$G = (V, E)$ \quad \omega : E \rightarrow \mathbb{R}^+$

We saw cut sparsifiers.

$H = (V, F)$ with $\omega' : F \rightarrow \mathbb{R}^+$ is an $\varepsilon$-cut sparsifier if

$\forall S \subseteq V$

\[
\omega'(S) \approx (1 \pm \varepsilon) \omega(S).
\]

$H$ is an $\varepsilon$-approximate spectral sparsifier if $\forall \overline{x} \in \mathbb{R}^n$

\[
\overline{x}^T L_H \overline{x} \approx (1 \pm \varepsilon) \overline{x}^T L_H \overline{x}
\]
Since $\bar{x}^T L \bar{x} = O(d(G))$
   when $\bar{x} = \chi_S$
   for $S \subseteq V$

an $\varepsilon$-spectral sparsifier is an
$\varepsilon$-cut sparsifier.

However, examples show that
a cut sparsifier is not always
a spectral sparsifier.

See below.
Example which shows cut spannifier does not imply spectral spannifier.

Fix \( n \) and some parameter \( k \).

\[ \begin{align*}
\text{Edges:} & \quad (i-j) \text{ mod } n \\
& \quad \leq k \\
& \quad \text{and } (1, \frac{n}{2})
\end{align*} \]

\( G' \) is graph \( G \)-edge \((1, \frac{n}{2})\).

Claim: \( \min \text{ cut of } G' \) is \( \Theta(k) \)

\( \Rightarrow G' \) is \( \varepsilon \)-cut spannifier for \( G \)

with \( \varepsilon = O\left(\frac{1}{k}\right) \).
Is $A'$ a good spectral sparsifier for $A$?

Consider $\bar{x} = (0, 1, 2, \ldots, \frac{n}{2} - 1, \frac{n}{2} - 1, \ldots, 1, 0)$

$$\bar{x}^T L A \bar{x} = \Theta(n K^3).$$

$$\bar{x}^T L_{2n} \bar{x} = \Theta(n K^3) + \left(\frac{n}{2} - 1\right)^2$$

If $k$ is constant, very different.

So edge $(1, \frac{n}{2})$ important to retain in spectral sparsifiers.
Random Sampling Band Scheme

[Spielman-Srivastava].

Inspired/motivated by cut specification

Recall \( \mathcal{L}_n = \sum_{e \in E} \mathcal{L}_e \)

Want space \( \mathcal{H} \) s.t.

\[
(1-\varepsilon) \mathcal{L}_n \leq \mathcal{L}_{\mathcal{H}} \leq (1+\varepsilon) \mathcal{L}_n.
\]

- Pick \( p_e \) for each edge \( e \) with probability \( p_e \).
- Sample each edge \( e \) independently with probability \( p_e \).
  If \( e \) chosen set \( w_e = \frac{w_e}{p_e} \)
Output $H$ as the chosen edge weights.

Proof relies on matrix Chernoff inequality.

Let $X_e = \frac{w_e}{p_e} 1_{e \in H} \text{ with } p_e = 0$ otherwise.

$$L_H = \sum_e X_e$$

$$\mathbb{E} [L_H] = L_a$$

Want to use matrix Chernoff inequalities to argue that
LH is an independent sum of random rank-1 matrices and hence by concentration
"behaves" like $L_a$ whp.

What should $p_e$ be?
Based on the setting of regular Chernoff bound for real values $p_e$ should correspond to the
"importance" of $L_e$ in the

sum $L_a = \sum_{e \in E} L_e$

Is there such a measure?
Turns out the $p_e = \text{Reff}(e)$ the effective resistance is quite useful here. Why?

Lemma 1: $L_e \leq \text{Reff}(e) \leq L_a$

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

To get concentration we will need to set $p_e = \frac{c \log n}{\varepsilon^2} \text{Reff}(e)$.

From Lemma 2 it is easy to see that # of edges chosen in expectation is $O\left(\frac{n \log n}{\varepsilon^2}\right)$. 
The main part is to understand Lemma 1. For this we need some background on what it means that

$$\hat{x}^T A \hat{x} \leq \hat{x}^T B \hat{x} \quad \forall \hat{x} \in \mathbb{R}^n.$$
Background on PSD matrices

Recall PSD matrices satisfy
\[ \mathbf{A} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \], where \( \mathbf{A} \in \mathbb{R}^{n \times n} \) is symmetric.

