

Title: Do Semidefinite Relaxations really solve Sparse PCA?

Danny Vilenchik

Estimating the leading principal components of high dimensional data, assuming those are sparse, is a problem of considerable interest, known as sparse-PCA. In this talk we review a semi-definite programming (SDP) approach, under a spiked covariance model with a single l_0 -sparse principal component. For this model, Amini and Wainwright recently proved that asymptotically, as the dimension and sample size tend to infinity, and a sufficiently sparse vector, the SDP solution is rank one and perfectly recovers the support of the sparse spike vector. In this talk we prove, in contrast, that once the spike sparsity exceeds a certain threshold, the SDP solution does not reflect the spike anymore. We complement our theoretical results with some simulations, and suggest a new approach for the sparse PCA problem that seems to outperform some popular techniques such as Diagonal Thresholding.