$G = (V,E)$ undirected graph.
$w : E \rightarrow \mathbb{R}^+$ edge weights.
Recall $H = (V,F)$ with $w' : F \rightarrow \mathbb{R}^+$ is an $\varepsilon$-approximate spectral sparsifier of $G$ if

$$(1-\varepsilon) L_G \leq L_H \leq (1+\varepsilon) L_G.$$

In previous lecture we saw a random sampling algorithm that shows that for any $\varepsilon > 0$ $\varepsilon$-approximate spectral sparsifiers with $O(\sqrt{n \log n})$ edges exist.
Recall the algorithm:

We see

\[ L_a = \sum_{e \in E} \mathbf{e} \mathbf{e}^T \]

where \( L_{uv} = (\mathbf{x}_u - \mathbf{x}_v) (\mathbf{x}_u - \mathbf{x}_v)^T \) is a rank-1 matrix corresponding to edge \( uv \).

We obtained \( H \) by importance sampling:

\[ L_H = \sum_{e \in E} \mathbf{x}_e \]

where \( \mathbf{x}_e = \frac{\mathbf{e}}{p(e)} \mathbf{e} \) with \( p(e) \) the probability.

\[ = 0 \quad \text{otherwise}. \]
And we chose $\frac{1}{e} = \frac{\ln n}{\varepsilon^2}$ Reff(e).

Since the "importance" or "influence" of we Le or Za is Reff(e) captured by the two lemmas.

1. we Le $\leq$ Reff(e) Za

2. $\sum \text{Reff}(e) = n - 1$.

The extra $\frac{\ln n}{\varepsilon^2}$ in the probability is standard when using Chernoff-type concentration inequalities and is necessary if one uses...
independent random sampling.

Two questions:
1. Can we derandomize?
2. Is $\frac{n \log n}{\varepsilon^2}$ a tight bound?

One standard derandomization strategy when using Chernoff-type inequalities is to use pessimistic estimators. In the context of Chernoff bounds, this yields a variant of the well-known Multiplicative
Weight Update method that (MWU)
relies on exponential potential function. In the context of Matrix Chernoff bound one needs to use the Matrix MWU which is technical but most likely yields a deterministic algorithm but since it essentially relies on the analysis of the randomized algorithm and will also lead to a sparsifier with \( \frac{n \ln n}{\delta^2} \) edges in the
Work Case.

[Batson-Spielman-Seiuradaran] used another potential function and an elegant analysis to obtain a deterministic algorithm that yields a sparsifier with $O\left(\frac{n}{\epsilon^2}\right)$ edges which is optimal. The initial algorithm was slow ($n^4$) but subsequent work has yielded faster deterministic and randomized algorithms.
Recall that $L_H \cong L_\alpha$ iff all eigenvalues of $L_H$ are roughly same as those of $L_\alpha$.

Even in the sampling algorithm we found it easier to work with $L_\alpha L_\alpha L_\alpha L_\alpha = \mathbb{I}$ which is identity on space $1 \to \mathbb{I}$.

The reason for this is that the min and max eigen values of $L_\alpha$ are $1$ and hence requiring that $\lambda_{\min}$ and $\lambda_{\max}$
are well-approximated ensures that all eigen values are approximated.

To this end this is the main theorem of DASS.

Main Theorem

Let \( \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_m \in \mathbb{R}^n \) s.t.

\[
\sum_{i=1}^{m} v_i v_i^T = I \quad \text{(are m (sparse) eigenvectors)}
\]

Then for any \( \varepsilon > 0 \) a set \( S \subseteq [m] \) and \( C_i, i \in S \) s.t.

\[
|S| = O\left( \frac{n}{\varepsilon^2} \right)
\]
Exercise: Use above theorem to derive $O\left(\frac{n}{\xi^2}\right)$ sized $\varepsilon$-approximate spectral spanifiers of a $n$-vector graph.

Hint:

$$L^+ L^+ = \sum_{e \in E} L^+ w_e L^+ = \sum_{e \in E} \sum_{i \in e} w_i v_i v_i^T.$$
Prof. Main Theme

The algorithmic idea is the following.
- Start with matrix $A = 0$ (zero).
- For some $# I$ steps do
  - Find an index $i \in [m]$ and a step size $c$ and update $A \leftarrow A + cv_i v_i^T$.

