Two proofs of the existance of Ramanujan graphs

Spectral Algorithms Workshop, Banff

Adam W. Marcus

Princeton University
adam.marcus@princeton.edu

August 1, 2016

Acknowledgements:

Joint work with:

Dan Spielman
Yale University

Nikhil Srivastava
University of California, Berkeley

Outline

Motivation

In spectral graph theory, we are interested in eigenvalues of some matrix (Laplacian, adjacency, etc).

Since graphs are a union of edges, there is a natural way to think of these matrices as a sum of smaller matrices.

Motivation

In spectral graph theory, we are interested in eigenvalues of some matrix (Laplacian, adjacency, etc).

Since graphs are a union of edges, there is a natural way to think of these matrices as a sum of smaller matrices.

Typically, we can figure out how the eigenvalues change when adding a fixed smaller matrix, but many times one might want to add edges to a graph *randomly*.

Motivation

In spectral graph theory, we are interested in eigenvalues of some matrix (Laplacian, adjacency, etc).

Since graphs are a union of edges, there is a natural way to think of these matrices as a sum of smaller matrices.

Typically, we can figure out how the eigenvalues change when adding a fixed smaller matrix, but many times one might want to add edges to a graph *randomly*.

In this talk, I will discuss the idea of adding randomly chosen matrices to other matrices and then show how the idea can be used to show existance of Ramanujan graphs in two ways.

Well-known techniques exist for bounding the eigenvalues of random sums of matrices.

Theorem (Matrix Chernoff, for example)

Let $\widehat{v}_1, \dots, \widehat{v}_n$ *be independent random vectors with* $\|\widehat{v}_i\| \leq 1$ *and* $\sum_i \widehat{v}_i \widehat{v}_i^T = \widehat{V}$*. Then*

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \leq \theta\right] \geq 1 - d \cdot e^{-nD(\theta \| \lambda_{max}(\mathbb{E}\widehat{V}))}$$

Similar inequalities by Rudelson (1999), Ahlswede–Winter (2002).

Well-known techniques exist for bounding the eigenvalues of random sums of matrices.

Theorem (Matrix Chernoff, for example)

Let $\widehat{v}_1, \dots, \widehat{v}_n$ *be independent random vectors with* $\|\widehat{v}_i\| \leq 1$ *and* $\sum_i \widehat{v}_i \widehat{v}_i^T = \widehat{V}$. *Then*

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \leq \theta\right] \geq 1 - d \cdot e^{-nD(\theta \| \lambda_{max}(\mathbb{E}\widehat{V}))}$$

Similar inequalities by Rudelson (1999), Ahlswede–Winter (2002).

All such inequalities have two things in common:

1. They give results with *high probability*
2. The bounds depend on the dimension

This will *always* be true — tight concentration (in this respect) depends on the dimension (consider n/d copies of basis vectors).

The goal

I want to find a bound on the eigenvalues that is *independent of dimension*.

The goal

I want to find a bound on the eigenvalues that is *independent of dimension*.

Furthemore, I want to keep the "probabilistic" nature:

Theorem

If $\widehat{\theta}$ is a random variable with finite support, then

$$\mathbb{P}\left[\widehat{\theta} \geq \mathbb{E}\widehat{\theta}\right] > 0 \quad \textit{and} \quad \mathbb{P}\left[\widehat{\theta} \leq \mathbb{E}\widehat{\theta}\right] > 0$$

In other words, I want to study one object (here $\mathbb{E}\widehat{\theta}$) and then be able to assert the existence of something at least as good (in both directions).

In fairy-tale land

So given a random frame $\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$, I would like to say:

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \geq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

and

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \leq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

In fairy-tale land

So given a random frame $\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$, I would like to say:

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \geq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

and

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \leq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

But this isn't true (pick just $\widehat{v}$ as $(0,1)$ or $(1,0)$ uniformly).

In fairy-tale land

So given a random frame $\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$, I would like to say:

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \geq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

and

$$\mathbb{P}\left[\lambda_{max}(\widehat{V}) \leq \lambda_{max}(\mathbb{E}\widehat{V})\right] > 0$$

But this isn't true (pick just $\widehat{v}$ as $(0, 1)$ or $(1, 0)$ uniformly).

So instead, we make an observation:

Observation

The eigenvalues of matrix are the roots of its characteristic polynomial. That is, if A is a $d \times d$ real, symmetric matrix with eigenvalues $\lambda_1, \dots, \lambda_d$, then

$$\chi_A(x) := \det[xI - A] = \prod_{i=1}^{d}(x - \lambda_i).$$

REAL fairy-tale land

So *now*, maybe we can do what we want in terms of polynomials!

REAL fairy-tale land

So *now*, maybe we can do what we want in terms of polynomials!

That is, given a random frame $\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$, maybe we can say:

$$\mathbb{P}\left[\text{maxroot}\left(\chi_{\widehat{V}}\right) \geq \text{maxroot}\left(\mathbb{E}\left[\chi_{\widehat{V}}\right]\right)\right] > 0$$

and

$$\mathbb{P}\left[\text{maxroot}\left(\chi_{\widehat{V}}\right) \leq \text{maxroot}\left(\mathbb{E}\left[\chi_{\widehat{V}}\right]\right)\right] > 0$$

REAL fairy-tale land

So *now*, maybe we can do what we want in terms of polynomials!

That is, given a random frame $\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$, maybe we can say:

$$\mathbb{P}\left[\text{maxroot}\left(\chi_{\widehat{V}}\right) \geq \text{maxroot}\left(\mathbb{E}\left[\chi_{\widehat{V}}\right]\right)\right] > 0$$

and

$$\mathbb{P}\left[\text{maxroot}\left(\chi_{\widehat{V}}\right) \leq \text{maxroot}\left(\mathbb{E}\left[\chi_{\widehat{V}}\right]\right)\right] > 0$$

Certainly this is nonsense, but let's play along with a toy problem:

Let A be a matrix and $\widehat{w}$ a random vector (taking values u or v uniformly).

What can we say about the eigenvalues of $A + \widehat{w}\widehat{w}^T$?

Still playing along

We would (naively) start by looking at the expected polynomial

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$

Why is this naive?

Still playing along

We would (naively) start by looking at the expected polynomial

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$

Why is this naive?

Adding polynomials is a function of the *coefficients* and we are interested in the *roots*.
In general, it is easy to get the coefficients from the roots but hard to get the roots from the coefficients.

Still playing along

We would (naively) start by looking at the expected polynomial

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$

Why is this naive?

Adding polynomials is a function of the *coefficients* and we are interested in the *roots*.
In general, it is easy to get the coefficients from the roots but hard to get the roots from the coefficients.

Example: $p(x) = (x-2)^2 - 1$ (has double root at 1) and $q(x) = (x+2)^2 - 1$ (has double root at -1).

$$p(x) + q(x) = x^2 + 6$$

does not have any real roots (roots are $\pm\sqrt{-6}$).

Unless...

Lemma (Separation Lemma)

Let $p_1, \ldots, p_k$ *be polynomials and* $[s, t]$ *an interval such that*

- *Each* $p_i(s)$ *has the same sign (or is* 0*)*
- *Each* $p_i(t)$ *has the same sign (or is* 0*)*
- *each* p_i *has exactly one real root in* $[s, t]$*.*

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

Unless...

Lemma (Separation Lemma)

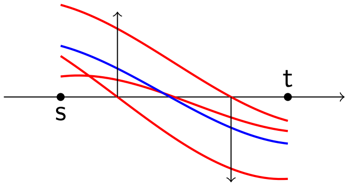
Let $p_1, \ldots, p_k$ be polynomials and $[s, t]$ an interval such that

- *Each $p_i(s)$ has the same sign (or is 0)*
- *Each $p_i(t)$ has the same sign (or is 0)*
- *each p_i has exactly one real root in $[s, t]$.*

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

Proof.

