

Title: Recent Progress in Algorithmic Robust Statistics via the Sum-of-Squares Method

Speaker: Pravesh K. Kothari (Carnegie Mellon University)

Abstract: Past five years have witnessed a sequence of successes in designing efficient algorithms for statistical estimation tasks when the input data is corrupted with a constant fraction of fully malicious outliers. The Sum-of-Squares (SoS) method for polynomial optimization has been an integral part of this story and is behind state-of-the-art robust learning algorithms for tasks such as estimating the mean, covariance, and higher moment tensors of a broad class of distributions, clustering and parameter estimation for spherical and non-spherical mixture models, linear regression, and list-decodable learning.

In this talk, I will attempt to demystify this (unreasonable?) effectiveness of the SoS method in robust statistics. I will argue that the utility of the SoS algorithm in robust statistics can be directly attributed to its capacity (via low-degree SoS proofs/certificates) to "reason about" analytic properties of probability distributions such as sub-gaussianity, hypercontractivity, and anti-concentration. I will discuss precise formulations of such statements, show how they lead to a principled blueprint for problems in robust statistics including the applications mentioned above, and point out natural gaps in our understanding of analytic properties within SoS, which, if resolved would yield improved guarantees for basic tasks in robust statistics.