Basically, it follows an iterative scheme where, in each step...
We add a small amount of some $\bar{V}_i V_i^T$ to current matrix $A$ to get it close to $I$.

Recall we want $\lambda_{\min}(A) \geq 1 + \varepsilon$

and $\lambda_{\max}(A) \leq 1 + \varepsilon$

at the end.

Of course we can simply add all of the $V_i$'s with step size $l$ but then we won't sparsify.

Idea is to use $O\left(\frac{n}{\varepsilon^2}\right)$ iterations

so we get a sparsifier.
Some mathematical background

\[ \text{Trace } (A) = \sum_{i=1}^{n} A_{ii} \quad \text{sum of diagonal entries of a } n \times n \text{ matrix.} \]

1. \( \text{Trace } (A) = \sum \lambda_i \) if eigenvalues.
   Why? \( \lambda \)'s are roots of characteristic poly det \( (A - \lambda I) = 0 \).

2. \( \text{Trace } (A+B) = \text{Trace } (A+B) \) clean.

3. Less clean: cyclic property
   \( \text{Trace } (AB) = \text{Trace } (BA) \) whenever
   \( AB \) is a square matrix
   \( A \) & \( B \) may not be square!}
Easy to verify.

**Lemma:**  \( \sum_{i=1}^{m} v_i^T M \bar{v}_i = \text{Tr}(\sum_{i=1}^{m} v_i v_i^T M) \)

and hence if \( \sum_{i=1}^{m} v_i v_i^T = I \) then

\[
\sum_{i=1}^{m} \bar{v}_i^T M \bar{v}_i = \text{Tr}(M).
\]

**Proof:**  \( \sum_{i=1}^{m} v_i^T M \bar{v}_i = \sum_{i=1}^{m} \text{Tr}(v_i v_i^T M \bar{v}_i) \)

\[
= \sum_{i=1}^{m} \text{Tr}(v_i v_i^T M) \quad \text{by cyclic prop}
\]

\[
= \text{Tr}(\sum_{i=1}^{m} v_i v_i^T M)
\]
\[ = \text{Tr} (IM) = \text{Tr} (M). \]

**Sherman-Morrison-Woodbury Formula**

**Theorem:** Let \( A \) be a symmetric non-singular matrix and let \( C \) be a real \# and \( \mathbf{v} \) be a vector. Then

\[
(A - c \mathbf{v} \mathbf{v}^T)^{-1} = A^{-1} + c \frac{A^{-1} \mathbf{v} \mathbf{v}^T A^{-1}}{1 - c \mathbf{v}^T A^{-1} \mathbf{v}}.
\]

**Proof:** Check by multiplying.

or view \( A \) as \( \Sigma \lambda \mathbf{u} \mathbf{u}^T \).

\( \Box \)
Barrier functions

We define two barrier functions, "upper" and "lower".

Let $A$ be a $n \times n$ symmetric matrix with eigenvalues

$$\lambda_1 \leq \lambda_2 \ldots \leq \lambda_n.$$

For $u > \lambda_n$ we let

$$\Phi_u(A) = \sum_{i=1}^{n} \frac{1}{u - \lambda_i}$$

$$= \text{Tr}(uI - A)^{-1}.$$ 

For $1 \leq \lambda_i \leq \lambda_1$ we let

$$\Phi_l(A) = \sum_{i=1}^{n} \frac{1}{\lambda_i - l}.$$
\[ T_2 \left( \left( A - l I \right)^{-1} \right). \]

As \( u \to \infty \) \( \phi_u (A) \to 0 \)
and as \( u \to \infty \) \( \phi_u (A) \to 0 \).

As \( l \to -\infty \) \( \phi_l (A) \to 0 \)
\( l \to \lambda_1 \) \( \phi_l (A) \to \infty \).

We need to understand how the barrier functions change as \( u \) and \( l \) change.
Claim: \( \phi_{u+s}(A) < \phi_u(A) \) for \( s > 0 \)

Claim: \( \lambda_n < u - \frac{1}{\phi_u(A)} \)

Claim: \( l + \frac{1}{\phi_e(A)} \leq \lambda_1 \)

More technical

Claim: Let \( l < \lambda_1 \) and \( s < \frac{1}{\phi_e(A)} \)

\[ \phi_{l+s}(A) \leq \frac{1}{\frac{1}{\phi_e(A)} - s} \]

Basic

Need algebra. See Spielman notes.
How do we use barrier function?

