CS5275 – The Algorithm Designer's Toolkit (S2 AY2025/26)

Lecture 5:

Expanders – Cheeger inequality

λ_2 and conductance

- **Recall:** $\lambda_2 = 0$ if and only if G has at least 2 connected components.

The second eigenvalue of the normalized Laplacian $\boldsymbol{N}$ of G.

λ_2 and conductance

- **Recall:** $\lambda_2 = 0$ if and only if G has at least 2 connected components.

  The second eigenvalue of the normalized Laplacian $\boldsymbol{N}$ of G.

- Intuitively, there should a robust version of this fact:
  - λ_2 is small if and only if G has a sparse cut.

λ_2 and conductance

- **Recall:** $\lambda_2 = 0$ if and only if G has at least 2 connected components.

  The second eigenvalue of the normalized Laplacian $\boldsymbol{N}$ of G.

- Intuitively, there should a robust version of this fact:
  - λ_2 is small if and only if G has a sparse cut.

- We will prove this: **Cheeger inequality:** $\frac{\lambda_2}{2} \leq \Phi(G) \leq \sqrt{2\lambda_2}$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Let $(S, V \setminus S)$ be a sparsest cut: $\Phi(S) = \Phi(G)$.

$$R_{\boldsymbol{N}}(\mathbf{1}_S) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)}$$

$$R_{\boldsymbol{N}}(\mathbf{1}_{V \setminus S}) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(V \setminus S)}$$

Rayleigh quotient of $\boldsymbol{x}$ with respect to $\boldsymbol{N}$:

$$R_{\boldsymbol{N}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{N} \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{d \sum_{v \in V} {x_v}^2}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Let $(S, V \setminus S)$ be a sparsest cut: $\Phi(S) = \Phi(G)$.

$$R_{\boldsymbol{N}}(\mathbf{1}_S) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)}$$

$$R_{\boldsymbol{N}}(\mathbf{1}_{V \setminus S}) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(V \setminus S)}$$

$$\lambda_2 = \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x})$$

$$\leq \max_{\boldsymbol{z} \in \mathrm{span}(\mathbf{1}_S, \mathbf{1}_{V \setminus S}) \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{z})$$

Rayleigh quotient of $\boldsymbol{x}$ with respect to $\boldsymbol{N}$:

$$R_{\boldsymbol{N}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{N} \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{d \sum_{v \in V} {x_v}^2}$$

Recall: $\lambda_2 = \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x})$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Let $(S, V \setminus S)$ be a sparsest cut: $\Phi(S) = \Phi(G)$.

$$R_{\boldsymbol{N}}(\mathbf{1}_S) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)}$$

$$R_{\boldsymbol{N}}(\mathbf{1}_{V \setminus S}) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(V \setminus S)}$$

$$\lambda_2 = \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x})$$

$$\leq \max_{\boldsymbol{z} \in \mathrm{span}(\mathbf{1}_S, \mathbf{1}_{V \setminus S}) \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{z})$$

$$\leq 2 \max\{R_{\boldsymbol{N}}(\mathbf{1}_S), R_{\boldsymbol{N}}(\mathbf{1}_{V \setminus S})\}$$

Rayleigh quotient of $\boldsymbol{x}$ with respect to $\boldsymbol{N}$:

$$R_{\boldsymbol{N}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{N} \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{d \sum_{v \in V} {x_v}^2}$$

Recall: $\lambda_2 = \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x})$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \mathrm{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\mathbf{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Let $(S, V \setminus S)$ be a sparsest cut: $\Phi(S) = \Phi(G)$.

$$R_{\boldsymbol{N}}(\mathbf{1}_S) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)}$$

$$R_{\boldsymbol{N}}(\mathbf{1}_{V\setminus S}) = \frac{|E(S, V \setminus S)|}{\mathrm{vol}(V \setminus S)}$$