\( \mathbf{A} \geq 0 \) is notation to indicate \( \mathbf{A} \) is PSD.

all eigen values are real and non-negative because of PSD.

The PSD property gives rise to a partial order on \( n \times n \) symmetric matrices.
We say \( A \geq B \) if
\[
\bar{x}^T A \bar{x} \geq \bar{x}^T B \bar{x} + \bar{x}^T C \bar{x}
\]
\[\Rightarrow A - B \geq 0.\]

\( A \geq B \) and \( B \geq C \)
\[\Rightarrow A \geq C.\]

also \( A \geq B \) \(\Rightarrow\) \( A + C \geq B + C \)

for all symmetric \( A, B, C \).

One can define for graphs
\( A \geq H \) if
\( L_A \geq L_H. \)
Observation: $A, B, C$ symmetric matrices $(n \times n)$.

$$A \succeq B \Rightarrow CAC^T \succeq CBC^T$$

If $C$ is non-singular then

$$CAC^T \succeq CBC^T \Rightarrow A \succeq B.$$ 

Proof: Check $\vec{x}^T A \vec{x} \geq \vec{x}^T B \vec{x}$
via $\vec{x}^T CAC^T \vec{x} \geq \vec{x}^T CBC^T \vec{x}$

$\Box$
Note that \( \bar{x}^T L \bar{x} = \sum_{i,j} w_{ij} (x_i - x_j)^2 \).

Therefore, if we reduce weights of edges, then we obtain a new graph \( H \) such that 
\[ G \supset H. \]

**Lemma:** Suppose \( L \geq c L_H \) for some constant \( c > 0 \). Then \[ \lambda_k(L) \geq \lambda_k(L_H) + k. \]

**Proof:** From Courant-Fischer theorem,
\[ \lambda_k(L) = \min_{S \subseteq \mathbb{R}^n} \max_{\tilde{x} \in S} \frac{\tilde{x}^T L \tilde{x}}{\tilde{x}^T \tilde{x}} \] subject to \( |S| = k \).
\[
\begin{align*}
\max_{S \subseteq \mathbb{R}^n} & \quad \frac{x^T L H x}{x^T x} \\
\text{subject to} & \quad \text{dim}(S) = k
\end{align*}
\]
Matrix Norms

Given an $n \times n$ matrix $A$, one can define several norms:

1) **Frobenius norm** $\|A\|_F$:

$$\|A\|_F = \sqrt{\sum_{ij} A_{ij}^2}$$

which is same as $l_2$ norm if $A$ viewed as a $n^2 \times 1$ vector.

2) **Spectral norm** based on how $A$ operates:

$$\|A\|_{\text{spectr}} = \max_{\|x\|_2 = 1} \|Ax\|_2$$

It is the largest singular value of $A$.

For PSD matrices it is $\sqrt{\lambda_{\text{max}}}$. 
Matrix Chernoff Inequality

Recall Chernoff inequality

\( X_1, X_2, \ldots, X_n \) independent random variables in \([0, R]\) (typically \( R = 1 \))

Let \( X = \sum_{i=1}^{n} X_i \).

\[ E[X] = \sum_{i=1}^{n} E[X_i] = \mu \text{ say.} \]

Let \( \mu_{\text{min}} \leq \mu \leq \mu_{\text{max}} \).

Then for all \( \delta > 0 \):

1. \( \Pr \left[ X > (1+\delta) \mu_{\text{max}} \right] \leq \left[ \frac{e^\delta}{(1+\delta)^{1+\delta}} \right] \frac{\mu_{\text{max}}}{R} \)

2. \( \Pr \left[ X \leq (1-\delta) \mu_{\text{min}} \right] \leq \left[ \frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right] \frac{\mu_{\text{min}}}{R} \).
For $0 \leq s \leq 1$ we can simplify:

\[
\frac{1}{e} \leq e^{s^2 \lambda_{\text{max}} / 3R}
\]

\[
\frac{1}{e} \leq e^{s^2 \lambda_{\text{min}} / 2R}
\]

Typically we assume $R = 1$. We can scale and get the results for general $R$ from the one for $R = 1$.

\underline{Matrix version}

Can we obtain concentration bounds for matrix sums?
Theorem [Trpp] Let $X_1, \ldots, X_k$ be independent, random symmetric real matrices of size $d \times d$ with $0 \leq X_i \leq R I$. $X = \sum_{i=1}^{n} X_i$.

Let $\mu_{\min} I \leq \mathbb{E}[X] \leq \mu_{\max} I$.