Initially, $A = 0 \cdot \lambda_1 = \lambda_2 \cdots = \lambda_n = 0$.

We set $l_0 = -n$

and $u_0 = n$

So that $\Phi_{u_0}(A) = \Phi_{l_0}(A) = 1$.

Our goal is eventually to make $A$ have the property that

$$\frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)} = O(1).$$

Ideally $1 - \varepsilon$.
Plan: Show that in each iteration we can find scalars $c$ and $v_i$ and $S_u$ and $S_e \leq 1$

$$\Phi_u(A + c v_i v_i^T) \leq \Phi_u(A) \leq 1$$
$$\Phi_e(A + 1 - v_i v_i^T) \leq \Phi_e(A) \leq 1$$

We want $S_u$ and $S_e$ to be fixed constants. Suppose we can do that. We can run algorithm for $\Omega(n)$ steps.

For concreteness say $S_u = 2, S_e = \frac{1}{3}$, and we run for $\Omega(n)$ steps.
At end \( u = 13n \) and \( \Phi(13n) \leq l \).

\[
\Rightarrow \quad \exists n \leq u - \frac{1}{\Phi_n(A)} \leq 13n
\]

\[
\ell \geq -n + 6n \times \frac{1}{3} \geq n.
\]

\[
\lambda_1 \geq \ell + \frac{1}{\Phi_e(A)} \geq n.
\]

\[
\Rightarrow \quad \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)} \leq 13.
\]

We need to scale \( A \) by \( 6n \) to normalize.
How to show existence of $c_{Su, Sd}$? Note that for given $c, v_i, Su, Sd$ we can compute all desired quantities and check if conditions hold. So issue is proving existence.

---

**Updating barrier function**

We want to change $A ightarrow A + c v_i v_i^T$ rank-1 update. How does $f_A$ change?
\[ \Phi_u(A) = \text{Tr} \left( (uI - A)^{-1} \right) \]
\[ \Phi_u(A + cu^T v^T) = \text{Tr} \left[ (uI - A - cu^T v^T)^{-1} \right] \]

We use SMW formula \( \Phi \).

\[ = \text{Tr} \left[ (uI - A)^{-1} + c \frac{(uI - A)^{-1} u v^T (uI - A)^{-1}}{1 - c u^T (uI - A)^{-1} v} \right] \]
\[ = \Phi_u(A) + \text{Tr} \left( u v^T (uI - A)^{-1} \right) \]
\[ = \Phi_u(A) + \frac{c u^T (uI - A)^{-1} v}{1 - c u^T (uI - A)^{-1} v} \]

Because \( \text{Tr} (XY) = \text{Tr} (YX) \).
\[
\phi_u(A) + \frac{c \, U^T (uI - A)^{-2} u}{1 - c U^T (uI - A)^{-1} u} \\
\Rightarrow \\
\text{the since } c > 0 \\
\text{and PSD } b(uI - A).
\]

So adding \( c \, U \, U^T \) increases \( \phi_u \). (Intuitively \( \lambda \)'s increase \\
\( \Rightarrow \sum \frac{1}{u - \lambda} \).)

To keep \( \phi_u(A + c \, U \, U^T) \leq 1 \) \\
we increase \( u \) to \( u' = u + \delta u \).

Q: how much should we increase \( u \) so?
\[ \phi_{u+\delta u}(A+c u \bar{v} \bar{v}^T) \leq \phi_u(A). \]

From above derivation:

\[ \phi_u(A+c u \bar{v} \bar{v}^T) \]

\[ = \Phi_u(A) + c \frac{u^T (u I - A)^{-2} \bar{v}}{1 - c v^T (u I - A)^{-1} v} \]

