Summer School

High Dimensional Probability and Applications

Hanoi, June 19-23, 2023

PROGRAM

	Monday June 19	Tuesday June 20	Wednesday June 21	Thursday June 22	Friday June 23
8:45-9:00	Opening				
9:00-10:00	Van Vu (hybrid)	Joseph Neeman	Manjunath Krishnapur	Ke Wang	**One-day Mini Workshop (hybrid)**
10:00-10:30	Coffee	Coffee	Coffee	Coffee	
10:30-11:30	Van Vu (hybrid)	Manjunath Krishnapur	Joseph Neeman	Manjunath Krishnapur	
11:30-14:00	Lunch	Lunch	Lunch	Lunch	
14:00-15:00	Shahar Mendelson (online)	Ke Wang	Shahar Mendelson (online)	Shahar Mendelson (online)	
15:00-15:30	Coffee	Coffee	Coffee	Coffee	
15:30-16:30	Manjunath Krishnapur	Joseph Neeman	Ke Wang	Joseph Neeman	

Manjunath Krishnapur: *Random Jacobi matrices*

Shahar Mendelson: *Introduction to statistical learning*

Joseph Neeman: *Gaussian isoperimetric and related inequalities*

Van Vu: *Perturbation theory of low rank matrices*

Ke Wang: *Concentration Inequalities*

[all times are Hanoi time GMT+7]
[venue: Room 301, A5 building, IMH-VAST, 18 Hoang Quoc Viet St. Hanoi, Vietnam]
[Link zoom online: Meeting ID: 822 1267 6698 Passcode: 123456
https://us06web.zoom.us/j/82212676698?pwd=K2tVeFYzSVRBMkQ4dDYrVE1sbFpFQT09]

Mini workshop on June 23, 2023

[all times are Hanoi time GMT+7]
[venue: Room 301, A5 building, IMH-VAST, 18 Hoang Quoc Viet St. Hanoi, Vietnam]
[Link zoom online: Meeting ID: 822 1267 6698 Passcode: 123456
https://us06web.zoom.us/j/82212676698?pwd=K2tVeFYzSVRBMkQ4dDYrVE1sbFpFQT09]

Morning session

09:00–09:45 **Ke Wang** (Hong Kong University of Science and Technology, Hongkong)
Random perturbation of low-rank matrices

09:45–10:30 **Manjunath Krishnapur** (Indian Institute of Science, India)
A proof of the KMT theorems

10:30–10:50 **Coffee break**

10:50–11:35 **Yen Do** (University of Virginia, USA)
Real roots of random polynomials

11:35–12:00 **Quyet Nguyen** (Institute of Mathematics Hanoi, VAST)
On the the graph distance for supercritical percolation model

12:00 – 14:00 Lunch

Afternoon session

14:00–14:45 **Duy Khanh Trinh** (Waseda University, Japan)
Spectral statistics of adjacency matrices of random simplicial complexes

14:45–15:30 **Oanh Nguyen** (Brown University, USA)
The contact process on large networks

15:30–15:50 **Coffee break**

15:50–16:35 **Joseph Neeman** (Austin Texas University, USA)
Minimal clusters in Gaussian and other spaces

16:35–17:20 **Hoi Nguyen** (Ohio State University, USA)
Random matrices: gaps of the eigenvalues

ABSTRACTS

Lectures: *Random Jacobi matrices*

Manjunath Krishnapur
Indian Institute of Science

A Jacobi matrix is a symmetric tridiagonal matrix with positive entries on the super-diagonal. They represent discrete difference operators of the second order, and hence arise in many situations in mathematics and physics. It turns out that a particular family of random Jacobi matrices has a very explicit eigenvalue distribution, known as the beta-log gas in a quadratic potential on the real line. Special cases of this ensemble are more commonly introduced via well-known full matrix models such as the GUE and GOE. However, it turns out that the Jacobi random matrix is much more useful in studying the log-gas than the full matrix models.

Assuming only familiarity with graduate level probability and analysis, we give an introduction to this subject. Some of the topics we cover are: Limiting spectral distributions such as Wigner's semicircle law; Bounds for the largest eigenvalue; Stochastic domination results; Limiting distributions and deviation probabilities at finite *n*; Selberg integral; Asymptotic spacing between zeros of orthogonal polynomials; Use of the deviation bounds in a problem of LIL for last passage percolation.

References:

1) Random matrix theory lecture notes from a course

2) Holcomb and Virag A Short Introduction to Operator Limits of Random Matrices

3) Basu, Hegde, Ganguly, Krishnapur https://arxiv.org/abs/1909.01333 (on LIL in last passage percolation)

Lectures: *Introduction to Statistical Learning*

Shahar Mendelson
Australian National University

Lectures: *Gaussian isoperimetric and related inequalities*

Joseph Neeman
Austin Texas University

The Gaussian isoperimetric inequality is a geometric result with many applications in probability, especially in high-dimensional concentration. We will introduce the inequality and its applications, survey some related inequalities, and make some connections to other areas of mathematics and computer science. The course will consist of four lectures:

I. Gaussian isoperimetry and concentration: we will see (but not yet prove) the basic inequality and some applications to concentration in high dimensions.
II. Functional inequalities and gradient bounds: we will give a functional generalization of the basic inequality, a proof, and some connections to inequalities in analysis.
III. Noise stability and algorithms: we will see a "non-local" generalization of the basic inequality and an application in computer science.
IV. The geometric measure theory perspective: we'll return to the basic inequality, but seen from the perspective of geometric measure theory.