Rayleigh quotient of $\boldsymbol{x}$ with respect to $\boldsymbol{N}$:

$$R_{\boldsymbol{N}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\top \boldsymbol{N} \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} = \frac{\sum_{\{u,v\}\in E}(x_u - x_v)^2}{d \sum_{v\in V} {x_v}^2}$$

$$\begin{aligned}
\lambda_2 &= \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x}\in\mathcal{V}\setminus\{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x}) \\
&\leq \max_{\boldsymbol{z}\in \mathrm{span}(\mathbf{1}_S, \mathbf{1}_{V\setminus S})\setminus\{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{z}) \\
&\leq 2 \max\{R_{\boldsymbol{N}}(\mathbf{1}_S), R_{\boldsymbol{N}}(\mathbf{1}_{V\setminus S})\} \\
&= 2 \cdot \frac{|E(S, V \setminus S)|}{\min\{\mathrm{vol}(S), \mathrm{vol}(V \setminus S)\}} \\
&= 2 \cdot \Phi(S) \\
&= 2 \cdot \Phi(G)
\end{aligned}$$

Recall: $\lambda_2 = \min_{2-\text{dim space } \mathcal{V}} \max_{\boldsymbol{x}\in\mathcal{V}\setminus\{\mathbf{0}\}} R_{\boldsymbol{N}}(\boldsymbol{x})$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \mathrm{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\mathbf{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Proof of the lemma:

- It suffices to prove it for the case of $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.

$$R_{\boldsymbol{M}}(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) \leq 2\max\{R_{\boldsymbol{M}}(\alpha \boldsymbol{x}), R_{\boldsymbol{M}}(\beta \boldsymbol{y})\} = 2\max\{R_{\boldsymbol{M}}(\boldsymbol{x}), R_{\boldsymbol{M}}(\boldsymbol{y})\}$$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \operatorname{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2\max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Proof of the lemma:

- It suffices to prove it for the case of $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.

$\boldsymbol{N}$ is a real symmetric $(n \times n)$-matrix:

- Eigenvalues: $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$
- Eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$

We write:

- $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$
- $\boldsymbol{y} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \operatorname{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

$$R_{\boldsymbol{N}}(\boldsymbol{z}) = \frac{\boldsymbol{z}^\top \boldsymbol{N} \boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}}$$
$$= \frac{\sum_{i=1}^{n} \lambda_i (a_i + b_i)^2}{\|\boldsymbol{x} + \boldsymbol{y}\|^2}$$

Proof of the lemma:

- It suffices to prove it for the case of $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.

$\boldsymbol{N}$ is a real symmetric $(n \times n)$-matrix:

- Eigenvalues: $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$
- Eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$

We write:

- $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$
- $\boldsymbol{y} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \operatorname{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Proof of the lemma:

- It suffices to prove it for the case of $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.

$\boldsymbol{N}$ is a real symmetric $(n \times n)$-matrix:

- Eigenvalues: $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$
- Eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$

We write:

- $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$
- $\boldsymbol{y} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i$

$$R_{\boldsymbol{N}}(\boldsymbol{z}) = \frac{\boldsymbol{z}^\top \boldsymbol{N} \boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}} = \frac{\sum_{i=1}^{n} \lambda_i (a_i + b_i)^2}{\|\boldsymbol{x} + \boldsymbol{y}\|^2} \leq \frac{2\sum_{i=1}^{n} \lambda_i \left(a_i^2 + b_i^2\right)}{\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2}$$

$\boldsymbol{x} \perp \boldsymbol{y} \implies \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$

$(a_i + b_i)^2 \leq 2a_i^2 + 2b_i^2$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \operatorname{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 1: $\frac{\lambda_2}{2} \leq \Phi(G)$

Proof of the lemma:

- It suffices to prove it for the case of $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$.