Thus to ensure

\[ \phi_u(A+c u \bar{v} \bar{v}^T) \leq \phi_u(A) \]

it suffices to have
\[
\Phi_u(A) - \Phi_u'(A) \geq \frac{c v^T (u' I - A)^{-2} \bar{u}^T}{1 - c \bar{u} (u' I - A) \bar{u}^T}
\]

\[
\Rightarrow \frac{1}{c} \geq \frac{v^T (u' I - A)^{-2} \bar{u}}{\Phi_u(A) - \Phi_u'(A)} + \bar{u}^T (u' I - A)^{-1} \bar{u}.
\]

Let \( U_A = \frac{(u' I - A)^{-2}}{\Phi_u(A) - \Phi_u'(A)} + (u' I - A)^{-1} \)

Lemma: Fix \( u, \delta u \) and \( \bar{v} \)
Then \( \max c \ s.t. \Phi_u(A + c u v^T) \leq \Phi_u(A) \) is given by
\[
\frac{1}{c} \geq v^T U_A \bar{v}.
\]
Very clean dependence on $A$, $u$ and $fu$ and $v$ is “outside”.

What happens when we increase $l$ to $l'$? $l'=l+\delta l$. $\Phi_\ell(A)$ increases, but when we add $c\bar{u}\bar{u}^T$ to $A$ $\Phi_\ell(A + (-\bar{u}\bar{u}^T))$ decreases (since $\ell$'s intuited to increase).

How much? Etc.
Want: $\phi_{\ell_1}(A + t \mathbf{v} \mathbf{v}^T) \leq \phi_{\ell_1}(A)$.

Lemma:
Let $L_A = \frac{(A - l'\mathbf{I})^{-2}}{\phi_{\ell_1}(A) - \phi_{\ell_1}(A)}$.

If $\frac{1}{c} \leq \mathbf{v}^T L_A \mathbf{v}$ then
$\phi_{\ell_1}(A + c \mathbf{v} \mathbf{v}^T) \leq \phi_{\ell_1}(A)$. 
**Inductive argument.**

Given \( u, A, \delta u, \delta \) and \( \vec{v} \) we have seen that

\[ \Phi_{u+\delta u} (A + c \; u \; \vec{v}^T) \leq \Phi_{u} (A) \]

if \( \frac{1}{c} \geq \vec{v}^T U_A \vec{v} \)

and \( \Phi_{\delta + \delta} (A + c \; u \; \vec{v}^T) \leq \Phi_{\delta} (A) \)

if \( \frac{1}{c} \leq \vec{v}^T L_A \vec{v} \).

Recall \( \sum_{i=1}^{m} \nu_i \; u_i \vec{v}_i^T = I \).
Fix $\delta_u$ and $\delta_e$.

Want to find scalar $c > 0$ and $\overline{V}_i \in S$ such we can update $A$ to $A + c \overline{V}_i \overline{V}_i^T$ without violating barriers.

Lemma: $\sum_{i=1}^{m} \overline{V}_i^T L_A \overline{V}_i \leq \frac{1}{\delta_u} + \phi_u(A)$.

Lemma: $\sum \overline{V}_i^T L_A \overline{V}_i \geq \frac{1}{\delta_e} - \frac{1}{\phi_e(A)} - \delta_e$.
Thus if we have that

\[ \frac{1}{S_u} + \phi_u(A) \leq \frac{1}{S_l} - \phi_l(A). \]

\[ \Rightarrow \exists \, \nu_i \text{ s.t. } \overline{\nu}_i^T U A \overline{\nu}_i \leq \overline{\nu}_i^T L_A \overline{\nu}_i \]

and

Then we can find c and \( \overline{\nu}_i \) to add to A.
Parametric setting

d fixed constant \( d = \frac{1}{e^2} \).

\( l_0 = -\sqrt{d} \ n \quad \phi_{l_0}(0) = \frac{1}{\sqrt{d}} \)

\( u_0 = \frac{d + \sqrt{d}}{\sqrt{d} - 1} \ n \quad \phi_{u_0}(0) = \frac{\sqrt{d} - 1}{d + \sqrt{d}} \)

\( \delta_l = 1 \quad \delta_u = \frac{\sqrt{d} + 1}{\sqrt{d} - 1} \).

