Polynomial Techniques in Combinatorial Linear Algebra

Adam W. Marcus

Princeton University
adam.marcus@princeton.edu

Triangle Lectures
March 30, 2019

Frequent coauthors (MSS):

Dan Spielman
Yale University

Nikhil Srivastava
University of California, Berkeley

My involvement partially supported by:

NSF Postdoctoral Research Fellowship

NSF CAREER Grant DMS-1552520

Outline

What is this “Combinatorial Linear algebra?”

Combinatorics: What properties can collections of “things” have?

Linear algebra: “things” = vectors or matrices

Focus of this talk: eigenvalues of symmetric* matrices

What is this “Combinatorial Linear algebra?”

Combinatorics: What properties can collections of “things” have?

Linear algebra: “things” = vectors or matrices

Focus of this talk: eigenvalues of symmetric* matrices

Elements of

- Random matrix theory
- Geometric combinatorics
- Real algebraic geometry**
- Free probability**

A fundamental question

Given (real) symmetric matrix

$$A = R \begin{bmatrix} \text{-}2 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 7 \end{bmatrix} R^T$$

and a collection of symmetric matrices

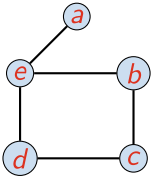
$$B_k = Q_k \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} Q_k^T$$

(R, Q_k orthogonal matrices).

Goal: minimize $\lambda_{max}(A + B_k)$

Example: Spectral Graph Theory

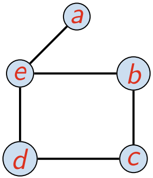
A is the Laplacian matrix of a graph, B_k Laplacian matrices of edges:



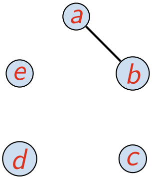
$$\begin{bmatrix} 1 & \cdot & \cdot & \cdot & -1 \\ \cdot & 2 & -1 & \cdot & -1 \\ \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & -1 & 2 & -1 \\ -1 & -1 & \cdot & -1 & 3 \end{bmatrix}$$

Example: Spectral Graph Theory

A is the Laplacian matrix of a graph, B_k Laplacian matrices of edges:



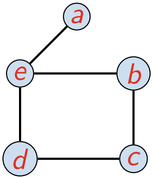
$\cup$



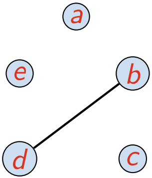
$$\begin{bmatrix} 1 & \cdot & \cdot & \cdot & -1 \\ \cdot & 2 & -1 & \cdot & -1 \\ \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & -1 & 2 & -1 \\ -1 & -1 & \cdot & -1 & 3 \end{bmatrix} + \begin{bmatrix} 1 & -1 & \cdot & \cdot & \cdot \\ -1 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Example: Spectral Graph Theory

A is the Laplacian matrix of a graph, B_k Laplacian matrices of edges:



$\bigcup$



$$\begin{bmatrix} 1 & \cdot & \cdot & \cdot & -1 \\ \cdot & 2 & -1 & \cdot & -1 \\ \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & -1 & 2 & -1 \\ -1 & -1 & \cdot & -1 & 3 \end{bmatrix} + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Example: Spectral Graph Theory

A is the Laplacian matrix of a graph, B_k Laplacian matrices of edges:

$$\begin{bmatrix} 1 & \cdot & \cdot & \cdot & -1 \\ \cdot & 2 & -1 & \cdot & -1 \\ \cdot & -1 & 2 & -1 & \cdot \\ \cdot & \cdot & -1 & 2 & -1 \\ -1 & -1 & \cdot & -1 & 3 \end{bmatrix} + \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot & 1 \end{bmatrix}$$

Can we guarantee some edge is “good”?

Probabilistic ansantz

If $x_1, \ldots, x_n$ are real numbers, then $\min_k\{x_k\} \leq \overline{x_k}$.

Compute an (easier?) expectation, get a bound.

Probabilistic ansantz

If $x_1, \ldots, x_n$ are real numbers, then $\min_k\{x_k\} \leq \overline{x_k}$.

