

# FOCM 2017 - Workshop B5

## Random Matrices

---

B5 - July 13, 14:30 – 15:20

### SPECTRA OF RANDOM REGULAR AND QUASI-REGULAR GRAPHS

**Ioana Dumitriu**

University of Washington, USA  
dumitriu@uw.edu

The spectra of random graphs and networks have been widely studied in the last two decades, both in theoretical contexts (relating to random matrices and universality) and practical ones (with applications to mixing, sampling, community detection). From a random matrix perspective, the “natural” random graphs have independent edges (Erdos-Renyi-like); however, regular graphs have also been shown to share many of the spectral properties of the independent-edge ones. Recently, some of these studies have been extended to quasi-regular random graphs. We will survey some of these results and potential applications to community detection.

---

B5 - July 13, 15:30 – 15:55

### RANDOM MATRICES, M-ESTIMATORS AND THE BOOTSTRAP

**Noureddine El Karoui**

UC, Berkeley, USA  
nkaroui@berkeley.edu

Random-matrices-inspired ideas and techniques can be used to understand a wide variety of problems in high-dimensional statistics, way beyond properties of correlation and covariance matrices. In this talk, I’ll discuss how RMT is a fundamental tool in understanding the properties of widely used methods such as M-estimators and the bootstrap in high-dimensional statistics. Interestingly, the corresponding analyses upend conventional statistical thinking: maximum likelihood methods are woefully inefficient (even among restricted classes of estimators) and the bootstrap typically fails to give statistically valid confidence statements, even for very simple problems. I’ll also discuss limitations of random matrix models for statistical applications.

*Joint work with Elizabeth Purdom (UC, Berkeley).*

---

B5 - July 13, 16:00 – 16:25

### STEIN’S METHOD FOR RANDOM MATRICES

**Elizabeth Meckes**

Case Western Reserve University, USA  
ese3@case.edu

Most of the random variables of interest which arise in random matrix theory come from random matrix ensembles in which dependence plays an important role. Stein’s method is a powerfully robust tool

for studying limiting behavior in such situations. I will describe a few applications of Stein’s method in random matrix theory, for distributional results (including joint work with S. Chatterjee), and for obtaining concentration inequalities (work of Chatterjee, Tropp–Mackey, and others).

---

B5 - July 13, 17:00 – 17:25

SOME RESULTS ON THE EIGENVECTORS OF RANDOM SYMMETRIC MATRICES

**Ke Wang**

Hong Kong University of Science and Technology, Hong Kong  
kewang@ust.hk

Eigenvectors of large matrices and graphs play an essential role in combinatorics and theoretical computer science. For instance, many properties of a graph can be deduced or estimated from its eigenvectors. It is conjectured that an eigenvector of a random symmetric matrix behaves like a random vector uniformly distributed on the unit sphere. I will talk about some recent partial results toward confirming this conjecture.

*Joint work with Sean O’Rourke (University of Colorado Boulder) and Van Vu (Yale University).*

---

B5 - July 13, 17:30 – 17:55

OPTIMALITY OF THE JOHNSON-LINDENSTRAUSS LEMMA

**Jelani Nelson**

Harvard University, USA  
minilek@seas.harvard.edu

One popular application of random matrices is the “random projections” tool offered by the Johnson-Lindenstrauss lemma. This lemma states that for any  $n, d > 1$  and any  $0 < \varepsilon < 1/2$ , and for any  $X \subset \ell_2^d$  of size  $n$ , there is a map  $f : X \rightarrow \ell_2^m$  with  $m = O(\varepsilon^{-2} \log n)$  with distortion  $1 + \varepsilon$ . All known proofs of this lemma operate by defining  $f(x) = \Pi x$  for some random matrix  $\Pi$ , with the distribution being used changing from proof to proof.

We show that this “random projection” method offered by the JL lemma is optimal for Euclidean dimensionality reduction. That is, for every  $n, d, \varepsilon$  as above where  $\varepsilon$  is not too small, there exists a set  $X \subset \ell_2^d$  with  $|X| = n$  such that any  $1 + \varepsilon$  distortion of embedding (even non-linear ones) of  $X$  into  $\ell_2^m$  must have  $m = \Omega(\varepsilon^{-2} \log n)$ .