Initially:

\[ \frac{1}{\delta_u} + \phi_u(A) = \frac{\sqrt{d} - 1}{\sqrt{d} + 1} + \frac{\sqrt{d} - 1}{d + \sqrt{d}}. \]

\[ = \frac{\sqrt{d} - 1}{\sqrt{d} + 1} \left(1 + \frac{1}{\sqrt{d}}\right). \]

\[ = \frac{\sqrt{d} - 1}{\sqrt{d}} = 1 - \frac{1}{\sqrt{d}} \]
\[ \frac{1}{\Delta t} - \frac{1}{\phi(0) - \delta t} = 1 - \frac{1}{\sqrt{d} - 1} \]

Hence we satisfy the condition.

Now run algorithm for \( d \) iterations. Let \( A \) be initial matrix.

\[ \frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)} \leq \frac{u_0 + d n \delta u}{l_0 + d n \delta c} \]

\[ \leq \frac{d + 2 \sqrt{d} + 1}{d - 2 \sqrt{d} + 1} \]

For \( d = \frac{1}{\varepsilon^2} \), above is
\[ \frac{\frac{1}{\varepsilon^2} + 2 \frac{1}{\varepsilon} + 1}{\frac{1}{\varepsilon^2} - 2 \frac{1}{\varepsilon} + 1} \]
\[ x = \frac{1 + 2\varepsilon + \varepsilon^2}{1 - 2\varepsilon + 3\varepsilon^2}. \]
Proof of Lemmas

Lemma: \[ \sum_{i=1}^{m} u_i^T u_A v_i \leq \frac{1}{\delta_u} + \phi_u(A). \]

Proof: Recall:
\[ \sum_{i=1}^{m} u_i^T u_A v_i = \text{Tr}(u_A) \]

\[ u_A = \frac{(u' I - A)^{-2}}{\phi_u(A)} + (u' I - A)^{-1} \]

\[ \text{Tr}((u' I - A)^{-1}) = \phi_u'(A) \leq \phi_u(A). \]

\[ \text{Tr}((u' I - A)^{-2}) ? \]
\[ \phi_u(A) = \sum_{i=1}^{n} \frac{1}{u - \lambda_i} \]

\[ \frac{d}{du} \phi_u(A) = -\sum \frac{1}{(u - \lambda_i)^2} \]

\[ = -\text{Tr}((uI-A)^{-2}) \]

Since \( \phi_u(A) \) is convex in \( u \), we have:

\[ \phi_u(A) - \phi_{u+\delta u}(A) \geq -\delta u \phi_u'(A) \]

\[ \geq \delta u \text{Tr}((u'\delta - A)^{-2}) \]

\[ = \frac{1}{\delta u} \geq \frac{\text{Tr}((u'\delta - A)^{-2})}{\phi_u(A) - \phi_{u+\delta u}(A)} \]

\[ \square \]
Lemma: \[ \sum_{i=1}^{m} V_i^T L_A V_i \geq \frac{1}{\delta \epsilon} - \phi(A). \]

Proof: Recall \[ \sum_{i=1}^{m} V_i^T L_A V_i = \text{Tr}(L_A). \]

\[ L_A = \frac{(A - \lambda I)^{-2}}{\phi(A) - \phi(A)} - (A - \lambda I)^{-1} \]

\[ \text{Tr}(L_A) = \text{Tr}(\nabla ) + \text{Tr}(\nabla ) \]
\[ T_x \left( A - \ell^1 I \right)^{-1} = \Phi_{\ell^1}(A) \preceq \Phi_{\ell}(A). \]

\[ T_x \left( (A - \ell^1 I)^{-2} \right). \]

\[ \Phi_{\ell}(A) = \sum_{i=1}^{n} \frac{1}{\lambda_i - \ell} \]

\[ \frac{d}{d\ell} \Phi_{\ell}(A) = \sum_{i=1}^{n} \frac{1}{(\lambda_i - \ell)^2} \]

\[ \Phi_{\ell}(A) \text{ is convex.} \quad \forall \ell. \]

\[ \Rightarrow \quad \frac{\Phi_{\ell+\delta \ell}(A) - \Phi_{\ell}(A)}{\delta \ell} \leq \Phi_{\ell+\delta \ell}^{-1}(A) \]

\[ \Rightarrow \quad T_x \left( (A - \ell^1 I)^{-2} \right) \frac{1}{\Phi_{\ell^1}(A) - \Phi_{\ell}(A)} \geq \frac{1}{\delta \ell}. \]