Compute an (easier?) expectation, get a bound.

We are interested in $\min_k \; \{ \; \lambda_{max}(A + B_k) \; \}$ so look at...

Probabilistic ansantz

If $x_1, \ldots, x_n$ are real numbers, then $\min_k\{x_k\} \leq \overline{x_k}$.

Compute an (easier?) expectation, get a bound.

We are interested in $\min_k \ \{ \ \lambda_{max}(A + B_k) \ \}$ so look at...

1. $\overline{\lambda_{max}(A + B)}$?

Probabilistic ansantz

If $x_1, \ldots, x_n$ are real numbers, then $\min_k\{x_k\} \leq \overline{x_k}$.

Compute an (easier?) expectation, get a bound.

We are interested in $\min_k \; \{ \; \lambda_{max}(A + B_k) \; \}$ so look at...

1. $\overline{\lambda_{max}(A + B)}$?
2. $\lambda_{max}(A + \overline{B})$?

For example:

$$\mathbb{E}_{Q \in O(3)} \left\{ Q^T \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} Q \right\} = \begin{bmatrix} 1/3 & \cdot & \cdot \\ \cdot & 1/3 & \cdot \\ \cdot & \cdot & 1/3 \end{bmatrix}$$

Probabilistic ansantz??

Let $\chi_A(x) = \det(x\mathbb{I} - A)$ be the *characteristic polynomial* of A, then

$$\lambda_{max}(A) = \text{maxroot}\left\{\chi_A(x)\right\}.$$

1. $\overline{\text{maxroot}\left\{\chi_{A+B}(x)\right\}}$?

2. $\text{maxroot}\left\{\chi_{A+\overline{B}}(x)\right\}$?

Probabilistic ansantz??

Let $\chi_A(x) = \det(x\mathbb{I} - A)$ be the *characteristic polynomial* of A, then

$$\lambda_{max}(A) = \text{maxroot}\left\{\chi_A(x)\right\}.$$

1. $\overline{\text{maxroot}\left\{\chi_{A+B}(x)\right\}}$?
2. $\text{maxroot}\left\{\chi_{A+\overline{B}}(x)\right\}$?
3. $\text{maxroot}\left\{\overline{\chi_{A+B}(x)}\right\}$ (an average of polynomials!)?

   Well defined?

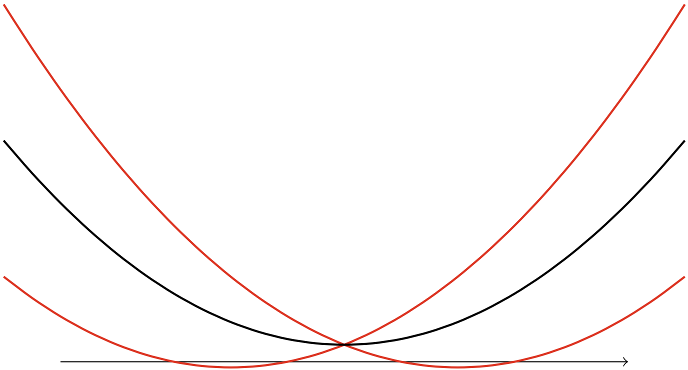
   Computable?

   Relevant?

Outline

Well-defined?

In general, real-rootedness is not closed under addition.



What is maxroot {} of something with complex roots?