*Joint work with Kasper Green Larhus (Aarhus University).*

---

B5 - July 14, 14:30 – 14:55

FINITE SIZE EFFECTS FOR SPACING DISTRIBUTIONS IN RANDOM MATRIX THEORY: CIRCULAR ENSEMBLES AND RIEMANN ZEROS

**Folkmar Bornemann**

Technical University of Munich, Germany  
bornemann@tum.de

According to Dyson's three fold way, from the viewpoint of global time reversal symmetry there are three circular ensembles of unitary random matrices relevant to the study of chaotic spectra in quantum mechanics. These are the circular orthogonal, unitary and symplectic ensembles, denoted COE, CUE and CSE respectively. For each of these three ensembles and their thinned versions, whereby each eigenvalue is deleted independently with probability  $1 - \xi$ , we take up the problem of calculating the first two terms in the scaled large  $N$  expansion of the spacing distributions. It is well known that the leading term admits a characterisation in terms of both Fredholm determinants and Painlevé transcendents. We show that modifications of these characterisations also remain valid for the next to leading term, and that they provide schemes for high precision numerical computations. In the case of the CUE there is an application to the analysis of Odlyzko's data set for the Riemann zeros, and in that case some further statistics are similarly analyzed.

*Joint work with Peter J. Forrester (University of Melbourne, Australia) and Anthony Mays (University of Melbourne, Australia).*

---

B5 - July 14, 15:00 – 15:25

### SELF-SIMILARITY IN THE SPECTRA OF RANDOM UNITARY MATRICES

**Mark Meckes**

Case Western Reserve University, USA  
mark.meckes@case.edu

I will present a rigorous result roughly stating that, for a range of mesoscopic scales, the eigenvalues of an  $n \times n$  random unitary matrix are statistically indistinguishable from those of a  $2n \times 2n$  matrix, suitably rescaled. This result is inspired by a conjecture made by Coram and Diaconis in a statistical study of the relationship between eigenvalues of large random unitary matrices and zeroes of the Riemann zeta function.

*Joint work with Elizabeth Meckes (Case Western Reserve University).*

---

B5 - July 14, 15:30 – 16:20

### VARIATIONS ON PCA

**Amit Singer**

Princeton University, United States  
amits@math.princeton.edu

Principal component analysis, or PCA, is a data analysis method used in nearly every scientific discipline. It allows the scientist to discover the main directions of variability in a dataset, known as the principal components. The principal components are used as factors to model the data, denoise and compress the data, and for visualization, clustering, or any other further analysis of the data.

Despite enormous progress in recent years on PCA in high dimensions, the theory and methods developed until now are mostly limited to simple applications involving the sample covariance matrix. Current methods for PCA are either statistically suboptimal or are not computationally scalable for data that is corrupted by non-additive, non-Gaussian random effects, as is routinely encountered in problems involving missing data, image deconvolution, and Poissonian noise.

In this talk we will discuss two variants of PCA for which we recently developed new statistical and computational frameworks: 1) PCA for exponential family distributions, and 2) PCA from noisy linearly reduced measurements.

We will discuss applications to cryo-electron microscopy, X-ray free electron lasers, low-rank matrix completion, and noisy deconvolution that motivated this work.

*Joint work with Lydia Liu (Princeton University), William Leeb (Princeton University), Edgar Dobriban (Stanford), Joakim Anden (Princeton University), Tejal Bhamre (Princeton University), Teng Zhang (University Central Florida).*

---

B5 - July 14, 17:00 – 17:25

## CONCENTRATION FOR COULOMB GASES AND COULOMB TRANSPORT INEQUALITIES

**Djalil Chafaï**

Université Paris-Dauphine, France

djalil@chafai.net

This talk will present a recent joint work with Mylène Maïda and Adrien Hardy on the non-asymptotic behavior of Coulomb gases in dimension two and more. Such gases are modeled by an exchangeable Boltzmann-Gibbs measure with a singular two-body interaction. We obtain concentration of measure inequalities for the empirical distribution of such gases around their equilibrium measure, with respect to bounded Lipschitz and Wasserstein distances. This implies macroscopic as well as mesoscopic convergence in such distances. In particular, we improve the concentration inequalities known for the empirical spectral distribution of Ginibre random matrices. Our approach is remarkably simple and bypasses the use of renormalized energy. It crucially relies on new inequalities between probability metrics, including Coulomb transport inequalities which can be of independent interest.

*Joint work with Mylène Maïda (Université Lille 1, France) and Adrien Hardy (Université Lille 1, France).*

---

B5 - July 14, 17:30 – 17:55

## EIGENVALUE ATTRACTION

**Ramis Movassagh**

IBM TJ Watson Research Center, USA

q.eigenman@gmail.com

We prove that the complex conjugate (c.c.) eigenvalues of a smoothly varying real matrix attract. We offer a dynamical perspective on the motion and interaction of the eigenvalues in the complex plane, derive their governing equations and discuss applications. C.c. pairs closest to the real axis, or those that are ill-conditioned, attract most strongly and can collide to become exactly real. We apply the results to various special cases including the Hatano-Nelson model and provide various numerical illustrations. It is our hope that the simple dynamical perspective herein might help better understanding of the aggregation and low density of the eigenvalues of real random matrices on and near the real line respectively.

