Advanced Graph Algorithms and Optimization	Spring 2023
Spectral Graph Theory	
R. Kyng & M. Probst	*Problem Set 3 — Wednesday, March 8*

These exercises will not count toward your grade, but you are encouraged to solve them all. This exercise sheet contains exercises relating to lectures in Weeks 3. We encourage you to start the exercises early so you have time to get through everything.

To get feedback, you must hand in your solutions by 23:59 on March 16. Both hand-written and LaTeX solutions are acceptable, but we will only attempt to read legible text.

Exercise 1

Let P_n be the path from vertex 1 to n and $G_{1,n}$ be the graph with only the edge between vertex 1 and n. Furthermore, assume that the edge between vertex i and $i+1$ has positive weight w_i for $1 \leq i \leq n-1$. Prove that

$$G_{1,n} \preceq \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) \sum_{i=1}^{n-1} w_i G_{i,i+1}.$$

Solution

Note that this inequality is a weighted version of the path inequality you have seen in class. We are going to apply the Cauchy-Schwarz inequality in a similar fashion as in the unweighted case. Let $\boldsymbol{x} \in \mathbb{R}^n$ be an arbitrary vector and define $\boldsymbol{\Delta}(i) = \boldsymbol{x}(i+1) - \boldsymbol{x}(i)$. We set $\boldsymbol{\gamma}(i) = \boldsymbol{\Delta}(i)\sqrt{w_i}$ and let $\boldsymbol{w}^{-\frac{1}{2}}$ denote the vector for which $\boldsymbol{w}^{-\frac{1}{2}}(i) = 1/\sqrt{w_i}$. Then, we have that

$$\sum_{i=1}^{n-1} \boldsymbol{\Delta}(i) = \boldsymbol{\gamma}^\top \boldsymbol{w}^{-\frac{1}{2}}, \quad \left\|\boldsymbol{w}^{-\frac{1}{2}}\right\|_2^2 = \sum_{i=1}^{n-1} \frac{1}{w_i}, \quad \text{and} \quad \|\boldsymbol{\gamma}\|_2^2 = \sum_{i=1}^{n-1} \boldsymbol{\Delta}(i)^2 w_i.$$

Therefore, we get

$$\begin{aligned}
\boldsymbol{x}^\top \boldsymbol{L}_{G_{1,n}} \boldsymbol{x} &= \left(\sum_{i=1}^{n-1} \boldsymbol{\Delta}(i)\right)^2 \\
&= \left(\boldsymbol{\gamma}^\top \boldsymbol{w}^{-\frac{1}{2}}\right)^2 \\
&\leq \left(\|\boldsymbol{\gamma}\|_2 \cdot \left\|\boldsymbol{w}^{-\frac{1}{2}}\right\|_2\right)^2 \\
&= \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) \sum_{i=1}^{n-1} \boldsymbol{\Delta}(i)^2 w_i \\
&= \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) \boldsymbol{x}^\top \left(\sum_{i=1}^{n-1} w_i \boldsymbol{L}_{G_{i,i+1}}\right) \boldsymbol{x}.
\end{aligned}$$

where we used the Cauchy-Schwarz inequality. Now, we can conclude that

$$G_{1,n} \preceq \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) \sum_{i=1}^{n-1} w_i G_{i,i+1}.$$

Exercise 2

In Chapter 4, we proved that

$$\lambda_2(T_d) \geq \frac{1}{(n-1)\log_2 n} \ .$$

Improve this bound to $\lambda_2(T_d) \geq 1/cn$ for some constant $c > 0$.

Hint: Use the result of previous exercise.

Solution

Let $w : E \to \mathbb{R}^+$ be a positive function where E is the edge set of T_d. We define $T_d^{i,j}$ to be the unique path between two vertices i and j in T_d. Furthermore, for any edge $e \in E$, we let G_e be the graph with n vertices and only edge e being present. Using Part A, we have that

$$\begin{aligned} K_n &= \sum_{i<j} G_{i,j} \\ &\preceq \sum_{i<j} \left(\left(\sum_{e \in T_d^{i,j}} \frac{1}{w(e)} \right) \left(\sum_{e \in T_d^{i,j}} w(e) G_e \right) \right) \\ &\preceq \left(\max_{i<j} \sum_{e \in T_d^{i,j}} \frac{1}{w(e)} \right) \left(\sum_{i<j} \sum_{e \in T_d^{i,j}} w(e) G_e \right) \end{aligned}$$