Example

For A in the Graph Laplacian example,

$$\chi_{A+B_1} = x^5 - 12x^4 + 51x^3 - 90x^2 + 55x$$

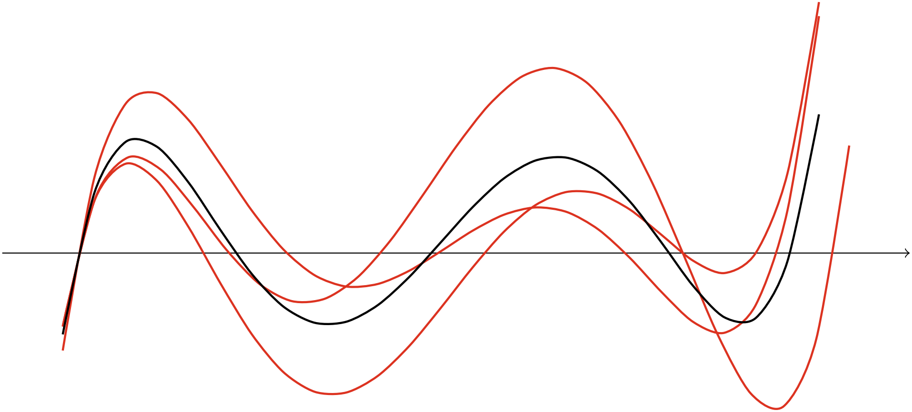
$$\chi_{A+B_2} = x^5 - 12x^4 + 50x^3 - 82x^2 + 40x$$

$$\chi_{A+B_3} = x^5 - 12x^4 + 49x^3 - 78x^2 + 40x$$

so

$$\frac{1}{3}\chi_{A+B_1} + \frac{1}{3}\chi_{A+B_2} + \frac{1}{3}\chi_{A+B_3} = x^5 - 12x^4 + 50x^3 - 83\frac{1}{3}x^2 + 45x$$

Plotted



Lemma (Root Separation Lemma)

Let $\{p_i\}$ *be degree* d *polynomials and let* $[s, t] \subseteq \mathbb{R}$ *an interval such that*

- $p_i(s) \geq 0$ *for all* i
- $p_i(t) \leq 0$ *for all* i
- *Each* p_i *has exactly one real root in* $[s, t]$.

Then every convex combination of $\{p_i\}$ *has exactly one real root in* $[s, t]$ *and it lies between the roots of some* p_a *and* p_b.

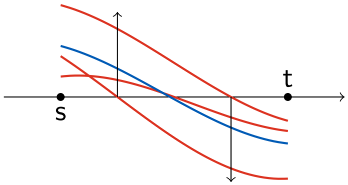
Lemma (Root Separation Lemma)

Let $\{p_i\}$ be degree d polynomials and let $[s, t] \subseteq \mathbb{R}$ an interval such that

- $p_i(s) \geq 0$ *for all* i
- $p_i(t) \leq 0$ *for all* i
- *Each* p_i *has exactly one real root in* $[s, t]$.

Then every convex combination of $\{p_i\}$ has exactly one real root in $[s, t]$ and it lies between the roots of some p_a and p_b.

Proof by picture:



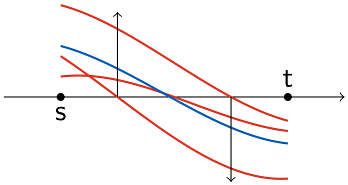
Lemma (Root Separation Lemma)

Let $\{p_i\}$ *be degree* d *polynomials and let* $[s, t] \subseteq \mathbb{R}$ *an interval such that*

- $p_i(s) \geq 0$ *for all* i
- $p_i(t) \leq 0$ *for all* i
- *Each* p_i *has exactly one real root in* $[s, t]$.

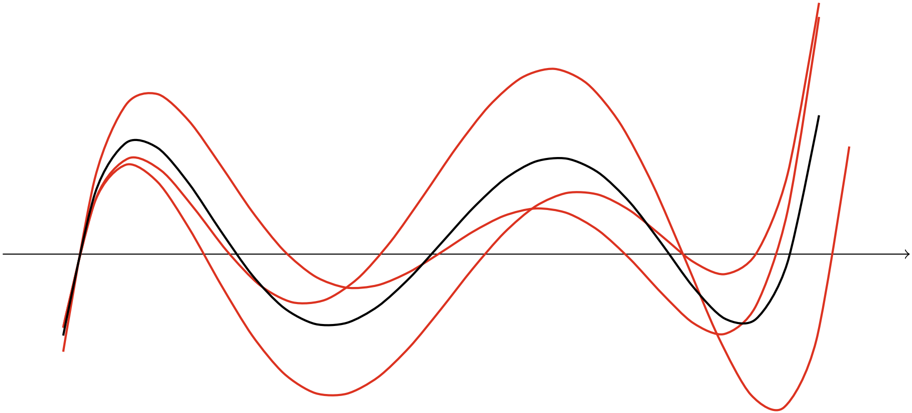
Then every convex combination of $\{p_i\}$ *has exactly one real root in* $[s, t]$ *and it lies between the roots of some* p_a *and* p_b.

