



The Hong Kong University of Science and Technology

Department of Mathematics

Seminar on Data Science

**Certifying Global Optimality of Graph Cuts via
Semidefinite Relaxation: A Performance Guarantee
for Spectral Clustering**

by

Dr. Shuyang Ling

Courant Institute of Mathematical Sciences

New York University

Abstract

Spectral clustering has become one of the most widely used clustering techniques when the structure of the individual clusters is non-convex or highly anisotropic. Yet, despite its immense popularity, there exists fairly little theory about performance guarantees for spectral clustering. This issue is partly due to the fact that spectral clustering typically involves two steps which complicated its theoretical analysis: first, the eigenvectors of the associated graph Laplacian are used to embed the dataset, and second, k-means clustering algorithm is applied to the embedded dataset to get the labels. This paper is devoted to the theoretical foundations of spectral clustering and graph cuts. We consider a convex relaxation of graph cuts, namely ratio cuts and normalized cuts, that makes the usual two-step approach of spectral clustering obsolete and at the same time gives rise to a rigorous theoretical analysis of graph cuts and spectral clustering. We derive deterministic bounds for successful spectral clustering via a spectral proximity condition that naturally depends on the algebraic connectivity of each cluster and the inter-cluster connectivity. Moreover, we demonstrate by means of some popular examples that our bounds can achieve near-optimality. Our findings are also fundamental for the theoretical understanding of kernel k-means. Joint work with Thomas Strohmer at UC Davis.

Date: Monday, 13 August 2018
Time: 3:00p.m. – 4:00p.m.
Venue: Room 5501 (lift 25, 26),
Academic Building, HKUST
All are welcome!