Qingqing Huang

+1 617.909.3319 | qqh@mit.edu | http://qingqinghuang.github.io

Affiliation

Laboratory of Information and Decision Systems,
Department of Electrical Engineering and Computer Sciences,
Massachusetts Institute of Technology

Education

Massachusetts Institute of Technology (MIT)

Ph.D candidate in Electrical Engineering **June 2013 – May 2016 (expected)**
Thesis Topic: Make impossible possible, make possible optimal, make optimal practical – algorithmic statistics for learning mixture models.
Thesis Committee: Munther Dahleh, Sham Kakade, Pablo Parrilo.

Master in Electrical Engineering **Sep 2011 – May 2013**
Thesis Topic: Efficiency-Risk Tradeoffs in Electricity Markets with Dynamic Demand Response
Research Advisor: Munther Dahleh

Hong Kong University of Science and Technology (HKUST)

Bachelor of Engineering in Electrical Engineering **Sep 2006 – June 2011**
Bachelor of Business Administration in Economics **Sep 2006 – June 2011**

Research Interests

Statistical learning theory, machine Learning, networked systems.

Ph.D thesis topic: I focus on a set of statistical learning problems regarding mixture models. Mixture models (examples include Gaussian Mixtures (GMM), topic models, and Hidden Markov Models (HMM)) serve to model the scenario where the underlying mechanism of each observed data sample belongs to a finite number of different sources. It is a class of powerful models which finds application in a wide range of unsupervised learning tasks, such as speech recognition, document classification, super-resolution imaging, community detection, and low rank matrix recovery for recommendation tasks.

The structural property of the distribution with the latent variable introduces non-convexity to the learning problem, making it much harder than the unstructured problems. To this end, I ask and attempt address three questions: Can we efficiently learn the model parameters, assuming some non-degeneracy of the instances? Can we achieve it with optimal sample complexity with fast algorithms? Can we make the learning algorithms also robust to model mis-specifications?

Future research directions: I will continue to research the broad field of statistical learning, in particular exploring the two aspects of *robustness* and *dynamics* of various learning questions. I am also particularly interested in the direction of application of statistics and machine learning to tackle real-life challenges in economics and sociology, in particular addressing the issue of sparse and noisy data in these fields, as in the non-asymptotic regime the conventional statistics and learning methods do not directly apply.

Research Experience

Graduate researcher at **Laboratory of Information and Decision Systems, MIT**
Thesis advisor : Munther Dahleh and Sham Kakade Sep 2011 – Present

Research Internship at **Machine Learning Group, Microsoft Research New England**
May 2014 – Aug 2014
Mentor: Sham Kakade May 2015 – Aug 2015

Visitor at **Big Data Lab, Baidu, Beijing**
Host: Tong Zhang Dec 2014

Research Assistant at **Wireless Communication Group, ECE, HKUST**
Research advisor: Vincent K.N. Lau June 2009 – Jan 2011

Teaching Experience

Department of Electrical Engineering and Computer Sciences, MIT

Teaching Assistant for 6.438 "Algorithms for Inference" — Fall 2013

It was a graduate level class with approximately 60 students with 2 TAs. TA duties include giving bi-weekly recitations, designing and administrating class project, holding weekly office hours, assisting the instructor to prepare homework assignments and exams.
Instructor: Devavrat Shah

Teaching Assistant for 6.207 "Networks" — Spring 2014

It was an undergraduate level class with approximately 80 students with 2 TAs. TA duties include giving bi-weekly recitations, holding weekly office hours, and assisting the instructor to prepare homework assignments and exams.

Teaching Assistant for 6.UAR "Prep for Undergrad Research" — Fall 2015

It was an undergraduate level class with approximately 180 students with 4 TAs. The TA duties include assisting the students in their individual research projects step by step, offering advice on project proposals and posters.

