FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES

JANUARY 29 AND FEBRUARY 5 2018

Caltech

FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES 2018

ALL TALKS ARE ONE HOUR LONG, AND WILL BE HELD IN **ANNENBERG 105**.
COFFEE BREAKS WILL BE HELD IN THE ANNENBERG LOBBY.

JANUARY 29

9:00am	Introduction	Venkat Chandrasekaran Caltech
9:15am	Approximate MCMC in Theory and Practice	James Johndrow Stanford University
10:15am	Coffee Break	
10:45am	Sparse Linear Algebra in the Exascale Era	Erin Carson New York University
1:00pm	Interactive Algorithms for Multiple Hypothesis Testing	Aaditya Ramdas UC Berkeley
2:00pm	Coffee Break	
2:30pm	Imaging the Invisible	Katie Bouman MIT
3:30pm	Coffee Break	
4:00pm	Machine Learning By the People, For the People	Nika Haghtalab Carnegie Mellon University

FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES 2018

ALL TALKS ARE ONE HOUR LONG, AND WILL BE HELD IN **ANNENBERG 105**.
COFFEE BREAKS WILL BE HELD IN THE ANNENBERG LOBBY.

FEBRUARY 5

9:00am	Introduction	Venkat Chandrasekaran Caltech
9:15am	Securing Computation on Untrusted Platforms	Justin Holmgren MIT
10:15am	Coffee Break	
10:45am	An Algorithm for Overcoming the Curse of Dimensionality in Hamilton-Jacobi Equations	Yat Tin (Raymond) Chow UCLA
1:00pm	Large Sample Asymptotics of Graph-Based Methods in Machine Learning: Mathematical Analysis and Implications	Nicolas Garcia Trillos Brown University
2:00pm	Coffee Break	
2:30pm	A Conditional Gaussian Framework for Uncertainty Quantification, Data Assimilation and Prediction of Complex Nonlinear Turbulent Dynamical Systems	Nan Chen New York University
3:30pm	Coffee Break	
4:00pm	Inverse Problems and Unsupervised Learning with Applications to Cryo-Electron Microscopy	Roy Lederman Princeton University

Imaging the Invisible

Katie Bouman
Massachusetts Institute of Technology

Imaging plays a critical role in advancing science. However, as science continues to push the boundaries of knowledge, traditional imaging approaches are reaching observational limits. In this talk I discuss how combining ideas from physics, signal processing, and machine learning has allowed us to transcend these limits in order to see people moving behind walls and take the first picture of a black hole.

It is theorized that the heart of the Milky Way galaxy is host to a evolving black hole. An image of this black hole could help to address a number of important scientific questions. Unfortunately, due to its small size, traditional imaging approaches require an Earth-sized radio telescope. In this talk, I discuss techniques we have developed to photograph the black hole using a network of telescopes scattered across the globe. Imaging the black hole's rapidly evolving structure with this computational telescope requires us to reconstruct video from sparse measurements, heavily corrupted by atmospheric error. Additionally, I present an evaluation process developed to establish confidence in reconstructions done with real scientific data. These methods and evaluation techniques are currently being applied in ongoing work to take the first picture of a black hole as part of the Event Horizon Telescope collaboration.

Biography:
Katie Bouman is a postdoctoral fellow in the Harvard-Smithsonian Center for Astrophysics. She recently received her Ph.D. in the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT. Before coming to MIT, she received her bachelor's degree in Electrical Engineering from the University of Michigan. The focus of her research is on using emerging computational methods to push the boundaries of interdisciplinary imaging.

Sparse Linear Algebra in the Exascale Era

Erin Carson
New York University

Sparse linear algebra problems, typically solved using iterative methods, are ubiquitous throughout scientific and data analysis applications and are often the most expensive computations in large-scale codes due to the high cost of data movement. Approaches to improving the performance of iterative methods typically involve modifying or restructuring the algorithm to reduce or hide this cost. Such modifications can, however, result in drastically different behavior in terms of convergence rate and accuracy. A clear, thorough understanding of how inexact computations, due to either finite precision error or intentional approximation, affect numerical behavior is thus imperative in balancing the tradeoffs between accuracy, convergence rate, and performance in practical settings.

