Fall 2026
BA990 & ECE590: Statistical inference on graphs
Syllabus

Schedule:	Wednesday 1:25pm-3:55pm, Fuqua Mosler Classroom
Professor:	Jiaming Xu jiaming.xu868@duke.edu, Fuqua W313
Office hours:	by appointment
Website:	https://faculty.fuqua.duke.edu/~jx77/Course_graphs.html

Note: The course is part of the broader Fall 2026 Duke IID Pop-Up Institute on Learning in Networks https://sites.duke.edu/learninginnetworks/. Students are encouraged to participate and get involved.

1 Content

An emerging research thread in statistics and machine learning deals with finding latent structures from data represented by graphs or matrices. This course will provide an introduction to mathematical and algorithmic tools for studying such problems. We will discuss information-theoretic methods for determining the fundamental limits. Complementing this objective of understanding the fundamental limits, another significant direction is to develop computationally efficient procedures that attain statistical optimality, or to understand the lack thereof. We will discuss computationally efficient procedures including spectral methods, low-degree polynomial methods, Linear/quadratic/semi-definite programming relaxations, message passing (belief propagation) algorithms, etc. Towards the end we will also discuss the recent trend of combining the statistical and algorithmic perspectives and the computational barriers in a series of statistical problems on large matrices and random graphs.

Specific topics will include spectral clustering, low-degree polynomial methods, planted clique and partition problem, community detection on stochastic block models, hidden Hamiltonian cycle problem, planted bipartite graph matching problem, noisy graph isomorphism, and detecting latent geometry or preferential attachment.

Tentative outline

1. **Introduction**: detection-recovery-estimation, sharp thresholds
2. **Spectral methods**: preliminaries from linear algebra, perturbation bound, application to clustering
3. **Planted clique**: degree test, spectral methods, low-degree polynomial methods, message passing algorithms
4. **Community detection**: stochastic block models, correlated recovery and mutual information, recovery thresholds, first and second moment methods, low-degree polynomial methods, non-backtracking matrices, belief propagation

5. **Semidefinite programming (SDP) relaxation**: KKT conditions and exact recovery threshold, Grothendieck inequality and consequences on clustering, robustness in semi-random models

6. **Hidden Hamiltonian cycle, bipartite graph matchings, and graph isomorphism problems**: sharp information-theoretic limits, analysis of linear/quadratic programming relaxations, Stieltjes transform of random matrix, applications to DNA sequencing, particle tracking, and network alignment.

7. **Detecting latent geometry or preferential attachment**: sharp detection and estimation thresholds for random geometric graphs and preferential attachment graphs.

8. **Computational limits**: low-degree polynomial methods, Polynomial-time randomized reduction, Planted dense subgraph problem

2 Administrivia

1. Course prerequisites: Maturity with probability theory and linear algebra. Familiarity with statistical theory, optimization, and algorithms.

2. Grading: 30% participation, 30% homework, 40% final project.

3. Class participation includes attending classes on time, raising good questions or providing good answers to questions, pointing out errors in the lecture notes, etc.

4. Final project: submit a report based on either reading a paper or a standalone research project. Some guidelines and project ideas will be posted later. Students are encouraged to apply the knowledge learned in this class to their application domains.

   - Team of 1-2 students
   - Mid-term project report due on October 14 (when classes resume after Fall Break)
   - Project presentation on November 18
   - Final project report due December 14

5. Materials: Lecture notes are available online `https://faculty.fuqua.duke.edu/~jx77/Course_graphs.html` and will be updated throughout the course. Additional reading materials will be posted online.

6. **Sept. 9 class will be in DeSanctis Classroom**.

7. **Oct. 14 class will be on Zoom**.

8. Classes will be video recorded using Zoom or Duke's Panopto system.