Assume that the root is in level 0 and the leaves are in level d. Then, for each edge e between level i and $i+1$ for $i \in \{0, 1, \ldots, d-1\}$, we define $w(e) = 2^i$. Therefore, we get

$$\max_{i<j} \sum_{e \in T_d^{i,j}} \frac{1}{w(e)} = 2 \sum_{i=0}^{d-1} 2^{-i} \leq 4.$$

For an edge e, let p_e be the number of paths $T_d^{i,j}$ which include edge e. Assume that e is between the i-th level and the $(i+1)$-th level, since e is a cut edge we have

$$p_e = \left(2^{d-i} - 1\right) \left(\left(2^{d+1} - 1\right) - \left(2^{d-i} - 1\right) \right) \leq 2^{d-i} \cdot 2^{d+1} = 2^{2d+1} \cdot 2^{-i}.$$

Therefore, we have that

$$\begin{aligned} K_n &\preceq 4 \sum_{i<j} \sum_{e \in T_d^{i,j}} w(e) G_e \\ &= 4 \sum_{e \in E} w(e) \cdot p_e \cdot G_e \\ &\preceq 4 \sum_{e \in E} 2^i \cdot 2^{2d+1} \cdot 2^{-i} \cdot G_e \\ &\preceq 4 \sum_{e \in E} n^2 \cdot G_e \\ &= 4n^2 T_d. \end{aligned}$$

Hence, we conclude that

$$\lambda_2(T_d) \geq \frac{1}{4n^2} \lambda_2(K_n) = \frac{1}{4n}.$$

Exercise 3

Find the conductance $\phi \in (0, 1]$ for the following graphs:

1. the complete graph K_n over n vertices.
2. the path graph P_n over n vertices.

Solution

Recall that

$$\phi(G) = \min_{\emptyset \subset S \subset V} \frac{|E(S, V \setminus S)|}{\min\{vol(S), vol(V \setminus S)\}}.$$

We assume n to be an even number, the case where n is odd is similar.

1. For any bipartition $(S, V \setminus S)$ of K_n with $k = |S| \leq n/2$, we have

   $$\phi(S) = \frac{k(n-k)}{k(n-1)} = \frac{n-k}{n-1}.$$

   Note that this ratio becomes smaller the larger k (for $k \leq n/2$), so the minimum conductance is achieved on any set with S of size $n/2$, which gives

   $$\phi(K_n) = \frac{n}{2(n-1)}.$$

2. Since ϕ is smaller with less cut edges, we thus use the bipartition of P_n with the form of $S = \{1, 2, \ldots, k\}$ and $V \setminus S = \{k+1, \ldots, n\}$, where we only have one cut edge, i.e., $|E(S, V \setminus S)| = 1$. Again, it is not hard to see that $\phi(G)$ is achieved when $|S| = n/2$ where

   $$\phi(S) = \frac{1}{1 + 2(n/2 - 1)} = \frac{1}{n-1} = \phi(P_n).$$

Exercise 4

Show that $\lambda_2(\boldsymbol{L}) \neq 0$ if and only if G is connected. Argue that the same applies for $\boldsymbol{N}$.

Solution

- "$\Rightarrow$"
  It is equivalent to prove that if G is disconnected, then $\lambda_2(\boldsymbol{L}) = 0$. We only prove in the case of G being an unweighted graph. Note that it is also true for weighted graphs. Since G is disconnected, denote $\{C_1, \cdots, C_k\}, k \geq 2$ as the connnected components of G. To prove $\lambda_2(\boldsymbol{L}) = 0$, we need to show there exists a vector $\boldsymbol{y} \perp \mathbf{1}$ s.t. $\boldsymbol{L}\boldsymbol{y} = \mathbf{0}$. We choose $\boldsymbol{y} = \mathbf{1}_{C_i} - \alpha\mathbf{1}$ as a test vector for the second eigenvalue of $\boldsymbol{L}$, where $\mathbf{1}_{C_i}$ is a vector with entries corresponding to vertices in $\mathcal{C}_i$ being 1 and the rest entries being 0, and α is some scalar. We can compute α to make $\boldsymbol{y} \perp \mathbf{1}$:

  $$(\mathbf{1}_{C_i} - \alpha\mathbf{1})^\top\mathbf{1} = \mathbf{1}_{C_i}^\top\mathbf{1} - \alpha\mathbf{1}^\top\mathbf{1} = |C_i| - \alpha|V| = 0,$$

  thus, $\alpha = |C_i|/|V|$. Also, we have,

  $$\begin{aligned} \boldsymbol{y}^\top\boldsymbol{L}\boldsymbol{y} =& (\mathbf{1}_{C_i} - \alpha\mathbf{1})^\top\boldsymbol{L}(\mathbf{1}_{C_i} - \alpha\mathbf{1}) \\ &= \mathbf{1}_{C_i}^\top\boldsymbol{L}\mathbf{1}_{C_i} && \boldsymbol{L}\mathbf{1} = \mathbf{0} \\ &= \sum_{(u,v)\in E}(\mathbf{1}_{C_i}(u) - \mathbf{1}_{C_i}(v))^2 \\ &= 0 && u \text{ and } v \text{ belong to the same connected component} \end{aligned}$$

  According to the Courant-Fischer Theorem, we have

  $$\lambda_2(\boldsymbol{L}) = \min_{\boldsymbol{x}\perp\mathbf{1},\boldsymbol{x}\neq\mathbf{0}} \frac{\boldsymbol{x}^\top\boldsymbol{L}\boldsymbol{x}}{\boldsymbol{x}^\top\boldsymbol{x}} \leq \frac{\boldsymbol{y}^\top\boldsymbol{L}\boldsymbol{y}}{\boldsymbol{y}^\top\boldsymbol{y}} = 0.$$

  Also knowing that $\boldsymbol{L}$ is PSD, we can conclude that $\lambda_2(\boldsymbol{L}) = 0$.

- "$\Leftarrow$"
  We prove by contradiction. We assume that G is connected and $\lambda_2(\boldsymbol{L}) = 0$. Then, we know there exists $\boldsymbol{y} \perp \mathbf{1}$ s.t. $\boldsymbol{L}\boldsymbol{y} = \mathbf{0}$, thus, $\boldsymbol{y}^\top\boldsymbol{L}\boldsymbol{y} = 0$. Since

  $$\boldsymbol{y}^\top\boldsymbol{L}\boldsymbol{y} = \sum_{(u,v)\in E}(\boldsymbol{y}(u) - \boldsymbol{y}(v))^2 = 0,$$

  we must have $\boldsymbol{y} = \alpha\mathbf{1}$ since G is connected, which is contradiactory to our assumption that $\boldsymbol{y} \perp \mathbf{1}$.

To argue that the same applies for $\boldsymbol{N}$, similar proof strategies can be used. Just noticing that the eigenvector corresponding to $\nu_1 = 0$ is $\psi_1 = \boldsymbol{D}^{\frac{1}{2}}\mathbf{1}$. (Details will be in the board notes for exercise session 4).

Exercise 5

A quite related concept to conductance is *sparsity*: we define the sparsity of a cut $\emptyset \subset S \subset V$ by

$$\sigma(S) = \frac{|E(S, V \setminus S)|}{\min\{|S|, |V \setminus S|\}}.$$

An alternative version of Cheeger's inequality relates the second eigenvalue of $\boldsymbol{L}$ (not $\boldsymbol{N}$) to the sparsity of the graph $\sigma(G) = \min_{\emptyset \subset S \subset V} \sigma(S)$:

$$\frac{\lambda_2(\boldsymbol{L})}{2} \leq \sigma(G) \leq \sqrt{2 d_{max} \cdot \lambda_2(\boldsymbol{L})}$$

where d_{max} is the maximum degree of any vertex in the graph.

Prove the lower bound on $\sigma(G)$, i.e. that $\frac{\lambda_2(\boldsymbol{L})}{2} \leq \sigma(G)$.

Hint: Follow closely the proof of of the lower bound in Cheeger's inequality and try to understand what has to be adapted.