In this talk, we focus on two general classes of iterative methods for solving linear systems: Krylov subspace methods and iterative refinement. We present bounds on the attainable accuracy and convergence rate in finite precision s-step and pipelined Krylov subspace methods, two popular variants designed for high performance. For s-step methods, we demonstrate that our bounds on attainable accuracy can lead to adaptive approaches that both achieve efficient parallel performance and ensure that a user-specified accuracy is attained. We present two such adaptive approaches, including a residual replacement scheme and a variable s-step technique in which the parameter s is chosen dynamically throughout the iterations. Motivated by the recent trend of multiprecision capabilities in hardware, we present new forward and backward error bounds for a general iterative refinement scheme using three precisions. The analysis suggests that on architectures where half precision is implemented efficiently, it is possible to solve certain linear systems up to twice as fast and to greater accuracy.

As we push toward exascale level computing and beyond, designing efficient, accurate algorithms for emerging architectures and applications is of utmost importance. We discuss extensions to machine learning and data analysis applications, the development of numerical autotuning tools, and the broader challenge of understanding what increasingly large problem sizes will mean for finite precision computation both in theory and practice.

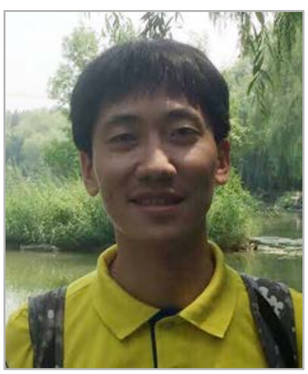
Biography:

Erin Carson is a Courant Instructor/Assistant Professor at the Courant Institute of Mathematical Sciences at New York University. She received her PhD at the University of California, Berkeley in 2015, advised by James Demmel and Armando Fox and supported by a National Defense Science and Engineering Graduate Fellowship. Her research interests lie at the intersection of high-performance computing, numerical linear algebra, and parallel algorithms, with a focus on analyzing tradeoffs between accuracy and performance in algorithms for large-scale sparse linear algebra.

A Conditional Gaussian Framework for Uncertainty Quantification, Data Assimilation and Prediction of Complex Nonlinear Turbulent Dynamical Systems

Nan Chen
New York University

A conditional Gaussian framework for uncertainty quantification, data assimilation and prediction of nonlinear turbulent dynamical systems is developed. Despite the conditional Gaussianity, the dynamics remain highly nonlinear and are able to capture strongly non-Gaussian features such as intermittency and extreme events in nature. The conditional Gaussian structure allows efficient and analytically solvable conditional statistics that facilitates the real-time data assimilation and prediction. The talk will include three applications of such conditional Gaussian framework. In the first part, a physics-constrained nonlinear stochastic model is developed, and is applied to the data assimilation and the prediction of the Madden-Julian oscillation with strongly intermittent features. The second part regards the state estimation and uncertainty quantification of multiscale and turbulent ocean flows using noisy Lagrangian tracers. Rigorous analysis shows that an exponential increase in the number of tracers is required for reducing the uncertainty by a fixed amount. This indicates a practical information barrier. In the last part of the talk, an efficient statistically accurate algorithm is developed that is able to solve the high dimensional Fokker-Planck equation with strong non-Gaussian features within the conditional Gaussian framework and beat the curse of dimensions.



Biography:

Nan Chen is currently a postdoc research associate at Courant Institute of Mathematical Sciences (CIMS) and Center of Atmosphere and Ocean Science (CAOS), New York University.

Chen received his PhD degree in Mathematics and Atmosphere and Ocean Science at CIMS and CAOS, New York University in 2016. His thesis advisor was Dr. Andrew J. Majda. Chen was awarded the Kurt O. Friedrichs prize in May 2016 for an out standing dissertation in mathematics. Chen's undergraduate major was Mechanical Engineering and he received a Master's degree from the School of Mathematical Sciences, Fudan University. He also spent one year visiting the Department of Scientific Computing, Florida State University, collaborating with Dr. Max Gunzberger and Dr. Xiaoming Wang.