By picture:



□

A ray of hope

So if we have the right structure, using characteristic polynomials could actually work!

A ray of hope

So if we have the right structure, using characteristic polynomials could actually work!

Pros:

- All eigenvalues are tracked in a compact form
- Maybe take advantage of polynomial techniques that “don’t make sense” to matrices

A ray of hope

So if we have the right structure, using characteristic polynomials could actually work!

Pros:

- All eigenvalues are tracked in a compact form
- Maybe take advantage of polynomial techniques that "don't make sense" to matrices

Cons:

- You lose rotation (how can we add without knowing rotation?)
- Have to worry about matrix operations that "don't make sense" to polynomials

A ray of hope

So if we have the right structure, using characteristic polynomials could actually work!

Pros:

- All eigenvalues are tracked in a compact form
- Maybe take advantage of polynomial techniques that "don't make sense" to matrices

Cons:

- You lose rotation (how can we add without knowing rotation?)
- Have to worry about matrix operations that "don't make sense" to polynomials

What do I mean by "polynomial techniques"?

Polynomial Techniques

Univariate polynomials inherit techniques from

- Convex Analysis
- Complex Analysis
- Combinatorics

Polynomial Techniques

Univariate polynomials inherit techniques from

- Convex Analysis
- Complex Analysis
- Combinatorics

Multivariate polynomials inherit techniques from

- (Real) Algebraic Geometry
- Matroid theory
- Control Theory

Polynomial Techniques

Univariate polynomials inherit techniques from

- Convex Analysis
- Complex Analysis
- Combinatorics

Multivariate polynomials inherit techniques from

- (Real) Algebraic Geometry
- Matroid theory
- Control Theory

Both inherit from recent work in polynomial geometry:

- Hyperbolic polynomials
- Stable polynomials

What you need to remember

We are interested in the eigenvalues of (random) matrix sums:

$$\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$$

What you need to remember

We are interested in the eigenvalues of (random) matrix sums:

$$\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$$

All known techniques for this require concentration of measure and (as a result) weaken as the dimension grows.

What you need to remember

We are interested in the eigenvalues of (random) matrix sums:

$$\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$$

All known techniques for this require concentration of measure and (as a result) weaken as the dimension grows.

We will look for new techniques by doing something seemingly absurd: study their (random) characteristic polynomials.

What you need to remember

We are interested in the eigenvalues of (random) matrix sums:

$$\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$$

All known techniques for this require concentration of measure and (as a result) weaken as the dimension grows.

We will look for new techniques by doing something seemingly absurd: study their (random) characteristic polynomials.

In the case that we have root separation, we actually have a chance for this to work.

What you need to remember

We are interested in the eigenvalues of (random) matrix sums:

$$\widehat{V} = \sum_i \widehat{v}_i \widehat{v}_i^T$$

All known techniques for this require concentration of measure and (as a result) weaken as the dimension grows.

We will look for new techniques by doing something seemingly absurd: study their (random) characteristic polynomials.

In the case that we have root separation, we actually have a chance for this to work.

In exchange for requiring extra structure, we are hoping to get some new "polynomial techniques" that we can use.

Outline

Return on investment

To find separating intervals, we can use results in polynomial theory.

Return on investment

To find separating intervals, we can use results in polynomial theory.

Let p be a real rooted polynomial of degree d and q a real rooted polynomial of degree $d-1$

$$p(x) = \prod_{i=1}^{d}(x-\alpha_i) \quad \text{and} \quad q(x) = \prod_{i=1}^{d-1}(x-\beta_i)$$

with $\alpha_1 \leq \cdots \leq \alpha_d$ and $\beta_1 \leq \cdots \leq \beta_{d-1}$.

Return on investment

To find separating intervals, we can use results in polynomial theory.

Let p be a real rooted polynomial of degree d and q a real rooted polynomial of degree $d-1$

$$p(x) = \prod_{i=1}^{d}(x - \alpha_i) \quad \text{and} \quad q(x) = \prod_{i=1}^{d-1}(x - \beta_i)$$

with $\alpha_1 \leq \cdots \leq \alpha_d$ and $\beta_1 \leq \cdots \leq \beta_{d-1}$.

We say q *interlaces* p if $\alpha_1 \leq \beta_1 \leq \alpha_2 \cdots \leq \alpha_{d-1} \leq \beta_{d-1} \leq \alpha_d$.

Think: The roots of q separate the roots of p.

Return on investment

To find separating intervals, we can use results in polynomial theory.

Let p be a real rooted polynomial of degree d and q a real rooted polynomial of degree $d-1$

$$p(x) = \prod_{i=1}^{d}(x - \alpha_i) \quad \text{and} \quad q(x) = \prod_{i=1}^{d-1}(x - \beta_i)$$

with $\alpha_1 \leq \cdots \leq \alpha_d$ and $\beta_1 \leq \cdots \leq \beta_{d-1}$.

We say q *interlaces* p if $\alpha_1 \leq \beta_1 \leq \alpha_2 \cdots \leq \alpha_{d-1} \leq \beta_{d-1} \leq \alpha_d$.

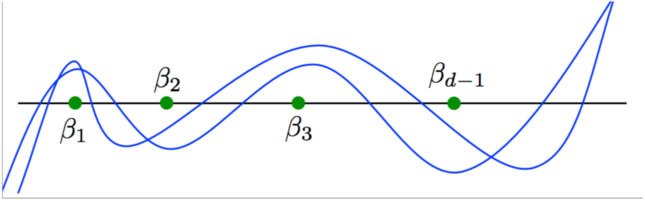
Think: The roots of q separate the roots of p.

Example: $p'(x)$ interlaces $p(x)$.

Common Interlacer

We say that degree d real rooted polynomials $p_1, \ldots, p_k$ have a *common interlacer* if there exists a q such that q interlaces *every* p_i simultaneously.

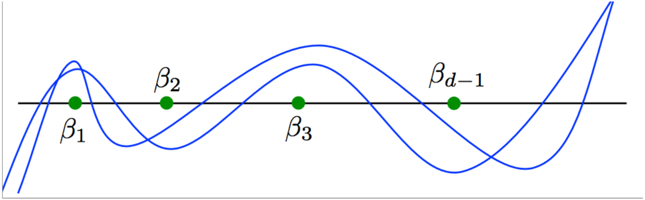
Think: the roots of q split up $\mathbb{R}$ into d intervals, each of which contains exactly one root of each p_i.



Common Interlacer

We say that degree d real rooted polynomials $p_1, \ldots, p_k$ have a *common interlacer* if there exists a q such that q interlaces *every* p_i simultaneously.

Think: the roots of q split up $\mathbb{R}$ into d intervals, each of which contains exactly one root of each p_i.



Note: if the p_i have a common interlacer (say q), then the intervals defined by the β_i can serve as separators for the lemma!