Solution

Observe that we can write

$$\sigma(G) = \min_{\emptyset \subset S \subset V, |S| \leq |V|/2} \frac{\boldsymbol{1}_S^\top \boldsymbol{L} \boldsymbol{1}_S}{\boldsymbol{1}_S^\top \boldsymbol{1}_S}.$$

Meanwhile, according to Courant-Fischer, we have

$$\lambda_2(\boldsymbol{L}) = \min_{\boldsymbol{x} \perp \boldsymbol{1}, \boldsymbol{x} \neq 0} \frac{\boldsymbol{x}^\top \boldsymbol{L} \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}}.$$

If for each $\boldsymbol{1}_S$, we can find a vector $\boldsymbol{y}_S \perp \boldsymbol{1}$ s.t.

$$\lambda_2(\boldsymbol{L}) \leq \frac{\boldsymbol{y}_S^\top \boldsymbol{L} \boldsymbol{y}_S}{\boldsymbol{y}_S^\top \boldsymbol{y}_S} \leq 2 \frac{\boldsymbol{1}_S^\top \boldsymbol{L} \boldsymbol{1}_S}{\boldsymbol{1}_S^\top \boldsymbol{1}_S},$$

then we are done.

We can choose $\boldsymbol{y}_S = \boldsymbol{1}_S - \alpha \cdot \boldsymbol{1}$ with $\alpha = |S|/|V|$, thus $\boldsymbol{y}_S \perp \boldsymbol{1}$. Then we compare the value of numerator and denominator separately.

1. For the numerator, we have

   $$\boldsymbol{y}_S^\top \boldsymbol{L} \boldsymbol{y}_S = (\boldsymbol{1}_S - \alpha \boldsymbol{1})^\top \boldsymbol{L} (\boldsymbol{1}_S - \alpha \boldsymbol{1}) = \boldsymbol{1}_S^\top \boldsymbol{L} \boldsymbol{1}_S,$$

   since we translate by the kernal of $\boldsymbol{L}$.

2. For the denominator, we have

$$\begin{aligned}
\boldsymbol{y}_S^\top \boldsymbol{y}_S &= (1-\alpha)^2|S| + \alpha^2|V \setminus S| \\
&= (1-2\alpha+\alpha^2)|S| + \alpha^2(|V|-|S|) \\
&= |S| - 2\alpha|S| + \alpha^2|V| \\
&= |S| - 2\cdot|S|^2/|V| + |S|^2/|V| \\
&= |S| - |S|^2/|V| \\
&\geq \frac{1}{2}\boldsymbol{1}_S^\top \boldsymbol{1}_S && |S| \leq \frac{|V|}{2}
\end{aligned}$$

Combining these two, we can obtain that for any S s.t. $\emptyset \subset S \subset V, |S| \leq |V|/2$, we have $\frac{\boldsymbol{y}_S^\top \boldsymbol{L}\boldsymbol{y}_S}{\boldsymbol{y}_S^\top \boldsymbol{y}_S} \leq 2\frac{\boldsymbol{1}_S^\top \boldsymbol{L}\boldsymbol{1}_S}{\boldsymbol{1}_S^\top \boldsymbol{1}_S}$, which completes our proof.

Exercise 6

In the lecture, we skipped various steps in the proof of Cheeger's inequality. Show that

1. $\boldsymbol{N}$ is symmetric and in fact PSD. We recommend to prove this by proving the following stronger statement: for any matrix $\boldsymbol{A}$ that is PSD, and any matrix $\boldsymbol{C}$, we have that $\boldsymbol{C}^\top \boldsymbol{A}\boldsymbol{C}$ is PSD.

2. Show that the normalization of $\boldsymbol{z}$ in the upper bound proof of Cheeger's inequality can only make the ratio we are interested in smaller. I.e. prove that $\frac{\boldsymbol{z}^\top \boldsymbol{L}\boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z}} \geq \frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}}$.

   Hint: Argue first about the transformation of $\boldsymbol{z}$ to $\boldsymbol{z}_c$. One way of relating their denominator is by minimizing over all choices of $\boldsymbol{z}_c$ for c. For $\boldsymbol{z}_c$ and $\boldsymbol{z}_{sc}$ you should be able to prove an equality.

3. We have also skipped proving that τ is sampled according to a valid probability distribution: Show that $\int_\tau \mathbb{P}[\tau = \ell]\, d\tau = 1$.