Chen's research lies in the contemporary applied mathematics, fluids and geophysics. Problems with large dimensions, turbulence and partial information are particularly what he is concerned with. Mathematical and physical problems in uncertainty quantification, data assimilation, applied stochastic analysis, inverse problems, high-dimensional data analysis and effective prediction all belong to his research topics. He is also devoted to proposing efficient and statistically accurate algorithms to ameliorate the curse of dimensionality for large-dimensional complex systems with strong non-Gaussian features. In addition, He is active in developing both dynamical and stochastic models and use these models to predict real-world phenomena related to atmosphere ocean science, climate and other complex systems such as the Madden-Julian Oscillation (MJO), the monsoon and the El Nino Southern Oscillation (ENSO) based on real observational data. Some of his works have been attracted by media attention.

An Algorithm for Overcoming the Curse of Dimensionality in Hamilton-Jacobi Equations

Yat Tin (Raymond) Chow
University of California, Los Angeles

In this talk, we discuss an algorithm to overcome the curse of dimensionality, in possibly non-convex/time/state-dependent Hamilton-Jacobi partial differential equations. They may arise from optimal control and differential game problems, and are generally difficult to solve numerically in high dimensions.

A major contribution of our works is to consider an optimization problem over a single vector of the same dimension as the dimension of the HJ PDE. To do so, we consider the new approach using Hopf-type formulas. The sub-problems are now independent, and they can be implemented in an embarrassingly parallel fashion. That is ideal for perfect scaling in parallel computing.

The algorithm is proposed to overcome the curse of dimensionality when solving high dimensional HJ PDE. Our method is expected to have application in control theory, differential game problems, and elsewhere. This approach can be extended to the computational of a Hamilton-Jacobi equation in the Wasserstein space, and is expected to have applications in mean field control problems, optimal transport and mean field games.

Biography:
I am currently a CAM Assistant Professor in Department of Mathematics, UCLA. I received my Bachelor's degree in 2009, my Master's degree in 2012, and my Ph.D. in Mathematics in 2015, from the Chinese University of Hong Kong. I have been working on overcoming the curse of dimensionality in solving Hamilton-Jacobi equations arising in optimal control and differential games, which may extend to problems including optimal transport and mean field games. I have also been working on optimization methods, e.g. coordinate update methods. Moreover, I have worked on inverse problems, in particular, coefficient determinations in partial differential equations, which has applications in many areas, for instance, tomography (e.g. electrical impedance tomography).

Large Sample Asymptotics of Graph-Based Methods in Machine Learning: Mathematical Analysis and Implications

Nicolas Garcia Trillos
Brown University

Many machine learning procedures aimed to extract information from data can be defined as precise mathematical objects that are constructed in terms of the data. It is often assumed that the data is “big” in complexity but also in quantity, and in this “large amount of data” setting, a basic mathematical concept that one can explore is that of *closure* of a given class of statistical procedures (i.e. what are the limiting procedures as the number of data points available goes to infinity.) In this talk, I will explore this notion in the context of graph-based methods. Examples of such methods include minimization of Cheeger cuts, spectral clustering, and graph-based bayesian semi-supervised learning, among others. I will introduce some of the mathematical ideas needed for the analysis, as well as show some of the implications of it: our results show statistical consistency of the methods, provide with quantitative information in the form of scaling of parameters and rates of convergence, imply qualitative properties at the discrete level, and suggest the use of appropriate algorithms.

Following the same line of thought, I will then show how ideas from optimal transport can be used to define a large class of evolution equations on graphs, and present some theoretical results that connect them with their continuum limits; I will then indicate the relationship between these mathematical constructs and the analysis of large sample asymptotics of other types of machine learning methodologies.

I will finish my talk suggesting some future directions for research, both at the theoretical level as well as at the practical level (analysis of real data).

