# The University of Chicago

# Department of Computer Science & Mathematics

# Combinatorics & Theoretical Seminar

PRESENTS:

**Ilias Diakonikolas**



**University of Southern California**

Title: Statistical Query Lower Bounds for High-Dimensional Unsupervised Learning

Abstract:

We describe a general technique that yields the first Statistical Query lower bounds for a range of fundamental high dimensional learning problems. Our main results are for the problems of (1) learning Gaussian mixture models, and (2) robust learning of a single Gaussian distribution. For these problems, we show a super-polynomial gap between the sample complexity and the computational complexity of any Statistical Query (SQ) algorithm for the problem. SQ algorithms are a class of algorithms that are only allowed to query expectations of functions of the distribution rather than directly access samples. This class of algorithms is quite broad: a wide range of known algorithmic techniques in machine learning are known to be implementable using SQs.

Our SQ lower bounds are attained via a unified moment-matching technique that is useful in other contexts. Our method yields tight lower bounds for a number of related unsupervised estimation problems, including robust covariance estimation in spectral norm, and robust sparse mean estimation. Finally, for the classical problem of robustly testing an unknown mean Gaussian, we show a sample complexity lower bound that scales linearly in the dimension. This matches the sample complexity of the corresponding robust learning problem and separates the sample complexity

of robust testing from standard testing. This separation is surprising because such a gap does not exist for the corresponding learning problem.

(Based on joint work with Daniel Kane (UCSD) and Alistair Stewart (USC).)

Host: Prof. Andrew Drucker

**Tuesday, October 3, 2017**

**Ry. 251 @ 3:30 pm**

**(Refreshments will be served before the talk in Ry. 255 @ 3pm)**