Twice-Ramanujan Sparsifiers

Nikhil Srivastava

MSR India / Yale

with Josh Batson and Dan Spielman
Sparsification

Approximate any graph $G$ by a sparse graph $H$.

- Nontrivial statement about $G$
- $H$ is faster to compute with than $G$
Cut Sparsifiers [Benczur-Karger’96]

$H$ approximates $G$ if for every cut $S \subset V$ the sum of weights of edges leaving $S$ is preserved.

Can find $H$ with $O(n \log n/\varepsilon^2)$ edges in $\tilde{O}(m)$ time.
The Laplacian (quick review)

\[ L_G = D_G - A_G = \sum_{ij \in E} c_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T \]

Quadratic form

\[ x : V \to \mathbb{R} \]

\[ x^T L_G x = \sum_{ij \in E} c_{ij}(x(i) - x(j))^2 \]

Positive semidefinite

Ker\( (L_G) = \text{span}(\mathbf{1}) \) if \( G \) is connected
Cuts and the Quadratic Form

For characteristic vector \( x_S \in \{0, 1\}^n \) of \( S \subseteq V \)

\[
x_S^T L_G x_S = \sum_{ij \in E} c_{ij} (x(i) - x(j))^2
= \sum_{ij \in (S, \overline{S})} c_{ij}
= \text{wt}_G(S, \overline{S})
\]

So BK says:

\[
1 - \epsilon \leq \frac{x^T L_H x}{x^T L_G x} \leq 1 + \epsilon \quad \forall x \in \{0, 1\}^n
\]
A Stronger Notion [ST’04]

For characteristic vector $x_S \in \{0, 1\}^n$, $S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} c_{ij} (x(i) - x(j))^2$$

$$= \sum_{ij \in (S, \overline{S})} c_{ij}$$

$$= wt_G(S, \overline{S})$$

So BK says:

$$1 - \epsilon \leq \frac{x^T L_H x}{x^T L_G x} \leq 1 + \epsilon$$

$\forall x \in \mathbb{R}^n$

$\forall x \in \{0, 1\}^n$
Why?
1. All eigenvalues are preserved

By Courant-Fischer,

\[(1 - \epsilon)\lambda_i(G) \leq \lambda_i(H) \leq (1 + \epsilon)\lambda_i(G)\]

\(G\) and \(H\) have similar eigenvalues.

For spectral purposes, \(G\) and \(H\) are equivalent.
(x^T L x says a lot)

2. Behavior of electrical flows.
   \( (x^T L x = \text{“energy” for potentials } x: V \to R) \)

3. Behavior of random walks: commute times, mixing time, etc.

4. ‘Relative condition number’ in lin-alg.

5. Fast linear system solvers.
   
   strong notion of approximation.
Examples
Example: Sparsify Complete Graph by Ramanujan Expander [LPS,M]

$G$ is complete on $n$ vertices. $\lambda_i(L_G) = n$

$H$ is $d$-regular Ramanujan graph. $\lambda_i(L_H) \sim d$

$\lambda_i\left(\frac{n}{d}L_H\right) \sim n$
Example: Sparsify Complete Graph by Ramanujan Expander [LPS,M]

$G$ is complete on $n$ vertices. $\lambda_i(L_G) = n$

$H$ is $d$-regular Ramanujan graph. $\lambda_i(L_H) \sim d$

$\lambda_i\left(\frac{n}{d}L_H\right) \sim n$

$$\frac{x^T \left(\frac{n}{d}L_H\right) x}{x^T L_G x} \sim 1$$

Each edge has weight $(n/d)$

So, $\frac{n}{d}H$ is a good sparsifier for $G$. 
Example: Dumbell

\[ K_n \xrightarrow{1} K_n \]

\[ d\text{-regular Ramanujan, times } n/d \xrightarrow{1} d\text{-regular Ramanujan, times } n/d \]
Example: Dumbell

\[ G = G_1 + G_2 + G_3 \]

\[ x^T G x = x^T G_1 x + x^T G_2 x + x^T G_3 x \]
Results
Results

*We can do this well for every $G$.  
(upto a factor of 2)*
Previously Known

Expanders/Ramanujan graphs exist:
“There are very sparse $H$ that look like $K_n$”
Previously Known

Expanders/Ramanujan graphs exist:

“There are very sparse $H$ that look like $K_n$”

degree $d$

$$1 \leq \frac{x^T L_H x}{x^T L_{K_n} x} = \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}}$$
New Result

Expanders/Ramanujan graphs exist:
“There are very sparse $H$ that look like $K_n$”

SPARSIFIERS EXIST:
“There are very sparse $H$ that look like any graph $G$.”