Biography:
Nicolas Garcia Trillos obtained his PhD in Mathematical Sciences from Carnegie Mellon University in 2015 and his bachelor’s degree in Mathematics from Los Andes University in Bogota, Colombia, in 2010. He is currently at Brown University for his last year of a three year post-doctoral position in the Division of Applied Mathematics. His research interests are in applied analysis, probability, and statistics. In his research work, he uses tools from mathematical analysis to study machine learning problems. He finds the concept of artificial intelligence fascinating from many different perspectives including literature. He likes to read short stories, play soccer and tennis, go on walks through city streets and forests, spend time with his wife and watch her put color on her canvases.

Machine Learning By the People, For the People

Nika Haghtalab
Carnegie Mellon University

Typical analysis of learning algorithms considers their outcome in isolation from the effects that they may have on the process that generates the data or the entity that is interested in learning. However, current technological trends mean that people and organizations increasingly interact with learning systems, making it necessary to consider these effects, which fundamentally change the nature of learning and the challenges involved. In this talk, I will explore three lines of research from my work on the theoretical aspects of machine learning and algorithmic economics that account for these interactions: learning optimal policies in game-theoretic settings, without an accurate behavioral model, by interacting with people; managing people's expertise and resources in data-collection and machine learning; and collaborative learning in a setting where multiple learners interact with each other to discover similar underlying concepts.

Biography:
Nika Haghtalab is a Ph.D. candidate at the Computer Science Department of Carnegie Mellon University, co-advised by Avrim Blum and Ariel Procaccia. Her research interests include learning theory and algorithmic economics. She is a recipient of the IBM and Microsoft Research Ph.D. fellowships and the Siebel Scholarship.

Securing Computation on Untrusted Platforms

Justin Holmgren
Massachusetts Institute of Technology

In today's networked world, weak devices increasingly rely on remote servers both to store data and to perform costly computations. Unfortunately, these servers may be easily hackable or otherwise untrustworthy. Therefore, without assuming honest behavior on the server's part, we would like to guarantee two basic security objectives:

1. (Correctness) It is possible to verify the correctness of the server's computations much more efficiently than by re-executing the computation.
2. (Privacy) A server learns nothing about the computation it performs, other than (perhaps) the output.

I will present recent results that achieve both these goals for arbitrary computations, and I will conclude with a discussion of open problems and future directions.

Biography:
Justin Holmgren is a graduating PhD student at MIT, advised by Professor Shafi Goldwasser. He has a general interest in theoretical computer science, especially cryptography and complexity. His research so far has mainly focused on developing provably secure protocols which remove or reduce trust requirements in client-server interactions. Within this area, he has particularly focused on the problem of securely outsourcing computation.

Approximate MCMC in Theory and Practice

James Johndrow
Stanford University

Rapid growth in the number of samples in typical datasets and the number of parameters in statistical and other mathematical models poses significant computational challenges. A popular strategy for reducing the computational cost of Markov chain Monte Carlo (MCMC) is to replace the exact Markov kernel with an approximation that is less costly to simulate. We give a number of results on the convergence properties of these approximating Markov chains and the performance of time-averages in approximating expectations of functions with respect to the target measure. The talk is structured around several canonical examples that both motivate the results and illustrate the power and limitations of this approach to scaling up MCMC. While the applications I discuss are mainly statistical, the results are applicable to large classes of Markov chains.

Biography:
James Johndrow works at the interface of statistics, applied probability, and computational science. His primary research interest is scalable Bayesian computation. Work in this area includes characterizing the computational cost of MCMC, developing measures of optimality that incorporate both statistical and computational performance of procedures, designing improved algorithms for estimation of high-dimensional statistical models, and constructing a theory of approximating Markov chains. Other topics include point processes with applications in multivariate extreme value theory and phylodynamics. James also has several active interdisciplinary projects in algorithmic fairness, population estimation and record linkage for human rights applications, A/B testing, population genetics, and treatment effect estimation for heavy-tailed distributions. James received a Ph.D. in Statistical Science from Duke in 2016, and is currently a Stein Fellow/Lecturer in the Statistics department at Stanford.