$\boldsymbol{N}$ is a real symmetric $(n \times n)$-matrix:

- Eigenvalues: $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$
- Eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$

We write:

- $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{v}_i$
- $\boldsymbol{y} = \sum_{i=1}^n b_i \boldsymbol{v}_i$

$$
\begin{aligned}
R_{\boldsymbol{N}}(\boldsymbol{z}) &= \frac{\boldsymbol{z}^\top \boldsymbol{N} \boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{z}} \\
&= \frac{\sum_{i=1}^n \lambda_i (a_i + b_i)^2}{\|\boldsymbol{x} + \boldsymbol{y}\|^2} \\
&\leq \frac{2 \sum_{i=1}^n \lambda_i \left(a_i^2 + b_i^2\right)}{\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2} \\
&= \frac{2(\boldsymbol{x}^\top \boldsymbol{N} \boldsymbol{x} + \boldsymbol{y}^\top \boldsymbol{N} \boldsymbol{y})}{\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2} \\
&\leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}
\end{aligned}
$$

Lemma:

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two orthogonal vectors.
- Let $\boldsymbol{z} \in \operatorname{span}(\boldsymbol{x}, \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$.

$$R_{\boldsymbol{N}}(\boldsymbol{z}) \leq 2 \max\{R_{\boldsymbol{N}}(\boldsymbol{x}), R_{\boldsymbol{N}}(\boldsymbol{y})\}$$

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:

- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:

- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Setting $\boldsymbol{x} = \boldsymbol{v}_2$ yields a cut with conductance at most $\sqrt{2\lambda_2}$.

$\Phi(G) \leq \sqrt{2\lambda_2}$

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:
- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Proof of the lemma:
- $\boldsymbol{x}' = \boldsymbol{x} - m\mathbf{1}$
  - m is the median of $\boldsymbol{x}$.
- $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$
  - $x_v^+ = \begin{cases} x_v' & \text{if} \quad x_v' > 0 \\ 0 & \text{otherwise} \end{cases}$
  - $x_v^- = \begin{cases} -x_v' & \text{if} \quad x_v' < 0 \\ 0 & \text{otherwise} \end{cases}$

At least half of the
vertices have zero values.

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:

- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Claim 1:
$R_N(\boldsymbol{x}) \geq R_N(\boldsymbol{x}') \geq \min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}$

Proof of the lemma:

- $\boldsymbol{x}' = \boldsymbol{x} - m\mathbf{1}$
  - m is the median of $\boldsymbol{x}$.
- $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$
  - $x_v^+ = \begin{cases} x_v' & \text{if} \quad x_v' > 0 \\ 0 & \text{otherwise} \end{cases}$
  - $x_v^- = \begin{cases} -x_v' & \text{if} \quad x_v' < 0 \\ 0 & \text{otherwise} \end{cases}$

At least half of the
vertices have zero values.

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:

- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Proof of the lemma:

- $\boldsymbol{x}' = \boldsymbol{x} - m\mathbf{1}$
  - m is the median of $\boldsymbol{x}$.
- $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$
  - $x_v^+ = \begin{cases} x_v' & \text{if } \ x_v' > 0 \\ 0 & \text{otherwise} \end{cases}$
  - $x_v^- = \begin{cases} -x_v' & \text{if } \ x_v' < 0 \\ 0 & \text{otherwise} \end{cases}$

At least half of the vertices have zero values.

Claim 1:

$R_N(\boldsymbol{x}) \geq R_N(\boldsymbol{x}') \geq \min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}$

Claim 2:

Let $\boldsymbol{y} \in \mathbb{R}^n_{\geq 0}$. There exists a threshold $t > 0$ such that $S_t = \{v : y_v \geq t\}$ satisfies:

$$\frac{|E(S_t, V \setminus S_t)|}{d|S_t|} \leq \sqrt{2R_N(\boldsymbol{y})}$$

Part 2: $\Phi(G) \leq \sqrt{2\lambda_2}$

Lemma:

This yields a cut with conductance at most $\sqrt{2\min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}} \leq \sqrt{2R_N(\boldsymbol{x})}$.

- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

Proof of the lemma:

- $\boldsymbol{x}' = \boldsymbol{x} - m\mathbf{1}$
  - m is the median of $\boldsymbol{x}$.
- $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$
  - $x_v^+ = \begin{cases} x_v' & \text{if} \quad x_v' > 0 \\ 0 & \text{otherwise} \end{cases}$
  - $x_v^- = \begin{cases} -x_v' & \text{if} \quad x_v' < 0 \\ 0 & \text{otherwise} \end{cases}$

At least half of the vertices have zero values.

Setting $\boldsymbol{y}$ to be $\boldsymbol{x}^+$ or $\boldsymbol{x}^-$

Claim 1:
$R_N(\boldsymbol{x}) \geq R_N(\boldsymbol{x}') \geq \min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}$

Claim 2:
Let $\boldsymbol{y} \in \mathbb{R}^n_{\geq 0}$. There exists a threshold $t > 0$ such that $S_t = \{v : y_v \geq t\}$ satisfies:

$$\frac{|E(S_t, V \setminus S_t)|}{d|S_t|} \leq \sqrt{2R_N(\boldsymbol{y})}$$

$$\Phi(S_t) = \frac{|E(S_t, V \setminus S_t)|}{d\min\{|S_t|, |V \setminus S_t|\}} = \frac{|E(S_t, V \setminus S_t)|}{d|S_t|} \leq \sqrt{2R_N(\boldsymbol{y})}$$

Proof of Claim 1

Lemma:
- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_{\boldsymbol{N}}(\boldsymbol{x})}$.

Claim 1:
$R_{\boldsymbol{N}}(\boldsymbol{x}) \geq R_{\boldsymbol{N}}(\boldsymbol{x}') \geq \min\{R_{\boldsymbol{N}}(\boldsymbol{x}^{+}), R_{\boldsymbol{N}}(\boldsymbol{x}^{-})\}$

- $\boldsymbol{x}' = \boldsymbol{x} - m\mathbf{1}$
  - m is the median of $\boldsymbol{x}$.

$$R_{\boldsymbol{N}}(\boldsymbol{x}') = \frac{(\boldsymbol{x} - m\mathbf{1})^{\top}\boldsymbol{N}(\boldsymbol{x} - m\mathbf{1})}{\|\boldsymbol{x} - m\mathbf{1}\|^2} = \frac{\boldsymbol{x}^{\top}\boldsymbol{N}\boldsymbol{x}}{\|\boldsymbol{x} - m\mathbf{1}\|^2} = \frac{\boldsymbol{x}^{\top}\boldsymbol{N}\boldsymbol{x}}{\|\boldsymbol{x}\|^2 + m^2} \leq \frac{\boldsymbol{x}^{\top}\boldsymbol{N}\boldsymbol{x}}{\|\boldsymbol{x}\|^2} = R_{\boldsymbol{N}}(\boldsymbol{x})$$

$\mathbf{1}$ is an eigenvector of $\boldsymbol{N}$ with a zero eigenvalue.

$\boldsymbol{x} \perp \mathbf{1}$

Proof of Claim 1

Lemma:
- Let $\boldsymbol{x}$ be any vector orthogonal to $\mathbf{1}$.
- There is a polynomial-time algorithm that computes a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$.

- $\boldsymbol{x}' = \boldsymbol{x}^+ - \boldsymbol{x}^-$
  - $x_v^+ = \begin{cases} x_v' & \text{if } \; x_v' > 0 \\ 0 & \text{otherwise} \end{cases}$
  - $x_v^- = \begin{cases} -x_v' & \text{if } \; x_v' < 0 \\ 0 & \text{otherwise} \end{cases}$

Claim 1:
$R_N(\boldsymbol{x}) \geq R_N(\boldsymbol{x}') \geq \min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}$

$$R_N(\boldsymbol{x}') = \frac{\sum_{\{u,v\}\in E}(x_u - x_v)^2}{d\|\boldsymbol{x}'\|^2} \geq \frac{\sum_{\{u,v\}\in E}\left((x_u^+ - x_v^+)^2 + (x_u^- - x_v^-)^2\right)}{d(\|\boldsymbol{x}^+\|^2 + \|\boldsymbol{x}^-\|^2)} \geq \min\{R_N(\boldsymbol{x}^+), R_N(\boldsymbol{x}^-)\}$$