Lectures: *Perturbation theory of low rank matrices*

Van Vu
Yale University and VinBigdata Institute

Perturbation theory is an integral part of applied mathematics and statistics. One of the main goals if this is to theory is to answer the question: How much does a spectral parameter (such as the leading eigenvector or eigenvalue) of a matrix change, subjected to a perturbation to the entries?

Answers to this question, such as Weyl inequality or Davis-Kahan sine theorem, are among the most applied mathematical results, even more so in recent years when the role of statistics has become so dominant in all fields of science. (Think about Principal component analysis or Google Page rank computation, for instance.)

A pervasive assumption about real data (in form of a matrix) is that it has low rank. This is appearing many kinds of data, from movie preferences, text documents, to survey data, medical records, and genomics. This motivates us to build up a study of perturbation theory for low rank matrices, which I have been working on in the last 15 years.

In this mini-course, I will first give a brief survey to current state of the art of the project. Next, I am going to discuss applications in many fields: clustering, data completion (Netflix type problems), numerical linear algebra, learning mixtures of distributions, and privacy. There will be a number of directions for further research, and also several interesting connections to random matrix theory.

Lectures: *Concentration Inequalities*

Ke Wang
Hong Kong University of Science and Technology

Lecture 1. Gaussian random variables, Gaussian tail, sub-Gaussian random variables, classical concentration inequalities.
Lecture 2. Hanson-Wright inequality and applications.
Lecture 3. Matrix Bernstein's inequality.

Prerequisites: linear algebra and basic probability.

Random perturbation of low-rank matrices

Ke Wang
Hong Kong University of Science and Technology

The analysis of large matrices is a key aspect of high-dimensional data analysis, with computing the singular values and vectors of a matrix being a central task. However, real-world data is often disturbed by noise, which affects the essential spectral parameters of the matrix. While classical deterministic theorems can provide accurate estimates for the worst-case scenario, this talk will focus on the case when the perturbation is random. By assuming that the data matrix has a low rank, optimal subspace perturbation bounds can be achieved under mild assumptions.

This talk is based on joint works with Sean O'Rourke and Van Vu.

A proof of the KMT theorems

Manjunath Krishnapur
Indian Institute of Science

The Komlós-Major-Tusnády theorem for simple symmetric random walk asserts that up to n steps, its path can be coupled to stay within distance log(n) of a Brownian motion run for time n. A second KMT theorem says that the empirical distribution function of *n i.i.d.* uniform random variables on [0,1] can be coupled to stay within $\log(n)/\sqrt{n}$ distance of a Brownian bridge. Modifying and simplifying an idea of Chatterjee, we present a somewhat simplified proof of these theorems. The key step is to couple Binomial and hypergeometric distributions among themselves by coupling Markov chains with these as stationary distributions.

Real roots of random polynomials

Yen Do
University of Virginia

Understanding the location and estimating the number of real roots for generic polynomial equations of high degrees are central topics in many studies. In this talk, I would like to introduce some basic concepts and some basic questions related to these themes, and a brief survey of recent progress.

On the the graph distance for supercritical percolation model

Quyet Nguyen
Institute of Mathematics, VAST

Considering supercritical Bernoulli percolation on Z^d, Garet and Marchand proved a diffusive concentration for the graph distance. In this talk, we sharpen this result by establishing the subdiffusive concentration inequality. In addition, we also discuss on the regularity of the time constant of the graph distance.

Spectral statistics of adjacency matrices of random simplicial complexes

Duy Khanh Trinh
Waseda University

Simplicial complexes which may be viewed as a higher dimensional generalization of graphs have been playing an important role in topological data analysis recently. This talk gives a brief introduction on the spectral statistics of the adjacency matrix of random simplicial complexes. We focus on two models: (1) Linial--Meshulam complexes and (2) random geometric complexes which are the generalization of Erdos--Renyi graphs and of random geometric graphs, respectively.

This is based on a joint work with Shu Kanazawa (Kyoto University and The Ohio State University).

The contact process on large networks

Oanh Nguyen
Brown University

We discuss several popular models of the contact process including the SIR, SIS, SIRS models and some of their exciting properties, recent developments, and open problems. Based on several joint work with Danny Nam, Allan Sly and ongoing work with Souvik Dhara.

Minimal clusters in Gaussian and other spaces

Joseph Neeman
Austin Texas University

The Gaussian isoperimetric inequality states that if we want to partition R^n into two sets with prescribed Gaussian measure while minimizing the Gaussian surface area of the interface between the sets, then the optimal partition is obtained by cutting R^n with a hyperplane. We prove an extension to more than two parts. For example, the optimal way to partition R^3 into three parts involves cutting along three rays that meet at 120-degree angles at a common point. This is the Gaussian analogue of the famous Double Bubble Theorem in Euclidean space, which describes the shape that results from blowing two soap bubbles and letting them stick together. We will also discuss some related results in Euclidean space and on the sphere.

Joint work with Emanuel Milman.

Random matrices: gaps of the eigenvalues

Hoi Nguyen
Ohio State University

Gaps between consecutive eigenvalues are a central topic in random matrix theory. We will discuss the tail distribution of these gaps in various random matrix models, along with several potential applications.