Degree $d$

$$1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}}$$

Avg. degree $2d$
New Result

Expanders/Ramanujan graphs exist:
“There are very sparse $H$ that look like $K_n$”

Sparsifiers exist:
“There are very sparse $H$ that look like any graph $G$.”

$1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}}$

Degree $d$

Avg. degree $2d$

Weighted subgraph
New Result

Expander/Ramanujan graphs exist:
“There are very sparse $H$ that look like $K_n$”

Sparsifiers exist:
“There are very sparse $H$ that look like any graph $G$.”

\[
1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}}
\]

degree $d$

avg. degree $2d$

weighted subgraph
New Result

Expanders/Ramanujan graphs exist:

“There are very sparse $H$ that look like $K_n$”

Sparsifiers exist:

“There are very sparse $H$ that look like any graph $G$.”

\[ 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq \frac{1 + \epsilon}{1 - \epsilon} \]

avg. degree \( \frac{8}{\epsilon^2} \)

weighted subgraph
The Method
The Method

(13-approximation with 6n edges.)
Step 1: Reduction to Linear Algebra
Goal

\[
L_G \xrightarrow{\text{Goal}} L_H
\]

\[
1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n
\]
Outer Product Expansion

Recall:

\[ L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T. \]
Outer Product Expansion

Recall:

$$L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T.$$

For a weighted subgraph $H$:

$$L_H = \sum_{e \in E} s_e b_e b_e^T$$

where $s_e = \text{wt}(e)$ in $H$. 
$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$
\[ 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n \]

\[ 1 \leq \lambda(L_{G}^{-1/2} L_H L_{G}^{-1/2}) \leq 13. \]
1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n

1 \leq \lambda(L_G^{-1/2} L_H L_G^{-1/2}) \leq 13.

1 \leq \lambda \left( \sum_{e \in E} s_e L_G^{-1/2} b_e b_e^T L_G^{-1/2} \right) \leq 13.
\[ 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n \]

\[ 1 \leq \lambda \left( L_G^{-1/2} L_H L_G^{-1/2} \right) \leq 13. \]

\[ 1 \leq \lambda \left( \sum_{e \in E} s_e L_G^{-1/2} b_e b_e^T L_G^{-1/2} \right) \leq 13. \]

\[ 1 \leq \lambda \left( \sum_{e \in E} s_e v_e v_e^T \right) \leq 13 \]

with \( v_e = L_G^{-1/2} b_e \).
A closer look at $\mathbf{v}_e$

$$\mathbf{v}_e = L_G^{-1/2} b_e.$$
A closer look at $v_e$

\[
\sum_e v_e v_e^T = L_G^{-1/2} \left( \sum_e b_e b_e^T \right) L_G^{-1/2} = I
\]
A closer look at $\mathbf{v}_e$

"decomposition of identity"

$m$ vectors in $\mathbb{R}^{n-1}$

$$\forall u \quad \sum_e \langle u, \mathbf{v}_e \rangle^2 = 1$$
Choosing a Subgraph
New Goal

\[ \forall u : 1 \leq \sum_e s_e \langle u, v_e \rangle^2 \leq 13 \]
New Goal

\[ 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n \]

\[ \forall u : 1 \leq \sum_e s_e \langle u, v_e \rangle^2 \leq 13 \]
Main theorem

If

\[ \sum_{e} v_{e}v_{e}^{T} = I_{n} \]

then there are scalars \( s_{e} \geq 0 \) with

\[ 1 \leq \lambda(\sum_{e} s_{e}v_{e}v_{e}^{T}) \leq 13 \]

and \( |\{ s_{e} \neq 0 \}| \leq 6n \).
Main theorem

If
\[ \sum_{e} v_{e} v_{e}^T = I_n \]
then there are scalars \( s_e \geq 0 \) with
\[ 1 \leq \lambda(\sum_{e} s_e v_{e} v_{e}^T) \leq \frac{d + 2\sqrt{d-1}}{d-2\sqrt{d-1}} \]
and \( \left| \{ s_e \neq 0 \} \right| \leq dn \)
Main theorem