Extra Curricular Activities

Publication Chair of *Graduate Student Association, MIT* — Jan 2013 – Dec 2013

In charge of website maintenance, poster design for publicizing events

Co-president of *Graduate Women in Course 6 (GW6), MIT* — Jan 2012 – Dec 2012

Organized extra curricular activities for graduate women in EECS department of MIT

Software Engineer Intern at *Yunzhou-Tech Company, China* — June 2011 – Aug 2011

Worked on navigation algorithm improvements for unmanned surface vehicles

IT Engineering at *Kwong Wah Hospital, Hong Kong* — Jan 2011 – May 2011

Developed a web-based medical image archiving system (Student Civic Fellow Program)

Project Assistant at *Heep Hong Society, Hong Kong* — June 2010 – Oct 2010

Collaborated to develop a computer-based learning package for autistic children

Programming

Proficient in Matlab, C++, Python.

Awards

Xerox-MIT Fellowship — 2012
Irwin Mark Jacobs and Joan Klein Jacobs Presidential Fellowship — 2011
Silver Medal in the National Physics Olympiad, China — 2005

Publications (Listed roughly in reverse chronological order. Authors are ordered alphabetically for CS theory papers)

Statistical learning theory

"Recovering Structured Probability Matrices" **H**, Sham Kakade, Wenhao Kong, Gregory Valiant, submitted to STOC 2016

"Learning Mixture of Gaussians in High dimensions " Rong Ge, **H**, Sham Kakade, appeared in 47th Annual Symposium on the Theory of Computing (STOC), 2015

"Super-Resolution off the Grid " **H**, Sham Kakade
appeared in the 30th Annual Conference on Neural Information Processing Systems (NIPS), 2015

"Minimal Realization Problems for Hidden Markov Models " **H**, Rong Ge, Sham Kakade, Munther Dahleh
Journal version submitted to IEEE Transactions on Signal Processing; conference version appeared in the 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton) , 2014.

"A Greedy Algorithm for Nonnegative Matrix and Tensor Factorization" **H**, Tong Zhang
working paper

Smart Grid technologies

"Efficiency-Risk Tradeoffs in Electricity Markets with Dynamic Demand Response" **H**, Mardavij Roozebehani, Munther Dahleh
Journal version appeared in IEEE Transactions on Smart Grid (TSG), 2014
Conference version appeared in IEEE Conference on Decision and Control (CDC), 2012

"Dynamic Fault Diagnosis in Power Grids Using Hidden Markov Models" **H**, Leilai Shao, Na Li
Journal version appeared in IEEE Transactions on Power System (TPS), 2015
Conference version appeared in IEEE American Control Conference (ACC), 2015

Miscellaneous

"H2-Based Network Volatility Measures" **H**, Ye Yuan, Jorge Goncalves, and Munther Dahleh
Conference version appeared in IEEE American Control Conference (ACC), 2014

"Queue-Aware Dynamic Clustering and Power Allocation for Network MIMO Systems via Distributed Stochastic Learning " Ying Cui, **H**, Vincent K.N. Lau
Journal version appeared in IEEE Transactions on Signal Processing (TSP), 2011

"Delay-Optimal Orthogonal Beam forming and Power Control for MIMO system with Reduced CSI Feedback " **H**, Ying Cui, Vincent K.N. Lau
Technical report, 2011

Abstract of selected publications

"Recovering Structured Probability Matrices"
We consider the problem of recovering a matrix B of size $M \times M$, which represents a probability distribution over M^2 outcomes, given access to independent sample "counts" generated according to the distribution B. How can structural properties of the underlying matrix B be leveraged to yield computationally efficient and information theoretically optimal reconstruction algorithms? When can accurate reconstruction be accomplished in the sparse data regime?
This basic problem lies at the core of a number of questions that are currently being considered by different communities, including community detection in sparse random graphs, learning structured probabilistic models such as topic models and hidden Markov models, as well as the efforts from the natural language processing community to compute "word embeddings". Many aspects of this problem – both in terms of learning and property testing, and on both the algorithmic and information theoretic sidesremain open.