Back to the toy problem

Recall our goal was to understand the roots of

$$\begin{aligned} p(x) &= \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x) \\ &= \frac{1}{2}q_0(x) + \frac{1}{2}q_1(x) \end{aligned}$$

Back to the toy problem

Recall our goal was to understand the roots of

$$\begin{aligned} p(x) &= \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x) \\ &= \frac{1}{2}q_0(x) + \frac{1}{2}q_1(x) \end{aligned}$$

We will say that p forms an *interlacing star* with $\{q_i\}$ if

1. p and $\{q_i\}$ have the same degree and are all real rooted
2. The leading coefficients of the $\{q_i\}$ have the same sign
3. The collection of polynomials $\{q_i\}$ has a common interlacer
4. p is a convex combination of the $\{q_i\}$

Back to the toy problem

Recall our goal was to understand the roots of

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$
$$= \frac{1}{2}q_0(x) + \frac{1}{2}q_1(x)$$

We will say that p forms an *interlacing star* with $\{q_i\}$ if

1. p and $\{q_i\}$ have the same degree and are all real rooted
2. The leading coefficients of the $\{q_i\}$ have the same sign
3. The collection of polynomials $\{q_i\}$ has a common interlacer
4. p is a convex combination of the $\{q_i\}$

Corollary

If p forms an interlacing star with $\{q_i\}$, then there exist i, j such that

$$\mathrm{k}^{\mathrm{th}}\mathrm{root}\,(q_i) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\,(p) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\,(q_j)$$

More help from polynomials

Polynomial theory gives us a nice characterization of interlacing:

Lemma (Chudnovsky–Seymour, among others)

Let $\{p_i\}$ *be a collection of degree* d *polynomials. The following are equivalent:*

- *Every polynomial in the convex hull of* $\{p_i\}$ *has* d *real roots.*
- *The collection* $\{p_i\}$ *has a common interlacer.*

More help from polynomials

Polynomial theory gives us a nice characterization of interlacing:

Lemma (Chudnovsky–Seymour, among others)

Let $\{p_i\}$ be a collection of degree d polynomials. The following are equivalent:

- *Every polynomial in the convex hull of $\{p_i\}$ has d real roots.*
- *The collection $\{p_i\}$ has a common interlacer.*

Recall (again) our equation

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$

More help from polynomials

Polynomial theory gives us a nice characterization of interlacing:

Lemma (Chudnovsky–Seymour, among others)

Let $\{p_i\}$ be a collection of degree d polynomials. The following are equivalent:

- *Every polynomial in the convex hull of $\{p_i\}$ has d real roots.*
- *The collection $\{p_i\}$ has a common interlacer.*

Recall (again) our equation

$$p(x) = \frac{1}{2}\chi_{A+uu^T}(x) + \frac{1}{2}\chi_{A+vv^T}(x)$$

If we could show that

$$p(x) = \lambda\chi_{A+vv^T}(x) + (1-\lambda)\chi_{A+uu^T}(x)$$

was real rooted for all $\lambda \in [0, 1]$, then we would get the interlacing for free.

Back to reality

But remember we are interested in graphs — that is, sums of possibly *multiple* random vectors.

p_{00} p_{01} p_{10} p_{11}

Back to reality

But remember we are interested in graphs — that is, sums of possibly *multiple* random vectors.

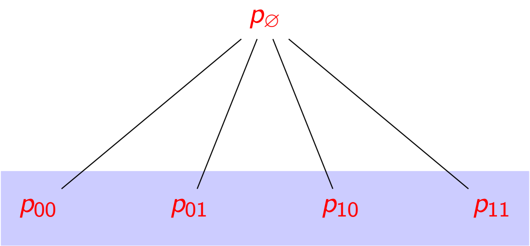
If all of the resulting characteristic polynomials had a common interlacer,

p_{00}	p_{01}	p_{10}	p_{11}

Back to reality

But remember we are interested in graphs — that is, sums of possibly *multiple* random vectors.

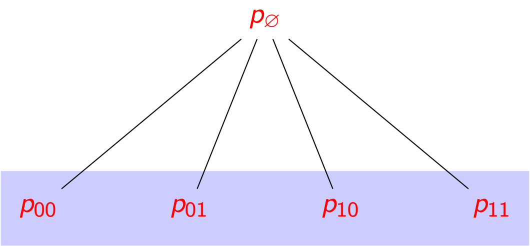
If all of the resulting characteristic polynomials had a common interlacer, we could study some convex combination and be able to use the lemma.



Back to reality

But remember we are interested in graphs — that is, sums of possibly *multiple* random vectors.

If all of the resulting characteristic polynomials had a common interlacer, we could study some convex combination and be able to use the lemma.



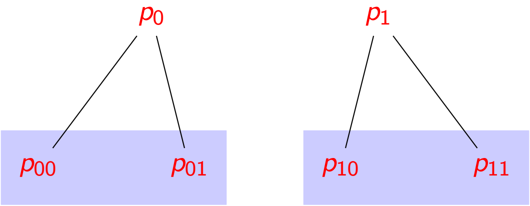
But in general they don't have a common interlacer...

Instead...

p_{00} p_{01} p_{10} p_{11}

Instead...

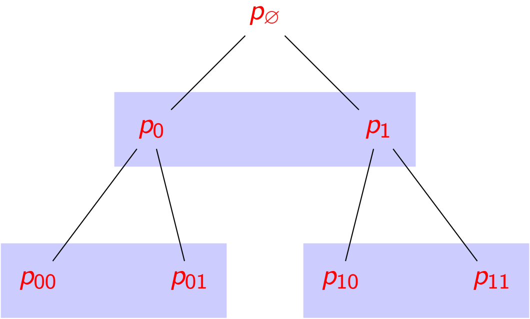
We can try to group them into smaller stars.



Instead...

We can try to group them into smaller stars.

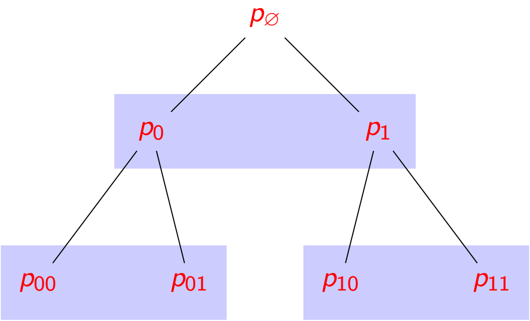
And then try to iterate.



Instead...

We can try to group them into smaller stars.

And then try to iterate.



We will call a rooted, connected tree where each node forms an interlacing star with its children an *interlacing family*.

The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} *and* p_{leaf_2} *such that*

$$\mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{leaf_1}\right) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{root}\right) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{leaf_2}\right).$$

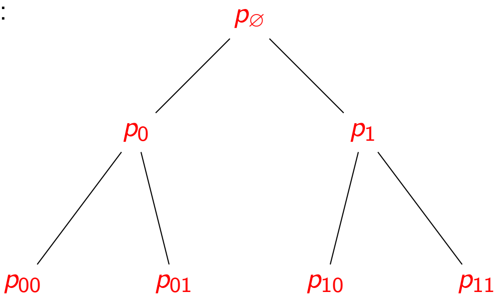
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} and p_{leaf_2} such that

$$\mathrm{k^{th}root}\left(p_{leaf_1}\right) \leq \mathrm{k^{th}root}\left(p_{root}\right) \leq \mathrm{k^{th}root}\left(p_{leaf_2}\right).$$

To find p_{leaf_i}:



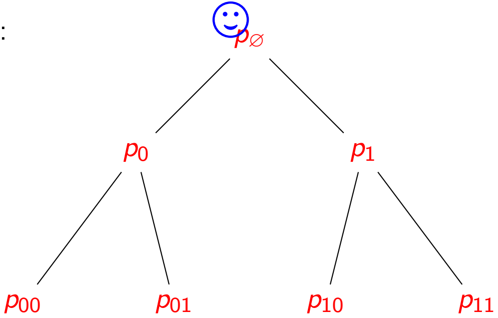
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} and p_{leaf_2} such that

$$\mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{leaf_1}\right) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{root}\right) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\left(p_{leaf_2}\right).$$

To find p_{leaf_i}:



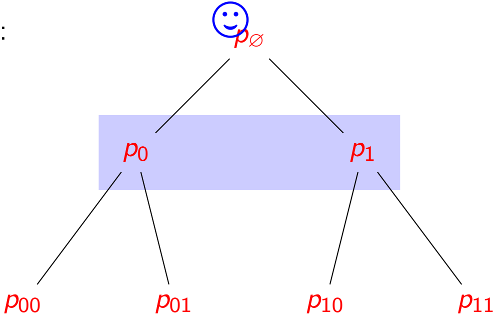
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} and p_{leaf_2} such that

$$\mathrm{k^{th}root}\,(p_{leaf_1}) \le \mathrm{k^{th}root}\,(p_{root}) \le \mathrm{k^{th}root}\,(p_{leaf_2})\,.$$

To find p_{leaf_i}:



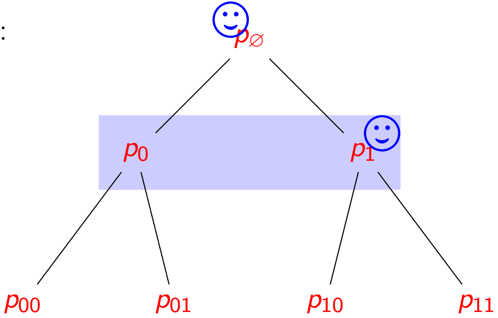
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} and p_{leaf_2} such that

$$\text{k}^{\text{th}}\text{root}\left(p_{leaf_1}\right) \leq \text{k}^{\text{th}}\text{root}\left(p_{root}\right) \leq \text{k}^{\text{th}}\text{root}\left(p_{leaf_2}\right).$$

To find p_{leaf_i}:



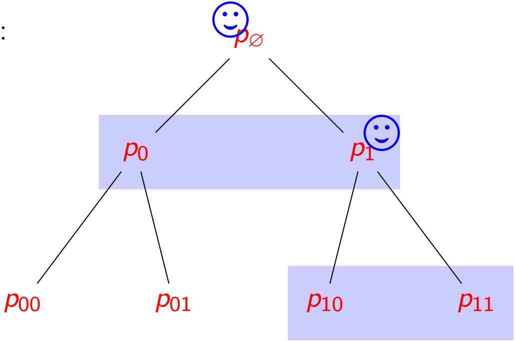
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} *and* p_{leaf_2} *such that*

$$\mathrm{k^{th}root}\,(p_{leaf_1}) \leq \mathrm{k^{th}root}\,(p_{root}) \leq \mathrm{k^{th}root}\,(p_{leaf_2})\,.$$

To find p_{leaf_i}:



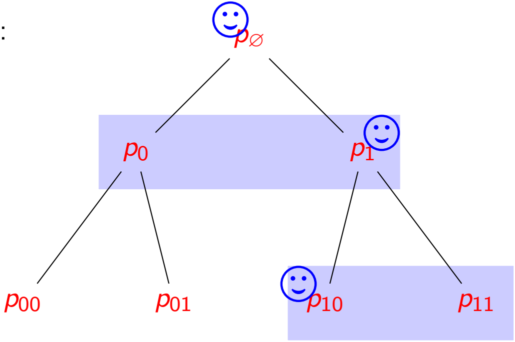
The punchline

Corollary

Every interlacing family contains leaf nodes p_{leaf_1} *and* p_{leaf_2} *such that*

$$\mathrm{k}^{\mathrm{th}}\mathrm{root}\,(p_{leaf_1}) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\,(p_{root}) \leq \mathrm{k}^{\mathrm{th}}\mathrm{root}\,(p_{leaf_2})\,.$$

To find p_{leaf_i}:



Outline

Building an interlacing family

Consider the sum of random vectors

$$\widehat{V} = \sum_{i=1}^{m} \widehat{v_i}\widehat{v_i}^T$$

where the $\widehat{v_i}$ have support size at most n.

Building an interlacing family

Consider the sum of random vectors

$$\widehat{V} = \sum_{i=1}^{m} \widehat{v}_i \widehat{v}_i^T$$

where the $\widehat{v}_i$ have support size at most n.

We will define a *choice vector* $\sigma \in [n]^m$ where σ_i is the index of a vector in the support of $\widehat{v}_i$. Then the characteristic polynomial of a fixed frame V in the support $\widehat{V}$ can be denoted

$$p_\sigma(x) = \chi_V(x).$$

Building an interlacing family

Consider the sum of random vectors

$$\widehat{V} = \sum_{i=1}^{m} \widehat{v}_i \widehat{v}_i^T$$

where the $\widehat{v}_i$ have support size at most n.

We will define a *choice vector* $\sigma \in [n]^m$ where σ_i is the index of a vector in the support of $\widehat{v}_i$. Then the characteristic polynomial of a fixed frame V in the support $\widehat{V}$ can be denoted

$$p_\sigma(x) = \chi_V(x).$$

We then define *partial choice* vectors $\sigma' \in [n]^k$ for $k < m$; the corresponding polynomial will be the conditional expectation:

$$p_{\sigma'} = \mathbb{E}_{\widehat{v}_{k+1},\ldots,\widehat{v}_d}\left[\chi(\widehat{V})(x) \mid \widehat{v}_i = v_i^{\sigma'_i} \text{ for } 1 \leq i \leq k\right]$$

Building an interlacing family

Consider the sum of random vectors

$$\widehat{V} = \sum_{i=1}^{m} \widehat{v}_i \widehat{v}_i^T$$

where the $\widehat{v}_i$ have support size at most n.

We will define a *choice vector* $\sigma \in [n]^m$ where σ_i is the index of a vector in the support of $\widehat{v}_i$. Then the characteristic polynomial of a fixed frame V in the support $\widehat{V}$ can be denoted

$$p_\sigma(x) = \chi_V(x).$$

We then define *partial choice* vectors $\sigma' \in [n]^k$ for $k < m$; the corresponding polynomial will be the conditional expectation:

$$p_{\sigma'} = \mathbb{E}_{\widehat{v}_{k+1},\dots,\widehat{v}_d} \left[\chi(\widehat{V})(x) \mid \widehat{v}_i = v_i^{\sigma'_i} \text{ for } 1 \leq i \leq k \right]$$

This forms an n-ary tree with fixed assignments at the leaves and $p_\varnothing = \mathbb{E}\left[\chi_{\widehat{V}}(x)\right]$ at the root.

Mixed characteristic polynomials

If we choose the vectors independently, the polynomials take a special form:

Mixed characteristic polynomials

If we choose the vectors independently, the polynomials take a special form:

Theorem

Let $\widehat{v}_1, \dots \widehat{v}_m$ be independent random vectors such that $\mathbb{E}\left[\widehat{v}_i \widehat{v}_i^T\right] = A_i$. Then

$$\mathbb{E}\left[\chi_{\widehat{V}}(x)\right] = \prod_{i=1}^{m}\left(1 - \frac{\partial}{\partial z_i}\right) \det\left[xI + \sum_{i=1}^{m} z_i A_i\right]\Bigg|_{z_1 = \dots = z_m = 0}$$

In particular, the expected polynomial does not depend on the vectors or the probabilities — only the expected outer product.

Mixed characteristic polynomials

If we choose the vectors independently, the polynomials take a special form:

Theorem

Let $\widehat{v}_1, \ldots \widehat{v}_m$ *be independent random vectors such that* $\mathbb{E}\left[\widehat{v}_i \widehat{v}_i^T\right] = A_i$*. Then*

$$\mathbb{E}\left[\chi_{\widehat{V}}(x)\right] = \prod_{i=1}^{m}\left(1 - \frac{\partial}{\partial z_i}\right) \det\left[xI + \sum_{i=1}^{m} z_i A_i\right]\Bigg|_{z_1=\cdots=z_m=0}$$

In particular, the expected polynomial does not depend on the vectors or the probabilities — only the expected outer product.

We call this a *mixed characteristic polynomial* and denote it $\mu[A_1, \ldots, A_m]$.