Inverse Problems and Unsupervised Learning with Applications to Cryo-Electron Microscopy

Roy Lederman
Princeton University

Cryo-Electron Microscopy (cryo-EM) is an imaging technology that is revolutionizing structural biology; the Nobel Prize in Chemistry 2017 was recently awarded to Jacques Dubochet, Joachim Frank and Richard Henderson "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".

Cryo-electron microscopes produce a large number of very noisy two-dimensional projection images of individual frozen molecules. Unlike related methods, such as computed tomography (CT), the viewing direction of each image is unknown. The unknown directions, together with extreme levels of noise and additional technical factors, make the determination of the structure of molecules challenging.

While other methods for structure determination, such as x-ray crystallography and nuclear magnetic resonance (NMR), measure ensembles of molecules, cryo-electron microscopes produce images of individual molecules. Therefore, cryo-EM could potentially be used to study mixtures of different conformations of molecules. Indeed, current algorithms have been very successful at analyzing homogeneous samples, and can recover some distinct conformations mixed in solutions, but, the determination of multiple conformations, and in particular, continuums of similar conformations (continuous heterogeneity), remains one of the open problems in cryo-EM.

I will discuss a one-dimensional discrete model problem, Heterogeneous Multireference Alignment, which captures many of the group properties and other mathematical properties of the cryo-EM problem. I will then discuss different components which we are introducing in order to address the problem of continuous heterogeneity in cryo-EM: 1. "hyper-molecules," the mathematical formulation of truly continuously heterogeneous molecules, 2. Computational and numerical tools for formulating associated priors, and 3. Bayesian algorithms for inverse problems with an unsupervised-learning component for recovering such hyper-molecules in cryo-EM.

Biography:

Roy Lederman is a postdoc at the Program in Applied and Computational Mathematics at Princeton University, working with Amit Singer. Before joining Princeton, he spent a year as a Gibbs Assistant Professor at Yale University, where he had completed his PhD in Applied Mathematics, working with Vladimir Rokhlin and Ronald Coifman. Roy holds a BSc in Electrical Engineering and a BSc in Physics from Tel Aviv University.

Interactive Algorithms for Multiple Hypothesis Testing

Aaditya Ramdas
UC Berkeley

Data science is at a crossroads. Each year, thousands of new data scientists are entering science and technology, after a broad training in a variety of fields. Modern data science is often exploratory in nature, with datasets being collected and dissected in an interactive manner. Classical guarantees that accompany many statistical methods are often invalidated by their non-standard interactive use, resulting in an underestimated risk of falsely discovering correlations or patterns. It is a pressing challenge to upgrade existing tools, or create new ones, that are robust to involving a human-in-the-loop.

In this talk, I will describe two new advances that enable some amount of interactivity while testing multiple hypotheses, and control the resulting selection bias. I will first introduce a new framework, STAR, that uses partial *masking* to divide the available information into two parts, one for selecting a set of potential discoveries, and the other for inference on the selected set. I will then show that it is possible to flip the traditional roles of the algorithm and the scientist, allowing the scientist to make post-hoc decisions after seeing the realization of an algorithm on the data. The theoretical basis for both advances is founded in the theory of martingales : in the first, the user defines the martingale and associated filtration interactively, and in the second, we move from optional stopping to optional *spotting* by proving uniform concentration bounds on relevant martingales.

This talk will feature joint work with (alphabetically) Rina Barber, Jianbo Chen, Will Fithian, Kevin Jamieson, Michael Jordan, Eugene Katsevich, Lihua Lei, Max Rabinovich, Martin Wainwright, Fanny Yang and Tijana Zrnic.

Biography:

Aaditya Ramdas is a postdoctoral researcher in Statistics and EECS at UC Berkeley, advised by Michael Jordan and Martin Wainwright. He finished his PhD in Statistics and Machine Learning at CMU, advised by Larry Wasserman and Aarti Singh, winning the Best Thesis Award in Statistics. A lot of his research focuses on modern aspects of reproducibility in science and technology — involving statistical testing and false discovery rate control in static and dynamic settings.