Our results apply to the setting where the $M \times M$ probability matrix B is of rank 2. We propose an efficient algorithm that accurately recovers the underlying matrix using $\Theta(M)$ samples. The linear sample complexity is optimal, up to constant factors, in an extremely strong sense: even testing basic properties of the underlying matrix, such as whether it has rank 1 or 2, requires $\Omega(M)$ samples.

"Super-Resolution off the Grid"

Super-resolution is the problem of recovering a superposition of point sources using bandlimited measurements, which may be corrupted with noise. This signal processing problem arises in numerous imaging problems, ranging from astronomy to biology to spectroscopy, where it is common to take (coarse) Fourier measurements of an object. Of particular interest is in obtaining estimation procedures which are robust to noise, with the following desirable statistical and computational properties: we seek to use coarse Fourier measurements (bounded by some *cutoff frequency*); we hope to take a (quantifiably) small number of measurements; we desire our algorithm to run quickly.

Suppose we have k point sources in d dimensions, where the points are separated by at least Δ from each other (in Euclidean distance). This work provides an algorithm with the following favorable guarantees:

– The algorithm uses Fourier measurements, whose frequencies are bounded by $O(1/\Delta)$ (up to log factors). Previous algorithms require a *cutoff frequency* which may be as large as $\Omega(\sqrt{d}/\Delta)$.

– The number of measurements taken by and the computational complexity of our algorithm are bounded by a polynomial in both the number of points k and the dimension d, with *no* dependence on the separation Δ. In contrast, previous algorithms depended inverse polynomially on the minimal separation and exponentially on the dimension for both of these quantities.

"Learning Mixture of Gaussians in High dimensions"

Efficiently learning mixture of Gaussians is a fundamental problem in statistics and learning theory. Given samples coming from a random one out of k Gaussian distributions in n dimensional space, the learning problem asks to estimate the means and the covariance matrices of these Gaussians. This learning problem arises in many areas ranging from the natural sciences to the social sciences, and has also found many machine learning applications.

Unfortunately, learning mixture of Gaussians is an information theoretically hard problem: in order to learn the parameters up to a reasonable accuracy, the number of samples required is exponential in the number of Gaussian components in the worst case. In this work, we show that provided we are in high enough dimensions, the class of Gaussian mixtures is learnable in its most general form under a smoothed analysis framework, where the parameters are randomly perturbed from an adversarial starting point.

In particular, given samples from a mixture of Gaussians with randomly perturbed parameters, when $n \geq \Omega(k^2)$, we give an algorithm that learns the parameters with polynomial running time and using polynomial number of samples.

The central algorithmic ideas consist of new ways to decompose the moment tensor of the Gaussian mixture by exploiting its structural properties. The symmetries of this tensor are derived from the combinatorial structure of higher order moments of Gaussian distributions (sometimes referred to as Isserlis' theorem or Wick's theorem). We also develop new tools for bounding smallest singular values of structured random matrices, which could be useful in other smoothed analysis settings.

"Efficiency-Risk Tradeoffs in Electricity Markets with Dynamic Demand Response"

In order to study the impact of dynamic demand response in the future smart grid, we examine in an abstract framework, how a tradeoff between efficiency and risk arises under different market architectures. We first examine the system performance under non-cooperative and cooperative market architectures. The statistics of the stationary aggregate demand processes show that, although the non-cooperative load scheduling scheme leads to an efficiency loss, the stationary distribution of the corresponding aggregate demand process has a smaller tail, resulting in less frequent aggregate demand spikes. Cooperative dynamic demand response, on the other hand, makes the market place more efficient at the cost of increased risk of aggregate demand spikes. The market architecture determines the locus of the system performance with respect to the tradeoff curve.

We also investigate how a properly designed real-time electricity pricing mechanism can help the system operator achieve a target tradeoff between efficiency and risk in a non-cooperative market. We further provide a convex characterization of the Pareto front of system performance measures, which serves as a benchmark of the tradeoffs for the system operator to evaluate the pricing rules.

Last updated: December 1, 2015