If

\[ \sum_{e} v_ev_e^T = I_n \]

then there are scalars \( s_e \geq 0 \) with

\[ 1 \leq \lambda(\sum_{e} s_e v_ev_e^T) \leq 13 \]

and \( |\{s_e \neq 0\}| \leq 6n \).
Step 2: Intuition for the proof
Main theorem

If

$$\sum_{e} v_ey_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_{e} s_ey_ev_e^T) \leq 13$$

and $|\{s_e \neq 0\}| \leq 6n$. 
Main theorem

If

\[ \sum_{e} v_e v_e^T = I_n \]
then there are scalars \( s_e \geq 0 \) with

\[ 1 \leq \lambda \left( \sum_{e} s_e v_e v_e^T \right) \leq 13 \]

and \( \left| \{ s_e \neq 0 \} \right| \leq 6n \)

will build this one vector at a time.
What happens when we add a vector?

\[ \lambda(A) \]
Interlacing

\[ \lambda(A) \]

\[ \lambda(A + vv^T) \]
More precisely

Characteristic Polynomial:

\[ p_A(x) = \det(xI - A) \]
More precisely

Characteristic Polynomial:

\[ p_A(x) = \det(xI - A) \]

Matrix-Determinant Lemma:

\[ p_{A + vv^T} = p_A \left( 1 + \sum_i \frac{\langle v, u_i \rangle^2}{\lambda_i - x} \right) \]
More precisely

Characteristic Polynomial:

\[
p_A(x) = \det(xI - A)
\]

Matrix-Determinant Lemma:

\[
p_{A + vv^T} = p_A \left( 1 + \sum \frac{\langle v, u_i \rangle^2}{\lambda_i - x} \right)
\]

\(\lambda(A + vv^T)\) are zeros of this.
Physical model of interlacing

\( \lambda_i = \text{positive unit charges resting at barriers on a slope} \)
Physical model of interlacing

\[ \langle v, u_i \rangle^2 = \text{charges added to barriers} \]

\[ \lambda(A + vv^T) \]
Physical model of interlacing

\[ \langle v, u_i \rangle^2 = \text{charges added to barriers} \]
Physical model of interlacing

Barriers repel eigs.

$$+ \langle v, u_n \rangle^2$$

$$+ \langle v, u_2 \rangle^2$$

$$+ \langle v, u_1 \rangle^2$$

$$\lambda_1$$

$$\lambda_2$$

$$\lambda_3$$

$$\lambda(A + vv^T)$$
Physical model of interlacing

Barriers repel eigs.

Inverse law repulsion

Gravity

$1 + \sum_i \frac{\langle v, u_i \rangle^2}{\lambda_i - x} = 0$
Physical model of interlacing

Barriers repel eigs.

\[ \lambda(A + uv^T) \]
Examples

\[ \lambda(A) \]
Ex1: All weight on $u_1$
Ex1: All weight on $\mathbf{u}_1$
Ex1: All weight on $u_1$
Ex1: All weight on $u_1$

$$\lambda(A + vv^T)$$
Ex2: Equal weight on $u_1, u_2$
Ex2: Equal weight on $u_1, u_2$
Ex2: Equal weight on $u_1, u_2$

\[
\lambda(A + vv^T)
\]
Ex3: Equal weight on all $u_1, u_2, \ldots, u_n$
Ex3: Equal weight on all $u_1, u_2, \ldots, u_n$

$$\lambda(A + vv^T)$$
Adding a balanced vector

\[ p_A + \nu \nu^t = p_A \left( 1 + \sum_i \frac{\langle \nu, u_i \rangle^2}{\lambda_i - x} \right) \]

\[ = p_A \left( 1 + \sum_i \frac{1}{\lambda_i - x} \right) \]

\[ = p_A - p'_A \]
Consider a random vector

If

$$\sum_e v_e v_e^T = I$$

For every $u_i$:

$$\sum_e \langle v_e, u_i \rangle^2 = 1.$$
Consider a random vector

If

\[ \sum_e v_e v_e^T = I \]

For every \( u_i \):

\[ \sum_e \langle v_e, u_i \rangle^2 = 1. \]

thus a ‘random’ vector has the same expected projection in every direction \( i \):

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m \]
Ideal proof

\[ A^{(0)} = 0 \]

\[ p^{(0)} = x^n \]
Ideal proof

$$E_e \langle v_e, u_i \rangle^2 = \frac{1}{m}$$

\[ A(0) = 0 \]
\[ p(0) = x^n \]
$A^{(1)} = 0 + vv^T$

$p^{(1)} = x^n - nx^{n-1}$

$\mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m}$
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A^{(2)} = A^{(1)} + vv^T \]

\[ p^{(2)} = x^n - 2nx^{n-1} + n(n - 1)x^{n-2} \]
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A^{(3)} = A^{(2)} + vv^T \]

\[ p^{(3)} = p^{(2)} - p^{(2)'} \]
\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A(i+1) = A(i) + vv^T \]