$\|\boldsymbol{x}'\|^2 = \|\boldsymbol{x}^+\|^2 + \|\boldsymbol{x}^-\|^2$
$(x_u - x_v)^2 \geq (x_u^+ - x_v^+)^2 + (x_u^- - x_v^-)^2$

Proof of Claim 2

- Assume $\max_v y_v = 1$. Multiplication by a scalar does not affect the Rayleigh quotient of a vector.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

We will prove:

$$\frac{\mathbb{E}[|E(S_t, V \setminus S_t)|]}{d\mathbb{E}[|S_t|]} \leq \sqrt{2R_N(\boldsymbol{y})}$$

Claim 2:
Let $\boldsymbol{y} \in \mathbb{R}^n_{\geq 0}$. There exists a threshold $t > 0$ such that $S_t = \{v : y_v \geq t\}$ satisfies:

$$\frac{|E(S_t, V \setminus S_t)|}{d|S_t|} \leq \sqrt{2R_N(\boldsymbol{y})}$$

Sub-claim 1: $d\mathbb{E}[|S_t|] = d\sum_v y_v^2$

Proof of Claim 2

Sub-claim 2: $\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sqrt{2\left(\sum_{\{u,v\}\in E}(y_u - y_v)^2\right)(d\sum_v y_v^2)}$

- Assume $\max_v y_v = 1$.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

$$\frac{\mathbb{E}[|E(S_t, V \setminus S_t)|]}{d\mathbb{E}[|S_t|]} \leq \sqrt{2 \cdot \frac{\sum_{\{u,v\}\in E}(y_u - y_v)^2}{d\sum_{v\in V} {y_v}^2}} = \sqrt{2R_N(\boldsymbol{y})}$$

We will prove:

$$\frac{\mathbb{E}[|E(S_t, V \setminus S_t)|]}{d\mathbb{E}[|S_t|]} \leq \sqrt{2R_N(\boldsymbol{y})}$$

Sub-claim 1: $d\mathbb{E}[|S_t|] = d\sum_v y_v^2$

Proof of Claim 2

Sub-claim 2: $\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sqrt{2\left(\sum_{\{u,v\}\in E}(y_u - y_v)^2\right)(d\sum_v y_v^2)}$

- Assume $\max_v y_v = 1$.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

Proof of Sub-claim 1:
For each vertex v, it is added to S_t when $t^2 \in [0, y_v^2]$,
which happens with probability y_v^2.

Linearity of expectations:

$$\mathbb{E}[|S_t|] = \sum_v y_v^2$$

Sub-claim 1: $d\mathbb{E}[|S_t|] = d\sum_v y_v^2$

Proof of Claim 2

Sub-claim 2: $\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sqrt{2\left(\sum_{\{u,v\}\in E}(y_u - y_v)^2\right)(d\sum_v y_v^2)}$

- Assume $\max_v y_v = 1$.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

Proof of Sub-claim 2:

For each edge $e = \{u, v\}$, it belongs to $E(S_t, V \setminus S_t)$ when

$$t^2 \in (\min\{y_u^2, y_v^2\}, \max\{y_u^2, y_v^2\}],$$

which happens with probability

$$|y_u^2 - y_v^2|.$$

Proof of Claim 2

Sub-claim 1: $d\mathbb{E}[|S_t|] = d\sum_v y_v^2$

Sub-claim 2: $\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sqrt{2\left(\sum_{\{u,v\}\in E}(y_u - y_v)^2\right)\left(d\sum_v y_v^2\right)}$

- Assume $\max_v y_v = 1$.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

Proof of Sub-claim 2:

For each edge $e = \{u, v\}$, it belongs to $E(S_t, V \setminus S_t)$ when

$$t^2 \in (\min\{y_u^2, y_v^2\}, \max\{y_u^2, y_v^2\}],$$

which happens with probability

$$|y_u^2 - y_v^2|.$$

$$\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sum_{\{u,v\}\in E} |y_u^2 - y_v^2|$$

$$= \sum_{\{u,v\}\in E} (y_u + y_v)|y_u - y_v|$$

Cauchy-Schwarz $\leq \sqrt{\sum_{\{u,v\}\in E} (y_u + y_v)^2} \sqrt{\sum_{\{u,v\}\in E} (y_u - y_v)^2}$

Proof of Claim 2

Sub-claim 1: $d\mathbb{E}[|S_t|] = d\sum_v y_v^2$

Sub-claim 2: $\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sqrt{2\left(\sum_{\{u,v\}\in E}(y_u - y_v)^2\right)\left(d\sum_v y_v^2\right)}$

- Assume $\max_v y_v = 1$.
- Select t randomly in such a way that t^2 is uniformly distributed in $[0,1]$.
- Consider $S_t = \{v : y_v \geq t\}$.

Proof of Sub-claim 2:

For each edge $e = \{u, v\}$, it belongs to $E(S_t, V \setminus S_t)$ when

$$t^2 \in (\min\{y_u^2, y_v^2\}, \max\{y_u^2, y_v^2\}],$$

which happens with probability

$$|y_u^2 - y_v^2|.$$

$$\mathbb{E}[|E(S_t, V \setminus S_t)|] = \sum_{\{u,v\}\in E} |y_u^2 - y_v^2|$$

$$= \sum_{\{u,v\}\in E} (y_u + y_v)|y_u - y_v|$$

$$\leq \underbrace{\sqrt{\sum_{\{u,v\}\in E} (y_u + y_v)^2}}_{} \sqrt{\sum_{\{u,v\}\in E} (y_u - y_v)^2}$$

$$\sum_{\{u,v\}\in E} (y_u + y_v)^2 \leq 2 \sum_{\{u,v\}\in E} (y_u^2 + y_v^2) = 2d \sum_v y_v^2$$

Cheeger vs. Leighton–Rao

- Cheeger inequality yields more practical algorithms.

Small and precise constants:

$$\frac{\lambda_2}{2} \le \Phi(G) \le \sqrt{2\lambda_2}$$

Simple and fast algorithm:

After sorting the vertices according to a given vector $\boldsymbol{x}$ orthogonal to $\mathbf{1}$, finding a cut with conductance at most $\sqrt{2R_N(\boldsymbol{x})}$ can be done in linear time.

Application: Image segmentation

A low-conductance cut

$\updownarrow$

A boundary where few high-similarity edges are cut

$\updownarrow$

A natural separation between regions

- Pixels = vertices.
- Edges = similarity between pixels.

Shi, Jianbo, and Jitendra Malik. “Normalized cuts and image segmentation.” *IEEE Transactions on pattern analysis and machine intelligence* (2000).

Application: Image segmentation

A low-conductance cut

A boundary where few high-similarity edges are cut

A natural separation between regions

- Pixels = vertices.
- Edges = similarity between pixels.

- This enables **image segmentation** by grouping similar pixels into meaningful regions.

- Applications:
  - Medical imaging: identifying organs, tumors, or tissue boundaries.
  - Robotics: scene understanding and object detection

Shi, Jianbo, and Jitendra Malik. "Normalized cuts and image segmentation." *IEEE Transactions on pattern analysis and machine intelligence* (2000).

Application: graph partitioning

- **Expander decomposition:**
  - Iteratively finding low-conductance cuts partitions the vertices into well-connected clusters with few inter-cluster edges.

  This tool has many applications in designing algorithms for **general graphs**.

- **Detecting communities in a social network:**
  - Splitting the network into clusters of tightly-connected users.

Extension 1: multi-way cuts

- **Recall:** $\lambda_k = 0$ if and only if G has at least k connected components.