---

B5 - July 15, 14:30 – 14:55

## THE SMALLEST SINGULAR VALUE OF ADJACENCY MATRICES OF RANDOM REGULAR DIGRAPHS

**Konstantin Tikhomirov**  
Princeton University, USA  
kt12@math.princeton.edu

Let  $A_n$  be the adjacency matrix of a random  $d$ -regular directed graph on  $n$  vertices, distributed uniformly on the set of all  $d$ -regular digraphs without multiple edges. By combining a specially developed chaining method with linear algebraic arguments and Littlewood-Offord-type anti-concentration inequalities, we derive a lower bound for the smallest singular value of matrices of the form  $A_n - z\text{Id}_n$ . Our result gives a bound  $s_{\min} \geq n^{-C}$  with probability at least  $1 - d^{-c}$ , for any  $d$  larger than an absolute constant and less than  $n/\log n$ . This result can be viewed as a step towards proving the circular law for uniform random digraphs in the “sparse” regime, complementing a recent result of Cook who established the circular law for  $d$  at least polylogarithmic in  $n$ .

*Joint work with A. E. Litvak (Alberta), A. Lytova (Alberta), N. Tomczak-Jaegermann (Alberta) and P. Youssef (Paris-Diderot).*

---

B5 - July 15, 15:00 – 15:25

#### FREE COMPONENT ANALYSIS

**Raj Rao**  
University of Michigan , USA  
Rajnrao@umich.edu

We describe a method for unmixing mixtures of ‘freely’ independent random variables in a manner analogous to the independent component analysis (ICA) based method for unmixing independent random variables from their additive mixture. Random matrices play the role of free random variables in this context so the method we develop, which we call Free component analysis (FCA), unmixes matrices from an additive mixture of matrices. We describe the theory – the various ‘contrast functions’, computational methods and compare FCA to ICA on data derived from real-world experiments.

*Joint work with Hao Wu (University of Michigan).*

---

B5 - July 15, 15:30 – 15:55

#### UNIVERSALITY IN NUMERICAL COMPUTATIONS WITH RANDOM DATA

**Thomas Trogdon**  
University of California, Irvine, USA  
ttrogdon@math.uci.edu

This talk will concern recent progress on the statistical analysis of numerical algorithms with random initial data. In particular, with appropriate randomness, the fluctuations of the iteration count (halting time) of numerous numerical algorithms have been demonstrated to be universal, i.e., independent of the distribution on the initial data. This phenomenon has given new insights into random matrix theory. Furthermore, recent estimates from random matrix theory allow for fluctuation limit theorems for simple algorithms and halting time estimates for others.

*Joint work with Percy Deift (New York University) and Govind Menon (Brown University).*

---

B5 - July 15, 16:00 – 16:25

**Alan Edelman**

MIT, USA

edelman@mit.edu

Over the many years of reading random matrix papers, it has become increasingly clear that the phenomena of random matrix theory can be difficult to understand in the absence of numerical codes to illustrate the phenomena. (We wish we could require that all random matrix papers that lend themselves to computing include a numerical simulation with publicly available code.) Of course mathematics exists without numerical experiments, and all too often a numerical experiment can be seen as an unnecessary bother. On a number of occasions, however, the numerical simulations themselves have an interesting twist of their own. This talk will illustrate a few of those simulations and illustrate why in particular the Julia computing language is just perfect for these simulations. Some topics we may discuss:

1. “Free” Dice
2. Tracy Widom ( $O(n^{1/3})$ )
3. Smallest Singular Value (High Parallel Monte Carlo)
4. Jacobians of Matrix Factorizations (Autodiff: Derivatives w/o Symbolic nor Finite Derivatives)

*Joint work with Bernie Wang (Amazon).*

---

B5 - Poster

NEW EXPRESSIONS FOR THE EXTREME EIGENVALUES OF  $\beta$  RANDOM MATRICES.

**Plamen Koev**

San Jose State University, United States of America

plamen.koev@sjsu.edu

We present new identities for the hypergeometric function of a matrix argument of parameter  $\beta$ , which lead to new expressions for the extreme eigenvalues of the  $\beta$  random matrix ensembles – Laguerre, Wishart, and MANOVA.

The new expressions have lower computational complexity than the existing ones. All known expressions are in terms of the hypergeometric function of a matrix argument, an infinite series of Jack functions summed over all integer partitions. Using algebraic identities for the hypergeometric function and we are able to reduce the set of partitions over which the summation needs to take place and in certain cases (e.g., the trace of the  $\beta$  Wishart matrix) derive new expressions which require the summation over partitions of just one part.

These new expressions allow for substantial computational savings that are at least an order of magnitude faster and more than that for “spiked” covariances.

*Joint work with Raymond Kan, University of Toronto.*

---