A world of mixed characteristic polynomials

Every polynomial we defined previously is a mixed characteristic polynomial.

A world of mixed characteristic polynomials

Every polynomial we defined previously is a mixed characteristic polynomial.

1. Normal characteristic polynomials (for an assignment $\sigma = v_1, \ldots, v_m$ with $\sum_i v_i v_i^T = V$)

$$p_\sigma(x) = \chi_V(x) = \mu[v_1 v_1^T, \ldots, v_m v_m^T](x)$$

A world of mixed characteristic polynomials

Every polynomial we defined previously is a mixed characteristic polynomial.

1. Normal characteristic polynomials (for an assignment $\sigma = v_1, \ldots, v_m$ with $\sum_i v_i v_i^T = V$)

$$p_\sigma(x) = \chi_V(x) = \mu[v_1 v_1^T, \ldots, v_m v_m^T](x)$$

2. The expected characteristic polynomial (with $\mathbb{E}\left[\widehat{v}_i \widehat{v}_i^T\right] = A_i$)

$$\mathbb{E}\left[\chi_{\widehat{V}}(x)\right] = \mu[A_1, \ldots, A_m](x)$$

A world of mixed characteristic polynomials

Every polynomial we defined previously is a mixed characteristic polynomial.

1. Normal characteristic polynomials (for an assignment $\sigma = v_1, \ldots, v_m$ with $\sum_i v_i v_i^T = V$)

$$p_\sigma(x) = \chi_V(x) = \mu[v_1 v_1^T, \ldots, v_m v_m^T](x)$$

2. The expected characteristic polynomial (with $\mathbb{E}\left[\widehat{v}_i \widehat{v}_i^T\right] = A_i$)

$$\mathbb{E}\left[\chi_{\widehat{V}}(x)\right] = \mu[A_1, \ldots, A_m](x)$$

3. The partial assignment polynomials

$$\begin{aligned} p_{\sigma'} &= \mathbb{E}_{\widehat{v}_{k+1}, \ldots, \widehat{v}_d}\left[\chi_{\widehat{V}}(x) \mid \widehat{v}_i = v_i^{\sigma'_i} \text{ for } 1 \le i \le k\right] \\ &= \mu[v_1 v_1^T, \ldots, v_k v_k^T, A_{k+1}, \ldots, A_m] \end{aligned}$$

Real stable polynomials

The advantage of having a multivariate formula is that we can utilize the theory of *real stable polynomials*, a multivariate extension of real rooted polynomials. Let
$\mathbb{H} = \{x \in \mathbb{C} \mid \Im(z_i) > 0\}$.

Real stable polynomials

The advantage of having a multivariate formula is that we can utilize the theory of *real stable polynomials*, a multivariate extension of real rooted polynomials. Let $\mathbb{H} = \{x \in \mathbb{C} \mid \Im(z_i) > 0\}$.

An n-variate polynomial p is called *stable* if it is never 0 in $\mathbb{H}^n$. (i.e. if $p(z_1, \dots, z_n) = 0$, then some z_i has nonnegative imaginary part). If, in addition, all coefficients of p are real, it is called *real stable*.

Real stable polynomials

The advantage of having a multivariate formula is that we can utilize the theory of *real stable polynomials*, a multivariate extension of real rooted polynomials. Let $\mathbb{H} = \{x \in \mathbb{C} \mid \Im(z_i) > 0\}$.

An n-variate polynomial p is called *stable* if it is never 0 in $\mathbb{H}^n$. (i.e. if $p(z_1, \dots, z_n) = 0$, then some z_i has nonnegative imaginary part). If, in addition, all coefficients of p are real, it is called *real stable*.

Two important properties:

- Univariate polynomials are real rooted if and only if they are real stable.
- Real stable polynomials are closed under substitution of reals $(z_1, z_2, \dots, z_n) \to (a, z_2, \dots, z_n)$ for $a \in \mathbb{R}$.

Similar to *hyperbolic polynomials*.

Real stable techniques

Real stability has been well studied in recent years. In particular,

Lemma

Let $A_1, \dots, A_m$ be Hermitian positive semidefinite matrices and $x_1 \dots x_m$ variables. Then

$$p(x_1, \dots, x_m) = \det\left[\sum_{i=1}^{m} x_i A_i\right]$$

is a real stable polynomial.

Real stable techniques

Real stability has been well studied in recent years. In particular,

Lemma

Let $A_1, \ldots, A_m$ be Hermitian positive semidefinite matrices and $x_1 \ldots x_m$ variables. Then

$$p(x_1, \ldots, x_m) = \det\left[\sum_{i=1}^{m} x_i A_i\right]$$

is a real stable polynomial.

Lemma

If $p(x_1, \ldots, x_m)$ is a multiaffine polynomial with real coefficients, then the following are equivalent:

1. *p is real stable*
2. $$\Delta_{ij}[p](x_1, \ldots, x_m) := \frac{\partial p}{\partial x_i}\frac{\partial p}{\partial x_j} - \frac{\partial^2 p}{\partial x_i \partial x_j} p \geq 0$$

Cutting to the chase

Theorem

Mixed characteristic polynomials are real rooted.

Cutting to the chase

Theorem

Mixed characteristic polynomials are real rooted.

Proof.

Follows directly from the formula:

$$\mu[A_1, \ldots, A_m](x) = \prod_{i=1}^{m} \left(1 - \frac{\partial}{\partial z_i}\right) \det\left[xI + \sum_{i=1}^{m} z_i A_i\right]\Bigg|_{z_1=\cdots=z_m=0}$$

□

Cutting to the chase

Theorem

Mixed characteristic polynomials are real rooted.

Proof.

Follows directly from the formula:

$$\mu[A_1, \dots, A_m](x) = \prod_{i=1}^{m} \left(1 - \frac{\partial}{\partial z_i}\right) \det\left[xI + \sum_{i=1}^{m} z_i A_i\right]\Bigg|_{z_1=\cdots=z_m=0}$$

□

This provides an easy way to generate interlacing families.

Corollary

Any tree of polynomials resulting from choosing independent random vectors forms an interlacing family.

Full Circle

So what have we accomplished?

Full Circle

So what have we accomplished?

We now have a “probabilistic” way to deal with roots of polynomials (under certain conditions).

Full Circle

So what have we accomplished?

We now have a "probabilistic" way to deal with roots of polynomials (under certain conditions).

In the case that we are choosing vectors independently and wanting to track the eigenvalues, those conditions are satisfied.

Full Circle

So what have we accomplished?

We now have a "probabilistic" way to deal with roots of polynomials (under certain conditions).

In the case that we are choosing vectors independently and wanting to track the eigenvalues, those conditions are satisfied.

Hence we have a "probabilistic" way to deal with eigenvalues. That is, for any given k, let R be the k^{th} root of the *expected characteristic polynomial* (under whatever product distribution you want). Then there exists

1. an assignment of the random vectors that has $\lambda_k \geq R$
2. an assignment of the random vectors that has $\lambda_k \leq R$

Outline

Expanders

Let's switch gears to talk about expander graphs.

Expanders

Let's switch gears to talk about expander graphs.

Expander graphs like to see new vertices — in this regard the "ultimate" d-regular expander is the d-regular infinite tree.

Expanders

Let's switch gears to talk about expander graphs.

Expander graphs like to see new vertices — in this regard the "ultimate" d-regular expander is the d-regular infinite tree.

So a "good" d-regular expander should be a "good" approximation of the d-regular tree.

In spectral graph theory, the definition of "good" is in terms of eigenvalues, and the spectrum of the d-infinite tree lies inside the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.

Expanders

Let's switch gears to talk about expander graphs.

Expander graphs like to see new vertices — in this regard the "ultimate" d-regular expander is the d-regular infinite tree.