Proof by picture:

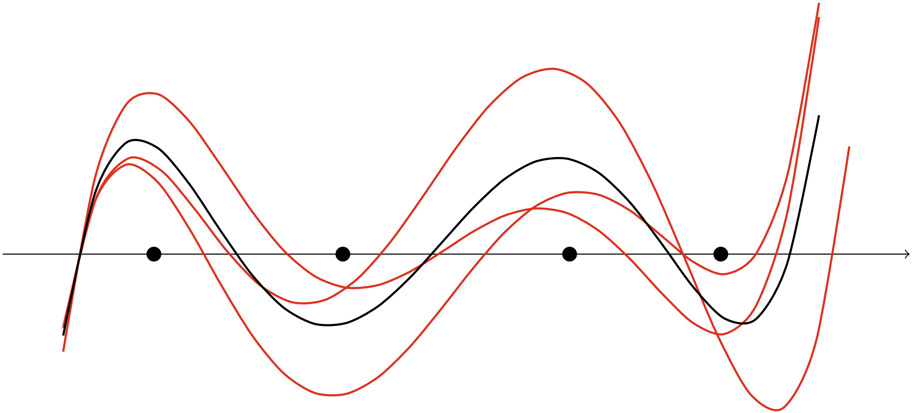


Note: if there are d such intervals, then $\sum_i p_i$ is real rooted.

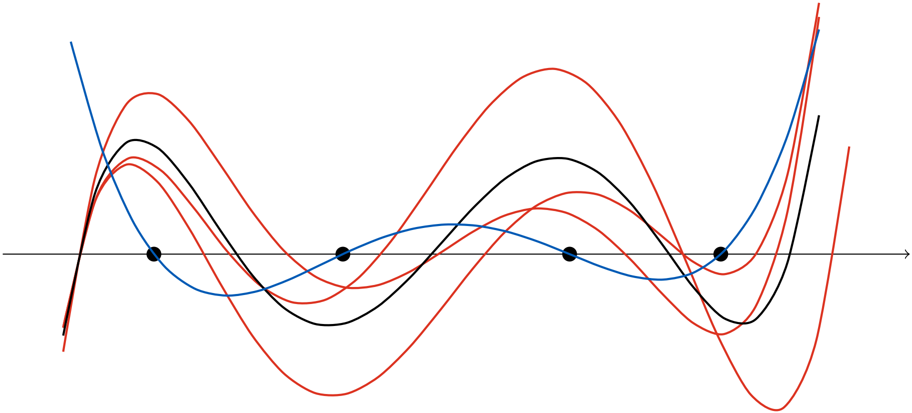
Plotted, again



Plotted, again



Plotted, again



d separating intervals $\Longleftrightarrow$ common interlacer

Corollary

If $\{p_k\}$ *are degree* d *real-rooted polynomials with a common interlacer:*

1. *Every convex combination* $\sum_k \alpha_k p_k$ *is real-rooted*
2. *There exist* i, j *such that* $\text{kth-root}\{p_i\} \leq \text{kth-root}\{\overline{p}\} \leq \text{kth-root}\{p_j\}$.

In particular, there exists an i such that $\text{maxroot}\,\{p_i\} \leq \text{maxroot}\,\{\overline{p}\}$.

Corollary

If $\{p_k\}$ are degree d real-rooted polynomials with a common interlacer:

1. *Every convex combination $\sum_k \alpha_k p_k$ is real-rooted*
2. *There exist i, j such that $\text{kth-root}\{p_i\} \leq \text{kth-root}\{\overline{p}\} \leq \text{kth-root}\{p_j\}$.*

In particular, there exists an i such that $\text{maxroot}\{p_i\} \leq \text{maxroot}\{\overline{p}\}$.