\[ p(i+1) = p(i) - p(i)' \]
Ideal proof

\[ A(i+1) = A(i) + \nu \nu^T \]

\[ p(i+1) = p(i) - p(i)' \]

\[ \mathbb{E}_e \langle \nu_e, u_i \rangle^2 = 1/m \]
\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m \]

\[ A(i+1) = A(i) + vv^T \]

\[ p(i+1) = p(i) - p(i)' \]
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[
A^{(i+1)} = A^{(i)} + vv^T
\]

\[
p^{(i+1)} = p^{(i)} - p^{(i)}'
\]
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A^{(i+1)} = A^{(i)} + vv^T \]

\[ p^{(i+1)} = p^{(i)} - p^{(i)'} \]
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A(i+1) = A(i) + vv^T \]

\[ p(i+1) = p(i) - p(i)' \]
Ideal proof

\[ E_e \langle \nu_e, u_i \rangle^2 = 1/m \]

\[
A(i+1) = A(i) + \nu \nu^T
\]

\[
p(i+1) = p(i) - p(i)'
\]
Ideal proof

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m \]

\[ A(i+1) = A(i) + vv^T \]

\[ p(i+1) = p(i) - p(i)' \]
\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ A(i+1) = A(i) + vv^T \]

\[ p(i+1) = p(i) - p(i)' \]

\[ \frac{\lambda_n(A)}{\lambda_1(A)} \leq 13? \]
\[ p(i) = \text{Laguerre}^{(i)} \]

\[ p(i+1) = p(i) - p(i)' \]

\[ E_e \langle v_e, u_i \rangle^2 = 1/m \]

\[ \frac{\lambda_n(A)}{\lambda_1(A)} \leq 13? \]
Punch Line

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ p^{(i)} = \text{Laguerre}^{(i)} \]

\[ p^{(i+1)} = p^{(i)} - p^{(i)'} \]

In \( dn \) steps:

\[ \frac{\lambda_n(A)}{\lambda_1(A)} \leq \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-2}} \]
find actual vectors that realize this ideal behavior.

\[ \mathbb{E}_e \langle v_e, u_i \rangle^2 = \frac{1}{m} \]

\[ p^{(i)} = \text{Laguerre}^{(i)} \]

\[ p^{(i+1)} = p^{(i)} - p^{(i)'} \quad \lambda_n(A) < 1 \]

In \( dn \) steps: \[ \frac{\lambda_n(A)}{\lambda_1(A)} \leq \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-2}} \]
Step 3: Actual Proof
(for 6n vectors, 13-approx)
Broad outline: moving barriers

\[ A = \emptyset \]
Step 1

\[ A = \emptyset \]

\[ +vv^T \quad v \in \{v_e\} \]
Step 1

\[ A = \emptyset \]

\[ + vv^T \quad v \in \{ v_e \} \]

\[ A = vv^T \]
Step 1

\[ A = \emptyset \]

\[ + v v^T \quad v \in \{v_e\} \]

\[ A = v v^T \]

\[ -n \quad 0 \quad n \]

\[ -n + \frac{1}{3} \quad 0 \quad n + 2 \]
Step 1

\[ A = \emptyset \]

\[ A = vvv^T \quad v \in \{v_e\} \]

+1/3

+n+1/3

-1

0

looser constraint

tighter constraint
Step $i+1$

$A^{(i)}$

$0$

$\leq \lambda_i \leq$
Step $i+1$

$A^{(i)}$

$\begin{array}{c}
+1/3 \\
\hline
+2 \\
\hline
\end{array}$

$\begin{array}{c}
\downarrow \quad \downarrow \\
0 \\
\downarrow \quad \downarrow \\
\uparrow \quad \uparrow \\
\begin{bmatrix} v \\ v \end{bmatrix}^T \\
\hline
\end{array}$
Step $i+1$

$A(i), A(i+1)$

$0$

$\leq \lambda_i \leq$
Step $i+1$

$A(i), A(i+1)$

$+\frac{1}{3}$

$+2$

$\text{vv}^T$
Step $i+1$

$A(i), A(i+1), A(i+2)$

$\lambda_i \leq \lambda_i \leq$
Step $i+1$

$A(i), A(i+1), A(i+2)$

$\text{+1/3}$

$\text{+2}$

$0$

$\mathbf{v}v^T$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3)$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$0 \leq \lambda_i \leq$
Step i+1

\[ A(i), A(i+1), A(i+2), A(i+3), \ldots \]
Step 6n

$A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n)$
Step 6n

\[ A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n) \]

13-approximation with 6n vectors.
Problem

need to show that an appropriate

\[ v_e v_e^T \]

always exists.
Problem

need to show that an appropriate $v_0 v_e^T v_e$ always exists.