Then for all $\delta > 0$,

1. $\mathbb{P} \left[ \lambda_{\max}(X) \geq (1+\delta) \mu_{\max} \right]$

   \[ \leq d \cdot \left[ \frac{e^{\delta}}{(1+\delta)^{g+\delta}} \right] \mu_{\max} \]

   \[ \leq d e^{-\delta^2 \mu_{\max} / 2R} \quad (\text{if } \delta \leq 1) \]

2. $\mathbb{P} \left[ \lambda_{\min}(X) \leq (1-\delta) \mu_{\min} \right]$

   \[ \leq d \left( \frac{e^{-\delta}}{(1-\delta)^{d-1}} \right) \mu_{\min} \]

   \[ \leq d \left( \frac{e^{-\delta}}{(1-\delta)^{d-1}} \right) \frac{\mu_{\min}}{R} \]
Useful corollary

\[ \lambda_{\min} = \lambda_{\max} = 1 \]

\[ \implies \mathbb{E}[X] = 1 \]

Then

\[ \mathbb{P}_{\mathbf{X}} [\lambda_{\min}(X) \leq (1 - \delta)] \leq e^{-\delta^2/2n} \]

\[ \mathbb{P}_{\mathbf{X}} [\lambda_{\max}(X) \geq (1 + \delta)] \leq e^{-\delta^2/3n} \]

\[ \implies \mathbb{P}_{\mathbf{X}} [(1 - \delta)I \leq X \leq (1 + \delta)I] \geq 1 - 2n e^{-\delta^2/3n} \]
We would like to apply Matrix Chernoff inequality to prove $L_H = L_a$. 

Recall that $L_H = L_a$ 

$(\Leftarrow)$ all eigenvalues are similar in both.

However Matrix Chernoff only quantifies $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$.

For identity matrix $I$ we have $\lambda_{\text{min}} = \lambda_{\text{max}} = 1$ and approximately preserving $\lambda_{\text{min}}$ & $\lambda_{\text{max}} \Leftarrow$ all eigenvalues preserved closely.
Transformation

Let $A$, $B$ be positive definite matrices.

$$A \leq (1+\varepsilon) B \iff B^{-\frac{1}{2}} A B^{-\frac{1}{2}} \leq (1+\varepsilon) I$$

We would like to apply above for PSD matrices but a PSD matrix may not have an inverse but pseudo inverse will work.

Claim: $L^H \leq (1+\varepsilon) L$

$$\iff L^H L + L L^H \leq L^H L + \varepsilon L L^H$$

$L^H$ is square root of pseudo inverse.
Let $\Pi = L^{+\frac{1}{2}} L L^{+\frac{1}{2}}$

$\Pi$ is not the identity matrix but is the identity operator in the space orthogonal to the null space of $L$ which in $L$.

\[
L^{+\frac{1}{2}} = \sum_{i=1}^{\nu} \frac{1}{\lambda_i} \bar{u}_i \bar{u}_i^T
\]

Hence for any $\bar{\nu} \perp \bar{I}$

\[
(L^{+\frac{1}{2}} L L^{+\frac{1}{2}}) \bar{v} = 0.
\]

\[
= \lambda_{\text{max}} \text{ and } \lambda_{\text{min}} \text{ is operator } \bar{\nu} = 1.
\]
Want to prove Lemma 1.  

\[ L_H \simeq (1 \pm \varepsilon) L_a. \]

Assume we = 1 \text{ per E for simplicity.}

\[ L_H = \sum_e X_e \quad \mathbb{E}[X_e] = 1e \]

\[ L_a = \sum_e L_e \]

Consider \( L_a \leftrightarrow L_H \leftrightarrow L_a \)

\[ = \sum_e L_a^{+\frac{1}{2}} X_e L_a^{+\frac{1}{2}} \]

\[ \mathbb{E}[L_a^{+\frac{1}{2}} L_H L_a^{+\frac{1}{2}}] = \mathbb{I} = L_a^{+\frac{1}{2}} L_a^{+\frac{1}{2}} \]
We want to use Tropp's inequalities.

