, the highest award bestowed by IEEE.

Pablo Parrilo and Mardavij Roozbehani, together with Amir Ali Ahmadi PhD ’11 (Princeton University) and Raphael Jungers (Université catholique de Louvain), received the 2015 SIAG/CST (SIAM Activity Group in Control and Systems Theory) Best SICON Paper Prize for “Joint Spectral Radius and Path-Complete Graph Lyapunov Functions.”

Graduate student Jonathan Perry was awarded a Facebook fellowship for fall 2015 to spring 2017 (thesis advisor: Devavrat Shah).

Yury Polyanskiy received the EECS Jerome H. Saltzer Teaching Award.

Devavrat Shah (together with collaborators Vivek Farias of the Sloan School of Management and Srikanth Jagabathula PhD ’11 of New York University’s Stern School of Business) received the INFORMS (Institute for Operations Research and the Management Sciences) Revenue Management and Pricing Section Prize for the paper “A Nonparametric Approach to Modeling Choice with Limited Data.”

Graduate student Jennifer Tang (with collaborators Da Wang, Yury Polyanskiy, and Professor Gregory Wornell) won first prize at the Claude Shannon Centennial Student Competition organized by Bell Labs for her work “Defect Tolerance: Fundamental Limits and Examples” (thesis advisor: Yury Polyanskiy).

John Tsitsiklis received the 2016 ACM SIGMETRICS (Special Interest Group for the Computer Systems Performance Evaluation Community) Achievement Award in recognition of his fundamental contributions to decentralized control and consensus, approximate dynamic programming, and statistical learning.

Caroline Uhler was awarded the 2015 Doherty Professorship in Ocean Utilization. Uhler also received the Charles E. Reed Faculty Initiative Fund Award.

Moe Win received an Institute of Advanced Study Natural Sciences and Technology Fellowship. Also, Win (with student coauthors Stefania Bartoletti and Wenhan Dai and colleague Andrea Conti) received the IEEE CWIT Student Paper Award (first place) at the 2015 IEEE Canadian Workshop on Information Theory.

Honors

Munther Dahleh was selected to be a 2017 fellow of the International Federation of Automatic Control.

Jonathan How was named a 2016 fellow of the American Institute of Aeronautics and Astronautics.

Sertac Karaman was promoted to associate professor effective July 1, 2016. In addition, Karaman was selected as the Class of '48 Career Development Chair, also effective July 1, 2016.

Sanjoy Mitter was elected a foreign fellow of the Indian National Academy of Engineering.

LIDS director Asuman Ozdaglar was appointed to the Joseph F. and Nancy P. Keithley Professorship in Electrical Engineering. Ozdaglar gave plenary talks on distributed and incremental optimization methods at the 2015 Conference on Neural Information Processing Systems and the 2015 International Symposium of Mathematical Programming; also, she presented a seminar at the Institute for Mathematics and Its Applications.

Pablo Parrilo was named an IEEE Fellow, the highest grade of membership in IEEE.

Devavrat Shah was promoted to full professor effective July 1, 2016.

Future Outlook

LIDS continues to be a world-leading center for foundational work in information and decision sciences. This has been a particularly important year, with LIDS joining the newly launched Institute for Data, Systems, and Society. As noted, LIDS faculty are playing a pivotal role in, among other areas, defining the new institute's intellectual agenda and designing new academic programs in statistics and complex systems. We are excited about leading new research projects in fields as diverse as energy, transportation, social networks, health care analytics, and financial systems in which we can combine our rigorous quantitative work with vast amounts of relevant data and detailed knowledge of relevant institutions and practices. With the addition of new faculty and researchers and interactions with a number of world-leading researchers we are planning to host this year, we expect LIDS activities to continue to grow and expand in 2016–2017 and anticipate that LIDS will play a leadership role in defining and addressing important societal problems both within IDSS and in the broader global community.

Asu Ozdaglar
Director
Professor of Electrical Engineering and Computer Science

Pablo Parrilo
Associate Director
Professor of Electrical Engineering and Computer Science