Theorem (Cauchy Interlacing)

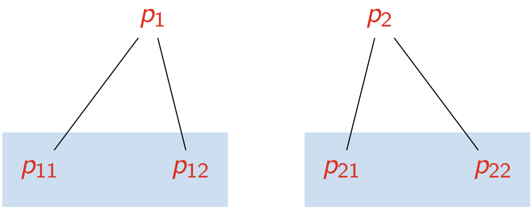
If $B_1, \dots, B_n$ satisfy $\text{rank}(B_i - B_j) \leq 2$ for all i, j, then the polynomials $p_i = \chi_{A+B_i}$ have a common interlacer.

What about $A + B_i + C_j$?

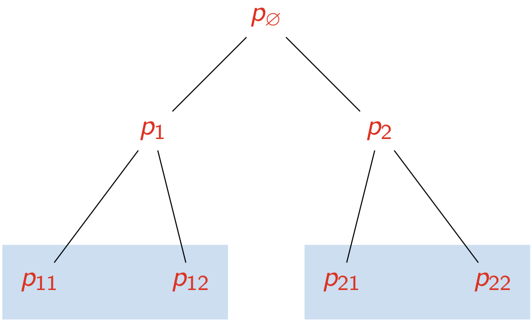
Iterate?

Let $B_i + C_j = D_{ij}$?

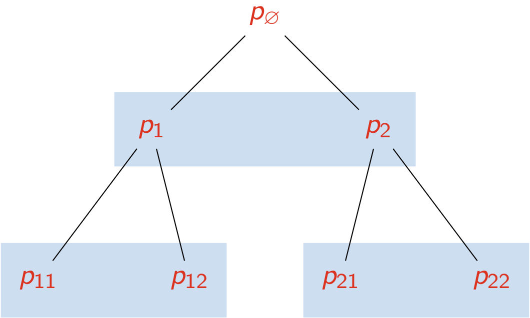
Think quantum (and iterate)



Think quantum (and iterate)



Think quantum (and iterate)



A rooted, connected tree of real-rooted polynomials where

1. Each parent is a convex combination of its children
2. Each set of children has a common interlacer

is called an *interlacing family*.

Putting it all together

Given symmetric matrices $C_1, \ldots, C_k$

1. Form $p_i = \chi_{C_i}$
2. (Try to) build an interlacing family on top of the p_i

There exist i such that $\lambda_{max}(C_i) \leq \text{maxroot}\{p_{top}\}$.

Theorem (MSS (2013) + Cohen(2015))

Let A be a symmetric matrix and let $\hat{B}_1, \ldots, \hat{B}_k$ be independent "random" symmetric matrices. Then there exists a matrix $Y \in \text{supp}(A + \hat{B}_1 + \cdots + \hat{B}_k)$ such that

$$\lambda_{max}(Y) \leq \text{maxroot}\left\{\mathbb{E}\left\{\chi_{A+\hat{B}_1,\ldots,\hat{B}_k}(x)\right\}\right\}$$

(called the "mixed characteristic polynomial").

Computing $\text{maxroot}\{\}$?

Outline

Picking vectors

Theorem (MSS (2013))

Let $v_1, \ldots, v_n \in \mathbb{R}^d$ *be a collection of vectors with*

$$\sum v_i v_i^T = \mathbb{I}$$

For all $k \leq d$, *there exists a set* $S \subseteq [n]$ *with* $|S| = k$ *such that*

$$\lambda_{max} \left(\sum_{i \in S} v_i v_i^T \right) \leq \left(1 + \sqrt{\frac{k}{d}} \right)^2 \left(\frac{d}{n} \right)$$

$\frac{k}{d}$: fraction of dimensions you want to span

$\frac{d}{n}$: average length of the v_i

Partitioning vectors

Theorem (MSS (2013))

Let $v_1, \ldots, v_n \in \mathbb{R}^d$ be a collection of vectors with

$$\sum v_i v_i^T = \mathbb{I} \qquad \text{and} \qquad \|v_i\|^2 \leq \epsilon.$$

For all $r > 1$, there exists a partition $\{S_1, S_2, \ldots, S_r\}$ so that

$$\lambda_{max}\left(\sum_{i \in S_k} v_i v_i^T\right) \leq \frac{1}{r}\left(1 + \sqrt{r\epsilon}\right)^2$$

for all k.