- Similar to Cheeger inequality, there is a robust version of this fact.

Extension 1: multi-way cuts

- **Recall:** $\lambda_k = 0$ if and only if G has at least k connected components.

- Similar to Cheeger inequality, there is a robust version of this fact.

$$\Phi_k(G) = \min_{\text{disjoint } S_1 \subseteq V, S_2 \subseteq V, \ldots, S_k \subseteq V} \max_{1 \le i \le k} \frac{|E(S_i, V \setminus S_i)|}{d|S_i|}$$

$$\frac{\lambda_k}{2} \le \Phi_k(G) \in O(k^2) \cdot \lambda_k$$

https://doi.org/10.1145/2665063

Extension 2: bipartiteness

- **Recall:** $\lambda_n = 2$ if and only if G has a bipartite connected component.

- Similar to Cheeger inequality, there is a robust version of this fact.

Extension 2: bipartiteness

- **Recall:** $\lambda_n = 2$ if and only if G has a bipartite connected component.

- Similar to Cheeger inequality, there is a robust version of this fact.

A graph $G = (V, E)$ is ϵ-close to bipartite if one can make it bipartite by removing at most an ϵ-fraction of its edges.

Theorem: If A graph $G = (V, E)$ is ϵ-close to bipartite, then $\lambda_n \geq 2(1 - \epsilon)$.

Extension 2: bipartiteness

- **Recall:** $\lambda_n = 2$ if and only if G has a bipartite connected component.
- Similar to Cheeger inequality, there is a robust version of this fact.

A graph $G = (V, E)$ is ϵ-close to bipartite if one can make it bipartite by removing at most an ϵ-fraction of its edges.

Theorem: If A graph $G = (V, E)$ is ϵ-close to bipartite, then $\lambda_n \geq 2(1 - \epsilon)$.

Proof:
- Suppose $G = (V, E)$ is ϵ-close to a bipartite graph with the bipartition $V = A \cup B$.
- Let $\boldsymbol{y} = \mathbf{1}_A - \mathbf{1}_B$.

$$2 - \lambda_n = \min_{x \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{d \sum_{v \in V} {x_v}^2} \leq \frac{\sum_{\{u,v\} \in E} (y_u + y_v)^2}{d \sum_{v \in V} {y_v}^2} \leq \frac{4\epsilon |E|}{2|E|} = 2\epsilon$$

Extension 2: bipartiteness

- **Recall:** $\lambda_n = 2$ if and only if G has a bipartite connected component.
- Similar to Cheeger inequality, there is a robust version of this fact.

A graph $G = (V, E)$ is ϵ-close to bipartite if one can make it bipartite by removing at most an ϵ-fraction of its edges.

Theorem: If A graph $G = (V, E)$ is ϵ-close to bipartite, then $\lambda_n \geq 2(1 - \epsilon)$.

Moreover, using approximate eigenvector computation, a set $D \subseteq E$ with $|D| \in O(\sqrt{\epsilon}) \cdot |E|$ such that $G' = (V, E \setminus D)$ is bipartite can be computed in polynomial time.

Outlook

- Given the eigenvector $\boldsymbol{v}_2$, we can obtain a cut of conductance at most $2\sqrt{\Phi(G)}$ in polynomial time.

- **Next:** The power method.
  - A fast algorithm for finding approximate eigenvectors.

For any constant $\epsilon > 0$, this allows us to obtain a cut of conductance at most $(1 + \epsilon) \cdot 2\sqrt{\Phi(G)}$ in polynomial time.

References

- **Main reference:**
  - Lecture 4.1 of https://sites.google.com/site/thsaranurak/teaching/Expander
  - Chapters 4 and 5 of https://lucatrevisan.github.io/books/expanders-2016.pdf

- **Additional/optional reading:** Extensions of Cheeger inequality
  - Chapters 6–8 of https://lucatrevisan.github.io/books/expanders-2016.pdf