is not strong enough to do the induction.
Problem

need to show that an appropriate

need a better way to measure quality of eigenvalues.

\[ \leq \lambda_i \leq \]

is not strong enough to do the induction.
The Upper Barrier

\[ \Phi^u(A) = \text{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i} \]
The Upper Barrier

\[ \Phi^u(A) = \text{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i} \]

\[ \Phi^u(A) \leq 1 \implies \lambda_{\max}(A) \ll u \]
The Upper Barrier

$$\Phi^u(A) = \text{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i}$$

- No $\lambda_i$ within dist. 1
- No 2 $\lambda_i$ within dist. 2
- No 3 $\lambda_i$ within dist. 3

No $k \lambda_i$ within dist. $k$

$$\Phi^u(A) \leq 1 \Rightarrow \lambda_{\text{max}}(A) \ll u$$
The Upper Barrier

\[ \Phi^u(A) = \text{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i} \]

No \( \lambda_i \) within dist. 1
No 2 \( \lambda_i \) within dist. 2
No 3 \( \lambda_i \) within dist. 3

\[ \Phi^u(A) \leq 1 \Rightarrow \lambda_{\text{max}}(A) \ll u \]
The Lower Barrier

\[ \Phi_\ell(A) = \text{Tr}(A - \ell I)^{-1} = \sum_i \frac{1}{\lambda_i - \ell} \]

\[ \Phi_\ell(A) \leq 1 \implies \lambda_{\text{min}}(A) \gg \ell \]
The Beginning

\[ A = \emptyset \]
The Beginning

\[ A = \emptyset \]

\[ \Phi^n(\emptyset) = \text{Tr}(nI)^{-1} = 1 \]

\[ \Phi_{-n}(\emptyset) = \text{Tr}(nI)^{-1} = 1. \]
Step $i+1$

$A(i), A(i+1), A(i+2)$

$\Phi^u(A) \leq 1$
$\Phi_\ell(A) \leq 1.$
Step i+1

$A(i), A(i+1), A(i+2)$

Lemma.

can always choose $s v v^T$ so that potentials do not increase

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1.$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3)$

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1$.
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1.$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1$. 
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1$
Step 6n

$A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n)$

$\Phi^u(A) \leq 1$
$\Phi_\ell(A) \leq 1.$
Step 6n

\[ A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n) \]

13-approximation with 6n vectors.
Lemma.

Can always choose $+svvv^T$ so that both potentials do not increase.

$\Phi^u(A) \leq 1$
$\Phi_\ell(A) \leq 1$. 

Goal
The Right Question

“Which vector should we add?”
The Right Question

“Which vector should we add?”

“Given a vector, how much of it can we add?”
Upper Barrier Update

Add $svv^T$ & set $u' \leftarrow u + 2$. 
Upper Barrier Update

Add \( svv^T \) & set \( u' \leftarrow u + 2 \).

\[
\Phi_{u'}(A + svv^T) = \text{Tr}(u'I - A - svv^T)^{-1}
\]
Upper Barrier Update

Add $svv^T$ & set $u' \leftarrow u + 2$.

$$
\Phi^{u'}(A + svv^T)
= \text{Tr}(u'I - A - svv^T)^{-1}
$$

$$
\text{Tr}(A + vv^T)^{-1} = \text{Tr}A^{-1} - \frac{v^TA^{-2}v}{1 + v^TA^{-1}v}
$$

Sherman-Morrisson
Upper Barrier Update

Add $s v v^T$ & set $u' \leftarrow u + 2$.

$\Phi^{u'}(A + s v v^T)$

$= \text{Tr}(u' I - A - s v v^T)^{-1}$

$= \Phi^{u'}(A) + \frac{v^T (u' I - A)^{-2} v}{1/s - v^T (u' I - A)^{-1} v}$
Upper Barrier Update

\[ \text{Add } \mathbf{s} \mathbf{v} \mathbf{v}^T \text{ & set } u' \leftarrow u + 2. \]

\[ \Phi^{u'}(A + s \mathbf{v} \mathbf{v}^T) \]

\[ = \text{Tr}(u'I - A - s \mathbf{v} \mathbf{v}^T)^{-1} \]

\[ = \Phi^{u'}(A) \frac{\mathbf{v}^T(u'I - A)^{-2} \mathbf{v}}{1/s - \mathbf{v}^T(u'I - A)^{-1} \mathbf{v}} \]

\[ \text{want } \leq \Phi^{u}(A). \]
How much of \( \mathbf{vv}^T \) can we add?