Resolves (via a result of Weaver) question of Kadison and Singer.

Asymptotically tight (as $d \to \infty$), witnessed by Marchenko–Pastur.

Aside

Previous best used matrix concentration:

$$\lambda_{max}\left(\sum_{i\in S_j} v_i v_i^T\right) \leq f(r,\delta)\frac{\log d}{\log\log d}$$

Theorem (Matrix Chernoff/Bernstein/Hoeffding/etc)

If $\hat{A}_1, \ldots, \hat{A}_n \in \mathbb{R}^{d\times d}$ are independent random self adjoint matrices then

$$\mathbb{P}\left[\lambda_{max}\left(\sum \hat{A}_i > t\right)\right] \leq d \cdot e^{-f(t,\hat{A}_1,\ldots,\hat{A}_n)}.$$

Claim: any sufficiently generic concentration ("high" probability) bound must depend on d.

The bad seed

Consider d copies of the diagonal matrices

$$\left\{ \begin{bmatrix} 1/d & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & \\ & 1/d & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1/d \end{bmatrix} \right\}$$

The bad seed

Consider d copies of the diagonal matrices

$$\left\{ \begin{bmatrix} 1/d & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & & & \\ & 1/d & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1/d \end{bmatrix} \right\}$$

Balls in bins

- d identical balls
- d different bins, chosen uniformly and independently
- X_i = the number of balls in bin i

Well known fact:

$$\mathbb{E}\left\{\max_i X_i\right\} = \Theta\left(\frac{\log d}{\log\log d}\right).$$

If we want asymptotic tightness:

$$\mathbb{P}\left[\max_i X_i = 1\right] = \frac{d!}{d^d} \approx \frac{1}{e^d}.$$

Any successful method needs to find small probability events.

Fractional approximations of operators

Theorem (Akemann, Weaver (2013))

Let $v_1, \ldots, v_n \in \mathbb{R}^d$ *be a collection of vectors with*

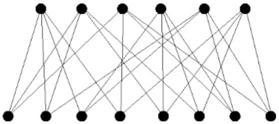
$$\sum v_i v_i^T \leq \mathbb{I} \qquad \text{and} \qquad \|v_i\|^2 \leq \epsilon$$

and let $0 \leq p_i \leq 1$*. There exist values* $\eta_1, \ldots, \eta_n \in \{0, 1\}$ *such that*

$$\left\| \sum_{i=1}^{n} \eta_i v_i v_i^T - \sum_{i=1}^{n} p_i v_i v_i^T \right\| = O(\epsilon^{1/8})$$

Integrality gaps in semi-definite programming?

Biregular bipartite Ramanujan graphs



Theorem (Gribinski, M (2017))

For all integers n, k, d, there exists bipartite graph $G = U \cup V$ such that

- $|U| = n$ *and* $|V| = kn$
- U *is* kd*-regular, and* V *is* d*-regular*
- $\lambda_2(A) \leq \sqrt{d-1} + \sqrt{kd-1}$

where A is the adjacency matrix of G.

Furthermore, it is constructible in time $\text{poly}(n)$ *for fixed k, d.*

Asymptotically optimal spectral expanders.

Partitioning into biregular Ramanujan graphs

Theorem (MSS (2013) + Boche–Wiese (2018))

For all integers n, d_1, d_2, there exists a coloring of the complete bipartite graph $K_{2^n d_1, 2^n d_2}$ using 2^n colors such that the adjacency matrix of every color class A_c satisfies

$$\lambda_2(A_c) \leq \sqrt{d_1 - 1} + \sqrt{d_2 - 1}.$$

Useful in building asymptotically optimal semantically secure codes.