So a "good" d-regular expander should be a "good" approximation of the d-regular tree.

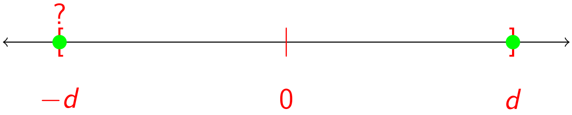
In spectral graph theory, the definition of "good" is in terms of eigenvalues, and the spectrum of the d-infinite tree lies inside the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.

Can we get a finite graph to have all eigenvalues inside this interval?

Ramaujan graphs

No — the adjacency matrix of a d-regular graph has either 1 or 2 (so-called) *trivial* eigenvalues

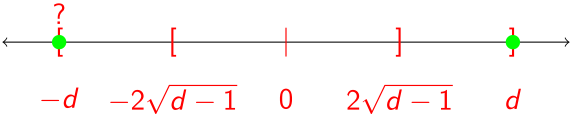
1. d is always the largest eigenvalue
2. G is bipartite if and only if $-d$ is an eigenvalue



Ramanujan graphs

No — the adjacency matrix of a d-regular graph has either 1 or 2 (so-called) *trivial* eigenvalues

1. d is always the largest eigenvalue
2. G is bipartite if and only if $-d$ is an eigenvalue

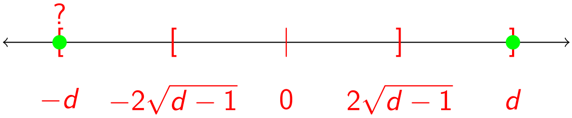


A d-regular graph with all nontrivial eigenvalues inside $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ is called a *Ramanujan graph* and an infinite collection (all d-regular) a *Ramanujan family*.

Ramanujan graphs

No — the adjacency matrix of a d-regular graph has either 1 or 2 (so-called) *trivial* eigenvalues

1. d is always the largest eigenvalue
2. G is bipartite if and only if $-d$ is an eigenvalue



A d-regular graph with all nontrivial eigenvalues inside $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ is called a *Ramanujan graph* and an infinite collection (all d-regular) a *Ramanujan family*.

Theorem (Alon, Boppana (1996))

No smaller interval can contain all nontrivial eigenvalues of an infinite collection of d-regular graphs.

Previous results

Theorem (Margulis, Lubotzky–Phillips–Sarnak (1988))

Ramanujan families exist for $d = p + 1$ where p is a prime number.

Previous results

Theorem (Margulis, Lubotzky–Phillips–Sarnak (1988))

Ramanujan families exist for $d = p + 1$ *where* p *is a prime number.*

Extended by Morganstern to $d = p^k + 1$, unknown for all other d.

All known constructions* are algebraic — they are Cayley graphs of highly structured groups.

Previous results

Theorem (Margulis, Lubotzky–Phillips–Sarnak (1988))

Ramanujan families exist for $d = p + 1$ where p is a prime number.

Extended by Morganstern to $d = p^k + 1$, unknown for all other d.

All known constructions* are algebraic — they are Cayley graphs of highly structured groups.

On the other hand, almost everything is almost Ramanujan:

Theorem (Friedman (2008))

For fixed d, there is a large enough n such that a randomly chosen d-regular graph on n vertices have a nontrivial spectrum inside the interval

$$[-2\sqrt{d-1} - \epsilon, 2\sqrt{d-1} + \epsilon]$$

with high probability.

Previous results

Theorem (Margulis, Lubotzky–Phillips–Sarnak (1988))

Ramanujan families exist for $d = p + 1$ where p is a prime number.

Extended by Morganstern to $d = p^k + 1$, unknown for all other d.

All known constructions* are algebraic — they are Cayley graphs of highly structured groups.

On the other hand, almost everything is almost Ramanujan:

Theorem (Friedman (2008))

For fixed d, there is a large enough n such that a randomly chosen d-regular graph on n vertices have a nontrivial spectrum inside the interval

$$[-2\sqrt{d-1} - \epsilon, 2\sqrt{d-1} + \epsilon]$$

with high probability.

Obvious question: are Ramanujan families really that special?

Idea 1: Covering space

It is easy to see that small graphs are Ramanujan.

1. K_d has all nontrivial eigenvalues -1
2. $K_{d,d}$ has all nontrivial eigenvalues 0

Idea 1: Covering space

It is easy to see that small graphs are Ramanujan.

1. K_d has all nontrivial eigenvalues -1
2. $K_{d,d}$ has all nontrivial eigenvalues 0

Bilu and Linial (2006) suggested treating the d-regular graphs as *quotients* of the d-infinite tree w.r.t. a covering map.

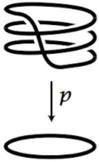
Idea 1: Covering space

It is easy to see that small graphs are Ramanujan.

1. K_d has all nontrivial eigenvalues -1
2. $K_{d,d}$ has all nontrivial eigenvalues 0

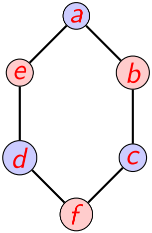
Bilu and Linial (2006) suggested treating the d-regular graphs as *quotients* of the d-infinite tree w.r.t. a covering map.

Their idea — construct a sequence of graphs from K_d (or $K_{d,d}$) to the d-infinite tree using a series of *lifts* (a common construction in algebraic topology).



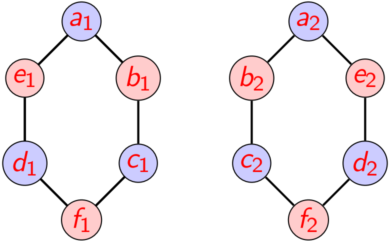
Lifting

Start with a (Ramanujan) graph.



Lifting

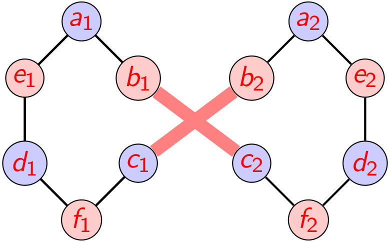
Start with a (Ramanujan) graph.



And make a copy (copies?) of it.

Lifting

Start with a (Ramanujan) graph.



And make a copy (copies?) of it. And perturb it.

Want to find perturbations that cause new graph to be good.

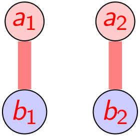
2-lifts

Bilu and Linial (2006) studied perturbations they called *2-lifts*.

2-lifts

Bilu and Linial (2006) studied perturbations they called *2-lifts*.

For each possible *signing* $s \in \{\pm\}^{|E|}$, form a graph by doing either

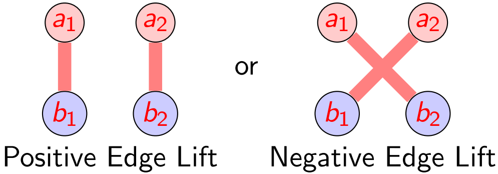


Positive Edge Lift

2-lifts

Bilu and Linial (2006) studied perturbations they called *2-lifts*.

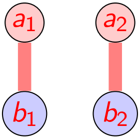
For each possible *signing* $s \in \{\pm\}^{|E|}$, form a graph by doing either



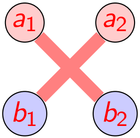
2-lifts

Bilu and Linial (2006) studied perturbations they called *2-lifts*.

For each possible *signing* $s \in \{\pm\}^{|E|}$, form a graph by doing either



or



Positive Edge Lift Negative Edge Lift

The action of each 2-lift can be described by its *signed adjacency matrix* A_s:

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{matrix} \quad \rightarrow \quad \begin{matrix} 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{matrix}$$

Main Eigenvalue lemma

Theorem (Bilu–Linial (2006))

Let G be a d-regular Ramanujan graph with n vertices and let s be a signing of G. If all eigenvalues of A_s lie in the interval

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

then the 2-lift G_s is a d-regular Ramanujan graph with $2n$ vertices.