Rearranging:

\[
\Phi^{u'}(A + svv^T) \leq \Phi^u(A)
\]

\[
\iff \frac{1}{s} \geq v^T \left( \frac{(u'\mathbf{I} - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + (u'\mathbf{I} - A)^{-1} \right) v
\]
How much of $\mathbf{vv}^T$ can we add?

Rearranging:

$$\Phi^{u'}(A + sv\mathbf{v}\mathbf{v}^T) \leq \Phi^u(A)$$

$$\iff$$

$$\frac{1}{s} \geq \mathbf{v}^T \left( \frac{(u'I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + (u'I - A)^{-1} \right) \mathbf{v}$$

$$\frac{1}{s} \geq \mathbf{U}_A \bullet \mathbf{vv}^T$$
The Lower Barrier

Similarly:

$$\Phi_{\ell'}(A + svv^T) \leq \Phi_\ell(A)$$

$$\iff$$

$$\frac{1}{s} \leq v^T \left( \frac{(A - \ell'I)^{-2}}{\Phi_{\ell'}(A) - \Phi_\ell(A)} - (A - \ell'I)^{-1} \right) v$$

$$\frac{1}{s} \leq L_A \bullet vv^T$$
Goal

Show that we can always add some vector while respecting both barriers.
Both Barriers

There is always a vector with

\[ U_A \bullet \vv \vv^T \leq L_A \bullet \vv \vv^T \]
Both Barriers

There is always a vector $v$ with

$$U_A \cdot vv^T \leq L_A \cdot vv^T$$

**can add**

**must add**
There is always a vector $v$ with

$$U_A \cdot vv^T \leq L_A \cdot vv^T$$

Then, can squeeze scaling factor in between:

$$U_A \cdot vv^T \leq \frac{1}{s} \leq L_A \cdot vv^T$$
Taking Averages

\[ \exists \mathbf{v}, U_A \cdot \mathbf{vv}^T \leq L_A \cdot \mathbf{vv}^T \]

\[ \sum_{\mathbf{v} \in \{\mathbf{v}_e\}} U_A \cdot \mathbf{vv}^T = U_A \cdot \left( \sum_e \mathbf{v}_e \mathbf{v}_e^T \right) \]

\[ = U_A \cdot I \]

\[ = \text{Tr}(U_A). \]
Taking Averages

$$\exists v, U_A \cdot vv^T \leq L_A \cdot vv^T$$

$$\sum_{v \in \{v_e\}} U_A \cdot vv^T = U_A \cdot \left( \sum_e v_e v_e^T \right)$$

$$U_A \cdot I = \text{Tr}(U_A).$$
Bounding $\text{Tr}(U_A)$

$$\frac{\text{Tr}(u' I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + \text{Tr}(u' I - A)^{-1}$$
Bounding $\text{Tr}(U_A)$

$$\frac{\text{Tr}(u'I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + \Phi^{u'}(A)$$
Bounding $\text{Tr}(U_A)$

$$\frac{\text{Tr}(u' I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + \leq \Phi^u(A)$$
Bounding $\text{Tr}(U_A)$

\[ \frac{\text{Tr}(u' I - A)^{-2}}{\Phi^u(A) - \Phi^{u'}(A)} + \leq 1 \text{ induction} \]
Bounding $\text{Tr}(U_A)$

\[
-\frac{\partial}{\partial u'} \Phi u'(A) \leq 1
\]

(Recall $\Phi^u(A) = \text{Tr}(uI - A)^{-1}$.)
Bounding $\text{Tr}(U_A)$

\[
\begin{aligned}
&-\frac{\partial}{\partial u'} \Phi u'(A) \\
\geq &\; \delta_u \left( -\frac{\partial}{\partial u'} \Phi u'(A) \right)
\end{aligned}
\]

\[\leq 1 \quad \text{induction} \]

(Recall $\Phi^u(A) = \text{Tr}(uI - A)^{-1}$.)
Bounding $\text{Tr}(U_A)$

\[
-\frac{\partial}{\partial u'} \Phi u'(A) \geq \delta_u \left( -\frac{\partial}{\partial u'} \Phi u'(A) \right)
\]