All color classes need to act like hash functions.

Not constructive.

Outline

Finite free probability

For $d \times d$ symmetric matrices A, B,

$$p(x) = \det(x\mathbb{I} - A) = \prod_i (x - a_i) \quad \text{and} \quad q(x) = \det(x\mathbb{I} - B) = \prod_i (x - b_i)$$

Additive convolution:

$$\begin{aligned}[p \boxplus q](x) &= \mathbb{E}_{Q \in O(d)} \left\{ \det(x\mathbb{I} - A - QBQ^T) \right\} \\ &= \mathbb{E}_{\pi \in S(d)} \left\{ \prod_i (x - a_i - b_{\pi(i)}) \right\}\end{aligned}$$

- Linear in coefficients of p and q
- Preserves real rootedness! (Borcea–Brändén)
- Gives unitarily invariant algebra of eigenvalues

Connection to Voiculescu Free Probability

Theorem (M, 2017)

Let A, B be symmetric matrices with (discrete) eigenvalue distributions μ_A, μ_B. Set

$$p(x) = \det(x\mathbb{I} - A) \qquad \text{and} \qquad q(x) = \det(x\mathbb{I} - B)$$

and let μ_{A+B} be the free convolution of $\mu(A)$ and $\mu(B)$.

1. $\lambda[p^n \boxplus q^n] \xrightarrow{D} \mu_{A+B}$
2. $\operatorname{conv}\operatorname{supp}\{\lambda[p^n \boxplus q^n]\} \subseteq \operatorname{conv}\operatorname{supp}\{\mu_{A+B}\}$

Connection to Voiculescu Free Probability

Theorem (M, 2017)

Let A, B be symmetric matrices with (discrete) eigenvalue distributions μ_A, μ_B. Set

$$p(x) = \det(x\mathbb{I} - A) \qquad \textit{and} \qquad q(x) = \det(x\mathbb{I} - B)$$

and let μ_{A+B} be the free convolution of $\mu(A)$ and $\mu(B)$.

1. $\lambda[p^n \boxplus q^n] \xrightarrow{D} \mu_{A+B}$
2. $\operatorname{conv}\operatorname{supp}\{\lambda[p^n \boxplus q^n]\} \subseteq \operatorname{conv}\operatorname{supp}\{\mu_{A+B}\}$

Edge of spectrum of μ_{A+B} is an upper bound on maxroot $\{p \boxplus q\}$.

Explicit bounds (MSS, 2015).

Further work:

1. Multiplicative convolution
2. Additive Brownian motion
3. Central Limit Theorem

With A. Gribinski:

1. Rectangular matrices (singular values)
2. Entropy, Fisher information

With B. Mirabelli:

1. Multiplicative Brownian motion, CLT
2. Non-Hermitian square matrices

Hyperbolic manifolds

Conjecture (M, Sarnak)

There exists an infinite sequence of compact hyperbolic manifolds $M_1, M_2, \ldots,$ *for which the Laplacians* $L_1, L_2, \ldots,$ *satisfy*

$$\lambda_2(L_i) \geq \frac{1}{4}.$$

"Analogue" of Ramanujan graphs:

1. $1/4 =$ spectral radius of the universal cover
2. $\lambda_1(L) = 0$ for all compact manifolds.

Method of interlacing zeta functions?

Random Matrix Theory

Ansatz (M)

Let $X(\beta)$ be a statement about random matrices for which

- $X(1)$ *is true over* $\mathbb{R}$
- $X(2)$ *is true over* $\mathbb{C}$
- $X(4)$ *is true over* $\mathbb{H}$

Then $\lim_{\beta \to \infty} X(\beta)$ *is true in finite free probability.*

Known:

1. Additive, multiplicative Brownian motions (Tao)
2. β-corners process (with V. Gorin)

Jack polynomials?

Thank you for your attention!