Main Eigenvalue lemma

Theorem (Bilu–Linial (2006))

Let G be a d-regular Ramanujan graph with n vertices and let s be a signing of G. If all eigenvalues of A_s lie in the interval

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

then the 2-lift G_s is a d-regular Ramanujan graph with $2n$ vertices.

Conjecture (Bilu–Linial (2006))

Every *d-regular graph contains a signing s for which the eigenvalues of A_s lie inside the interval*

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

Main Eigenvalue lemma

Theorem (Bilu–Linial (2006))

Let G be a d-regular Ramanujan graph with n vertices and let s be a signing of G. If all eigenvalues of A_s lie in the interval

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

then the 2-lift G_s is a d-regular Ramanujan graph with $2n$ vertices.

Conjecture (Bilu–Linial (2006))

Every *d-regular graph contains a signing s for which the eigenvalues of A_s lie inside the interval*

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

We prove the conjecture for every *bipartite* graph G.

Bipartite Adjacency Matrices

What is so special about being bipartite?

Bipartite Adjacency Matrices

What is so special about being bipartite?

In this case, the signed adjacency matrix can be written in block form

$$\left(\begin{array}{c|c} 0 & B \\ \hline B^T & 0 \end{array}\right)$$

causing eigenvalues/vectors to come in pairs

$$v_i = [u_i \mid u_i] \quad \text{and} \quad v_{n-i} = [u_i \mid -u_i]$$

for $1 \leq i \leq n/2$ and so the eigenvalues satisfy $\lambda_i = -\lambda_{n-i}$.

Bipartite Adjacency Matrices

What is so special about being bipartite?

In this case, the signed adjacency matrix can be written in block form

$$\left(\begin{array}{c|c} 0 & B \\ \hline B^T & 0 \end{array}\right)$$

causing eigenvalues/vectors to come in pairs

$$v_i = [u_i \mid u_i] \quad \text{and} \quad v_{n-i} = [u_i \mid -u_i]$$

for $1 \leq i \leq n/2$ and so the eigenvalues satisfy $\lambda_i = -\lambda_{n-i}$.

Corollary

A bipartite signed adjacency matrix A_s has all of its eigenvalues in the interval

$$[-2\sqrt{d-1}, 2\sqrt{d-1}]$$

if and only if all of its eigenvalues are at most $2\sqrt{d-1}$.

Main idea

For each signing, we consider the characteristic polynomial of the signed adjacency matrix.

Main idea

For each signing, we consider the characteristic polynomial of the signed adjacency matrix.

These correspond to picking either

- $(\delta_i + \delta_j)$, or
- $(\delta_i - \delta_j)$

independently for each edge (v_i, v_j).

Main idea

For each signing, we consider the characteristic polynomial of the signed adjacency matrix.

These correspond to picking either

- $(\delta_i + \delta_j)$, or
- $(\delta_i - \delta_j)$

independently for each edge (v_i, v_j).

Corollary

The tree corresponding to these polynomials forms an interlacing family.

Main idea

For each signing, we consider the characteristic polynomial of the signed adjacency matrix.

These correspond to picking either

- $(\delta_i + \delta_j)$, or
- $(\delta_i - \delta_j)$

independently for each edge (v_i, v_j).

Corollary

The tree corresponding to these polynomials forms an interlacing family.

Hence it suffices to bound the largest root of the expected characteristic polynomial.

The expected characteristic polynomial

Theorem (Godsil–Gutman (1981))

For any graph G,

$$\mathbb{E}_{s \in \{\pm\}^m} \chi_{A_s}(x) = \mu_G(x),$$

the matching polynomial of G.

The expected characteristic polynomial

Theorem (Godsil–Gutman (1981))

For any graph G,

$$\mathbb{E}_{s \in \{\pm\}^m} \chi_{A_s}(x) = \mu_G(x),$$

the matching polynomial of G.

In the original paper, Heilmann and Lieb also proved the following bound:

Theorem (Heilmann–Lieb (1972))

Let G be a graph with maximum degree Δ. Then

$$\mathrm{maxroot}\left(\mu_G\right) \leq 2\sqrt{\Delta - 1}$$

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Proof.

Set $G_0 = K_{d,d}$ (which is bipartite, d-regular, and Ramanujan for any d). Given G_i, form G_{i+1} as follows:

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Proof.

Set $G_0 = K_{d,d}$ (which is bipartite, d-regular, and Ramanujan for any d). Given G_i, form G_{i+1} as follows:

For each possible signing $s \in \{\pm 1\}^{|E_i|}$, form the polynomials p_s. By our theorem, this is an interlacing family.

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Proof.

Set $G_0 = K_{d,d}$ (which is bipartite, d-regular, and Ramanujan for any d). Given G_i, form G_{i+1} as follows:

For each possible signing $s \in \{\pm 1\}^{|E_i|}$, form the polynomials p_s. By our theorem, this is an interlacing family.

Combining this with Godsil–Gutman and Heilmann–Lieb ensures some p_{s^*} such that $\mathrm{maxroot}(p_{s^*}) \leq \mathrm{maxroot}(p_\varnothing) \leq 2\sqrt{d-1}$

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Proof.

Set $G_0 = K_{d,d}$ (which is bipartite, d-regular, and Ramanujan for any d). Given G_i, form G_{i+1} as follows:

For each possible signing $s \in \{\pm 1\}^{|E_i|}$, form the polynomials p_s. By our theorem, this is an interlacing family.

Combining this with Godsil–Gutman and Heilmann–Lieb ensures some p_{s^*} such that $\mathrm{maxroot}(p_{s^*}) \leq \mathrm{maxroot}(p_\varnothing) \leq 2\sqrt{d-1}$

Set G_{i+1} to be the 2-lift given by s^* — this is bipartite, d-regular, and (by Bilu and Linial) Ramanujan — and continue.

Piecing things together

Theorem

There exist bipartite Ramanujan families of degree d for any d.

Proof.

Set $G_0 = K_{d,d}$ (which is bipartite, d-regular, and Ramanujan for any d). Given G_i, form G_{i+1} as follows:

For each possible signing $s \in \{\pm 1\}^{|E_i|}$, form the polynomials p_s. By our theorem, this is an interlacing family.

Combining this with Godsil–Gutman and Heilmann–Lieb ensures some p_{s^*} such that $\mathrm{maxroot}(p_{s^*}) \leq \mathrm{maxroot}(p_\varnothing) \leq 2\sqrt{d-1}$

Set G_{i+1} to be the 2-lift given by s^* — this is bipartite, d-regular, and (by Bilu and Linial) Ramanujan — and continue.

Extended to d-lifts by Hall, Puder, and Sawin (2015).

$\square$

Idea 2: Free group

The $2d$-infinite tree is the Cayley graph of the free group on d elements.

Idea 2: Free group

The $2d$-infinite tree is the Cayley graph of the free group on d elements.

You can "build" a free group on d elements by using the free product

$$F_d = \underbrace{F_1 * F_1 * \cdots * F_1}_{d \text{ times}}$$

Idea 2: Free group

The $2d$-infinite tree is the Cayley graph of the free group on d elements.

You can "build" a free group on d elements by using the free product

$$F_d = \underbrace{F_1 * F_1 * \cdots * F_1}_{d \text{ times}}$$

Intuitively, you would build a d-regular graph a similar way:

$$G_d = \underbrace{M_1 \cup M_2 \cup \cdots \cup M_d}_{d \text{ times}}$$

where each M_i is a perfect matching.

Idea 2: Free group

The $2d$-infinite tree is the Cayley graph of the free group on d elements.