We set \( p_e = \frac{\text{Reff}(e)}{\mathcal{R}} \)

where \( \mathcal{R} = \frac{c e^2}{\ln n} \)

Let \( Y_e = L_e^{+1/2} X_e L_e^{-1/2} \)

We saw that\( L_e \preceq \text{Reff}(e) L_e \)

\[ L_e^{+1/2} X_e L_e^{-1/2} \preceq \text{Reff}(e) L_e^{+1/2} \]

Recall \( Y_e = \frac{1}{p_e} L_e^{+1/2} X_e L_e^{-1/2} \)

with \( p_e \) such that \( p_e = 0 \) otherwise.
Hence $Y_e = \frac{R}{\text{Reff}(e)} \cdot L_a \cdot L_e \cdot L_a$.

Hence $\|Y_{e\ell}\| \leq \frac{R}{\text{Reff}(e)} \cdot \text{Reff}(e) \leq R$.

We apply Matrix Chernoff to the sum $\sum L_{a\ell} \cdot Y_{e\ell} \cdot L_a$.

$E[C_{\ell}] = \overline{T_{\ell}}$.

$\lambda_{\text{min}}(T_{\ell}) = \lambda_{\text{max}}(T_{\ell}) = 1$

viewed as a symmetric operator over $n$-dimensional space.
Orthogonal to $I$.

$$\| L^{1/2} y e L^{1/2} \| \leq R$$

$$\Rightarrow \text{with prob. } 1 - 2n e^{-\frac{3^2}{8R}}$$

we have

$$(1-\epsilon) L^{1/2} e L^{1/2} L^{1/2} H L^{1/2} \leq (1+\epsilon) L^{1/2} e L^{1/2}$$

If $R \geq \frac{C k n}{\epsilon^2}$ for sufficiently large $c$ then done. Why.

$\square$
Proof of Lemma

Lemma 1: \( L_e = \text{Reff}(e) \cdot L_a \).

Proof: Want to prove

Consider \( \bar{v}^T L_e \bar{v} = (v_i - v_j)^2 \)

where \( e = (i, j) \)

Wlog assume \( v_i = 1 \) and \( v_j = 0 \),
in which case \( \bar{v}^T L_e \bar{v} = 1 \).

What is \( \bar{v}^T L_a \bar{v} \)?

\[
\bar{v}^T L_a \bar{v} \geq \min_{v_i = 1, v_j = 0} \sum_{a \neq b} (v_a - v_b)^2

\geq \text{Ceff}(ij) = \frac{1}{\text{Reff}(ij)}

(\text{volumes}).
Above is a sketch. See notes.

Lemma 2: \( \sum_{e} \text{Reff}(e) = n-1 \)

Proof: \( \text{Reff}(S,t) = (X_s - X_t)^T L^+ (X_s - X_t) \)

Recall.

\[ \sum_{u,v} \text{Reff}(u,v) = \sum_{u,v} (X_u - X_v)^T L^+ (X_u - X_v) \]

[Simplicity assume \( w_e = 1 \), can create parallel edges if necessary.]
We recall $\text{Trace}(A) = \sum_i A_{ii}$, the sum of diagonal entries.

$\text{Trace}(A) =$ sum of eigenvalues of $A$ (even when $A$ may have complex values)

Follows from fact that eigenvalues are solutions to $\det(A - \lambda I) = 0$

Coefficient $\beta \sum_i \lambda_i = \sum_i A_{ii}$.

$$\sum_{\text{UVCE}} \text{Reff}(u,v) = \sum_{\text{UVCE}} (x_i - x_{ij})^T L^T (x_i - x_{ij})$$

$$= \sum_{\text{UVCE}} \text{Trace} \left( \begin{array}{c} L^T \\ \end{array} \right)$$
(by cyclic property of trace) \[ \text{Trace}(ABC) = \text{Trace}(CAB) = \text{Trace}(CA(\Lambda + AC)) \]

\[ = \sum_{\text{UVCE}} \text{Trace}\left( (X_i - Xu)(X_i - Xu)^{\top}L^+ \right) \]

\[ = \text{Trace}\left( L L^+ \right) \]

Recall \[ L = \sum_{i=1}^{n} \lambda_i \bar{u}_i \bar{u}_i^{\top} \]

\[ L^+ = \sum_{i=2}^{n} \frac{1}{\lambda_i} \bar{u}_i \bar{u}_i^{\top} \]

\[ \lambda_1 = 0, \quad \lambda_2 = \sigma_{\text{trace}} \]
\[ LL^T = \sum_{i=2}^{\infty} u_i u_i^T \]

So essentially identify in the space orthogonal to \( \bar{u}_1 \).

So eigen values are \( 0, 1, \ldots, 1 \)

\[ \text{Trace} (LL^T) = n-1. \]