$\leq 1$

induction

convexity

$\text{Tr}(U_A) \leq \frac{1}{\delta_u} + 1$
Taking Averages

\[ \exists v, U_A \bullet vv^T \leq L_A \bullet vv^T \]

\[ \sum_{v \in \{v_e\}} U_A \bullet vv^T \leq \frac{1}{\delta_u} + 1. \]
Taking Averages

\[ \exists \mathbf{v}, U_A \cdot \mathbf{vv}^T \leq L_A \cdot \mathbf{vv}^T \]

\[ \sum_{\mathbf{v} \in \{v_e\}} U_A \cdot \mathbf{vv}^T \leq \frac{1}{\delta_u} + 1. \]

\[ \sum_{\mathbf{v} \in \{v_e\}} L_A \cdot \mathbf{vv}^T \geq \frac{1}{\delta_l} - 1. \]
Taking Averages

$$\exists v, \quad U_A \bullet vv^T \leq L_A \bullet vv^T$$

$$\sum_{v \in \{v_e\}} U_A \bullet vv^T \leq \frac{1}{2} + 1.$$

$$= \frac{3}{2}$$

$$\sum_{v \in \{v_e\}} L_A \bullet vv^T \geq \frac{1}{\delta \ell} - 1.$$
Taking Averages

\[ \exists v, U_A \cdot vv^T \leq L_A \cdot vv^T \]

\[ \sum_{v \in \{v_e\}} U_A \cdot vv^T \leq \frac{1}{2} + 1. \]

\[ = \frac{3}{2} \]

\[ \sum_{v \in \{v_e\}} L_A \cdot vv^T \geq \frac{1}{\frac{1}{3}} - 1. \]

\[ = 2 \]
Taking Averages

\[ \exists v, U_A \bullet vv^T \leq L_A \bullet vv^T \]

\[ \sum_{v \in \{v_e\}} U_A \bullet vv^T \leq \frac{1}{2} + 1. \quad = \frac{3}{2} \]

\[ \sum_{v \in \{v_e\}} L_A \bullet vv^T \geq \frac{1}{\frac{1}{3}} - 1. \quad 2 \]
Step $i+1$

$A(i), A(i+1), A(i+2)$

Lemma.
can always choose $\sum_{i} vvv^T$ so that potentials do not increase.

$\Phi^u(A) \leq 1$

$\Phi_{\ell}(A) \leq 1.$
Step i+1

$A(i), A(i+1), A(i+2), A(i+3)$

$\Phi^u(A) \leq 1$

$\Phi_\ell(A) \leq 1$
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$$\Phi^u(A) \leq 1$$
$$\Phi^\ell(A) \leq 1.$$
Step i+1

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$\Phi^u(A) \leq 1$

$\Phi^\ell(A) \leq 1$. 
Step $i+1$

$A(i), A(i+1), A(i+2), A(i+3), \ldots$

$\phi^u(A) \leq 1$

$\phi_\ell(A) \leq 1.$
Step 6n

\[ A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n) \]

\[ \Phi^u(A) \leq 1 \]

\[ \Phi_\ell(A) \leq 1. \]
Step 6n

$A(i), A(i+1), A(i+2), A(i+3), \ldots, A(6n)$

13-approximation with 6n vectors.
Main Sparsification Theorem

If

$$\sum_{e} v_{e}v_{e}^{T} = I_{n}$$

then there are scalars $s_{e} \geq 0$ with

$$1 \leq \lambda(\sum_{e} s_{e}v_{e}v_{e}^{T}) \leq 13$$

and $|\{s_{e} \neq 0\}| \leq 6n$. 
Sparsification of Graphs

\[ 1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n \]

\[ \forall u : 1 \leq \sum_e s_e \langle u, v_e \rangle^2 \leq 13 \]
Twice-Ramanujan

Fixing $dn$ steps and tightening parameters gives

$$\frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}}.$$  

(zeros of Laguerre polynomials).
Major Themes

• Electrical model of **interlacing** is useful

• Can use barrier potential to **iteratively** construct matrices with desired spectra

• Analysis of progress is **greedy / local**

• Requires **fractional weights** on vectors
Sparsification of PSD Matrices

Theorem. If

\[ V = \sum_i v_i v_i^T \]

then there are scalars \( s_i \geq 0 \) for which

\[ V \preceq \sum_i s_i v_i v_i^T \preceq (1 + \varepsilon)V \]

and at most \( n/\varepsilon^2 \) are nonzero.
To put this in context...

Given: \[ V = \sum_{i \leq m} v_i v_i^T \]

Spectral Theorem: If rank(V) = n then

\[ V = \sum_{i \leq n} \lambda_i u_i u_i^T \]

for eigenvectors \( u_i \).
To put this in context...

Given: \[ V = \sum_{i \leq m} v_i v_i^T \]

Spectral Theorem: If \( \text{rank}(V) = n \) then

\[ V = \sum_{i \leq n} \lambda_i u_i u_i^T \]

\( u_i \) need not be ‘meaningful’ directions...
(e.g., \( v_i = \) edges of graph)
To put this in context...