You can "build" a free group on d elements by using the free product

$$F_d = \underbrace{F_1 * F_1 * \cdots * F_1}_{d \text{ times}}$$

Intuitively, you would build a d-regular graph a similar way:

$$G_d = \underbrace{M_1 \cup M_2 \cup \cdots \cup M_d}_{d \text{ times}}$$

where each M_i is a perfect matching.

Can we build a Ramanujan graph this way?

What's happening

Perfect matchings on $2n$ vertices have eigenvalues 1 and -1 (each n times).

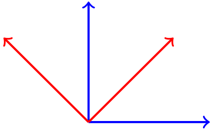
What's happening

Perfect matchings on $2n$ vertices have eigenvalues 1 and -1 (each n times).

A union of graphs is a sum of matrices... and the eigenvalues of $A + B$ depend on

1. the eigenvalues of A
2. the eigenvalues of B
3. the dot product of the corresponding eigenvectors

To keep eigenvalues low, you want the eigenvectors to be as "orthogonal as possible"



Free probability

As always, we can try to add A and B randomly.

As dimension the matrix gets larger, probability of having aligned eigenvectors goes down.

Free probability

As always, we can try to add A and B randomly.

As dimension the matrix gets larger, probability of having aligned eigenvectors goes down.

In the infinite limit, “orthogonal as possible” happens with probability one — the corresponding eigenvalue distributions are known as the *free convolution*.

Free probability

As always, we can try to add A and B randomly.

As dimension the matrix gets larger, probability of having aligned eigenvectors goes down.

In the infinite limit, "orthogonal as possible" happens with probability one — the corresponding eigenvalue distributions are known as the *free convolution*.

The free convolution of d copies of a Bernoulli random variable is known as the *Kesten–McKay law*:

$$d\mu(x) = \frac{d}{2\pi} \frac{\sqrt{4(d-1) - x^2}}{d^2 - x^2} \mathbf{1}_{[\lambda_-, \lambda_+]} \, dx$$

where $\lambda_\pm = \pm 2\sqrt{d-1}$.

Finite free probability

Can we do something similar in finite dimensions?

Finite free probability

Can we do something similar in finite dimensions?

Yes — let A and B be $n \times n$ matrices with

$$\det[xI - A] = p(x) \quad \text{and} \quad \det[xI - B] = q(x).$$

Finite free probability

Can we do something similar in finite dimensions?

Yes — let A and B be $n \times n$ matrices with

$$\det [xI - A] = p(x) \quad \text{and} \quad \det [xI - B] = q(x).$$

The *finite free convolution* of p and q is

$$[p \boxplus_n q](x) = \int \det \left[xI - A - RBR^T \right] d\mu(R)$$

where R is a Haar–distributed orthogonal matrix.

Finite free probability

Can we do something similar in finite dimensions?

Yes — let A and B be $n \times n$ matrices with

$$\det [xI - A] = p(x) \quad \text{and} \quad \det [xI - B] = q(x).$$

The *finite free convolution* of p and q is

$$[p \boxplus_n q](x) = \int \det \left[xI - A - RBR^T \right] d\mu(R)$$

where R is a Haar–distributed orthogonal matrix.

The finite free convolution of real rooted polynomials always has real roots! (Borcea, Brändén)

Adjacency matrices

Can we do something similar with *adjacency matrices*?

Adjacency matrices

Can we do something similar with *adjacency matrices*?

Theorem (Quadrature for Laplacian matrices)

Let A and B be $n \times n$ matrices such that $A\mathbb{1} = a\mathbb{1}$ and $B\mathbb{1} = b\mathbb{1}$ (where $\mathbb{1}$ is the all-1 vector) such that

$$\det[xI - A] = (x - a)p(x) \quad \text{and} \quad \det[xI - B] = (x - b)q(x).$$

Then

$$\mathbb{E}_P\left[\det\left[xI - A - PBP^T\right]\right] = (x - a - b)[p \boxplus_{n-1} q]$$

where P is a uniformly distributed permutation matrix.

Adjacency matrices

Can we do something similar with *adjacency matrices*?

Theorem (Quadrature for Laplacian matrices)

Let A and B be $n \times n$ matrices such that $A\mathbb{1} = a\mathbb{1}$ and $B\mathbb{1} = b\mathbb{1}$ (where $\mathbb{1}$ is the all-1 vector) such that

$$\det[xI - A] = (x-a)p(x) \quad \textit{and} \quad \det[xI - B] = (x-b)q(x).$$

Then

$$\mathbb{E}_P\left[\det\left[xI - A - PBP^T\right]\right] = (x-a-b)[p \boxplus_{n-1} q]$$

where P is a uniformly distributed permutation matrix.

Since this is also an expected characteristic polynomial, we can use interlacing families!

Interlacing Families

Theorem (Quadrature for Laplacian matrices, extended)

Let A and B be $n \times n$ matrices with $A\mathbb{1} = a\mathbb{1}$ and $B\mathbb{1} = b\mathbb{1}$ and

$$\det[xI - A] = (x - a)p(x) \quad \textit{and} \quad \det[xI - B] = (x - b)q(x).$$

Then

$$\mathbb{E}_P\left[\det\left[xI - A - PBP^T\right]\right] = (x - a - b)[p \boxplus_{n-1} q](x) := r(x)$$

and there exists a P_0 such that the largest root of

$$\mathbb{E}_P\left[\det\left[xI - A - P_0BP_0^T\right]\right]$$

is smaller than the largest root of $r(x)$.

Interlacing Families

Theorem (Quadrature for Laplacian matrices, extended)

Let A and B be $n \times n$ matrices with $A\mathbb{1} = a\mathbb{1}$ and $B\mathbb{1} = b\mathbb{1}$ and

$$\det[xI - A] = (x - a)p(x) \quad \textit{and} \quad \det[xI - B] = (x - b)q(x).$$

Then

$$\mathbb{E}_P\left[\det\left[xI - A - PBP^T\right]\right] = (x - a - b)[p \boxplus_{n-1} q](x) := r(x)$$

and there exists a P_0 such that the largest root of

$$\mathbb{E}_P\left[\det\left[xI - A - P_0BP_0^T\right]\right]$$

is smaller than the largest root of $r(x)$.

Proof uses random “swaps” to build an interlacing family.

Putting everything together

Lastly, we need to compare the free and finite free convolutions:

Theorem

The largest root of the finite free convolution of two eigenvalue distributions always lies inside the spectrum of the free convolution of those distributions.

Putting everything together

Lastly, we need to compare the free and finite free convolutions:

Theorem

The largest root of the finite free convolution of two eigenvalue distributions always lies inside the spectrum of the free convolution of those distributions.

Putting everything together, gives

Theorem

For all d and all even n, there exists a d-regular bipartite Ramanujan graph on n vertices.

Disclosure: because of the bipartiteness, the actual proof requires a more complicated convolution, but the proof idea is similar.

Outline

Further research

Typically, this is when I beg the audience to think about trying to make these polynomial techniques constructive.

Further research

Typically, this is when I beg the audience to think about trying to make these polynomial techniques constructive.

Theorem (Cohen, 2015)

The Ramanujan graphs provided by the second proof are constructible in polynomial time.

Further research

Typically, this is when I beg the audience to think about trying to make these polynomial techniques constructive.

Theorem (Cohen, 2015)

The Ramanujan graphs provided by the second proof are constructible in polynomial time.

Still open:

1. Can the method of interlacing polynomials be extended to bound the largest root from above and the smallest root from below *simultaneously*?

2. Other connections to free probability and/or random matrix theory?

3. Find more interlacing families!!

Thanks

Thank you for inviting me to speak today.

Thanks

Thank you for inviting me to speak today.

And thank you for your attention!