Given: \[ V = \sum_{i \leq m} v_i v_i^T \]

Spectral Theorem: If \( \text{rank}(V) = n \) then
\[ V = \sum_{i \leq n} \lambda_i u_i u_i^T \]

This Theorem. Can find scalars \( s_i \) so that:
\[ V \sim \epsilon \sum_i s_i v_i v_i^T \]
Open Questions

• The Ramanujan bound
• Unweighted sparsifiers for $K_n$
• A faster algorithm
• Directed graphs? (must be weaker notion)
• The Kadison-Singer Conjecture
Open Questions

• The Ramanujan bound
• Unweighted sparsifiers for $K_n$
• A faster algorithm
• Directed graphs? (must be weaker notion)
• The Kadison-Singer Conjecture
Sums of Outer Products

\[ A = \sum_{i \leq m} v_i v_i^T \quad v_i \in \mathbb{R}^n, m \gg n \]
Sums of Outer Products

\[ A = \sum_{i \leq m} v_i v_i^T \quad v_i \in \mathbb{R}^n, m \gg n \]

- elementary /
- “meaningful” directions
Sums of Outer Products

\[ A = \sum_{i \leq m} v_i v_i^T \quad v_i \in \mathbb{R}^n, m \gg n \]

• **Q**: Can we write \( A \) as a *sparse* sum?
Sums of Outer Products

\[ A = \sum_{i \leq m} v_i v_i^T \quad v_i \in \mathbb{R}^n, \, m \gg n \]

• **Q:** Can we write \( A \) as a \textit{sparse} sum?

• **A:** Yes, Spectral Theorem:

\[ A = \sum_{i \leq n} \lambda_i u_i u_i^T \text{ for eigvecs } u_i \]
Sums of Outer Products

\[ A = \sum_{i \leq m} v_i v_i^T \quad v_i \in \mathbb{R}^n, m \gg n \]

• **Q:** Can we write A as a *sparse* sum?
• **A:** Yes, Spectral Theorem:

\[ A = \sum_{i \leq n} \lambda_i u_i u_i^T \text{ for eigvecs } u_i \]

• **Good:** only \( n \) terms = optimal!
• **Bad:** \( u_i \) hard to interpret in terms of \( v_i \)
Example: Graphs

Undirected graph $G(V,E)$

Laplacian Matrix

$$L_G = \sum_{ij \in E} (e_i - e_j)(e_i - e_j)^T$$
Example: Graphs

Undirected graph $G(V,E)$

Laplacian Matrix

$$L_G = \sum_{i,j \in E} (e_i - e_j)(e_i - e_j)^T$$

Interested in `sparse approximation`: write $L_G$ as sum of a small number of edges.
Example: Graphs

Undirected graph $G(V,E)$

Laplacian Matrix

$$L_G = \sum_{i,j \in E} (e_i - e_j)(e_i - e_j)^T$$

Interested in `sparse approximation`: write $L_G$ as sum of a small number of edges.

Spectral Thm: $L_G = \sum_{i=1:n} \lambda_i u_i u_i^T$

Sparse, but $u_i$ do not correspond to edges... :(
Spectral Sparsification [BSS’09]

Theorem. Given $A = \sum_{i \leq m} v_i v_i^T$

there are nonnegative weights $s_i \geq 0$ s.t.

$$A \sim \tilde{A} = \sum_i s_i v_i v_i^T$$

and at most $1.1n$ of the $s_i$ are nonzero.
Spectral Sparsification [BSS’09]

Theorem. Given $A = \sum_{i \leq m} v_i v_i^T$
there are nonnegative weights $s_i \geq 0$ s.t.

\[ A \sim \tilde{A} = \sum_i s_i v_i v_i^T \]

and at most $1.1n$ of the $s_i$ are nonzero.

• same $v_i$

• cf. $n$ terms for $\lambda_i u_i u_i^T$
Back to Graphs

Undirected graph $G(V,E)$

Laplacian Matrix

$$L_G = \sum_{i,j \in E} (e_i - e_j)(e_i - e_j)^T$$
Back to Graphs

Undirected graph $G(V,E)$

Laplacian Matrix

$$L_G = \sum_{i,j \in E} (e_i - e_j)(e_i - e_j)^T$$

Apply Theorem:

$$\tilde{L}_G = \sum_{i,j \in E} s_i (e_i - e_j)(e_i - e_j)^T$$

$$L \sim \tilde{L}$$

\( \leq 1.1n \text{ edges!} \)