   Hint: Recall the properties of $\boldsymbol{z}_{sc}$.

4. Show that

   $$\mathbb{E}_\tau[|E(S_\tau, V \setminus S_\tau)|] \leq \sum_{\{i,j\}\in E} |\boldsymbol{z}_{sc}(i) - \boldsymbol{z}_{sc}(j)| \cdot (|\boldsymbol{z}_{sc}(i)| + |\boldsymbol{z}_{sc}(j)|)$$

   by concluding the argument in the proof.

5. Standard Probabilistic Method: Consider a random variable X with a discrete distribution and let Ω be the sample space. Argue that there exists an $\omega \in \Omega$ with $X(\omega) \geq \mathbb{E}[X]$.

   Hint: Recall the definition of expectation of a discrete random variable.

6. Using the probabilisitic method for Cheeger's Inequality: recall that in our proof, we want to argue that $\frac{\mathbb{E}_\tau\left[\boldsymbol{1}_S^\top \boldsymbol{L}\boldsymbol{1}_S\right]}{\mathbb{E}_\tau\left[\boldsymbol{1}_S^\top \boldsymbol{D}\boldsymbol{1}_S\right]} \leq \sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}}}$ implies that there exists an S with $\frac{\boldsymbol{1}_S^\top \boldsymbol{L}\boldsymbol{1}_S}{\boldsymbol{1}_S^\top \boldsymbol{D}\boldsymbol{1}_S} \leq \sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}}}$. There are two ways to prove this (feel free to choose just one):

(a) you can prove this claim by considering $\mathbb{E}_\tau\left[\mathbf{1}_S^\top \boldsymbol{L}\mathbf{1}_S\right] \leq \sqrt{2 \cdot \frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}}} \cdot \mathbb{E}_\tau\left[\mathbf{1}_S^\top \boldsymbol{D}\mathbf{1}_S\right]$. Use only linearity of expectation to obtain an expression with a single $\mathbb{E}_\tau$ and apply the probabilistic method, or

(b) you can prove that for any two discrete random variables $X, Y > 0$ with the same distribution, we have that there exists an $\omega \in \Omega$ with

$$\frac{X(\omega)}{Y(\omega)} \leq \frac{\mathbb{E}[X]}{\mathbb{E}[Y]}.$$

Solution

1. For any $\boldsymbol{x} \neq \mathbf{0}$, we have

$$\begin{aligned}
\boldsymbol{x}^\top(\boldsymbol{C}^\top \boldsymbol{A}\boldsymbol{C})\boldsymbol{x} =& (\boldsymbol{C}\boldsymbol{x})^\top \boldsymbol{A}(\boldsymbol{C}\boldsymbol{x}) \\
=& \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{y} && \text{denote } \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} \\
\geq& 0 && \boldsymbol{A} \text{ is PSD}
\end{aligned}$$

which shows that, for any matric $\boldsymbol{C}$, $\boldsymbol{C}^\top \boldsymbol{A}\boldsymbol{C}$ is PSD if $\boldsymbol{A}$ is PSD. Therefore, $\boldsymbol{N}$ is PSD since $\boldsymbol{N} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{L}\boldsymbol{D}^{-\frac{1}{2}}$ and $\boldsymbol{L}$ is PSD. $\boldsymbol{N}$ is also symmetric since $\boldsymbol{N}^\top = (\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{L}\boldsymbol{D}^{-\frac{1}{2}})^\top = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{L}\boldsymbol{D}^{-\frac{1}{2}} = \boldsymbol{N}$ since $\boldsymbol{L}$ is symmetric and $\boldsymbol{D}^{-\frac{1}{2}}$ is diagonal.

2. According to the lecture, $\boldsymbol{z}_{sc}$ is $\boldsymbol{z}$ after centering and scaling operation (renumbering is w.l.o.g.). We can thus express

$$\boldsymbol{z}_{sc} = \beta(\boldsymbol{z} - \alpha\mathbf{1}), \text{ where } \boldsymbol{z} \perp \boldsymbol{d}.$$

To prove $\frac{\boldsymbol{z}^\top \boldsymbol{L}\boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z}} \geq \frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}}$, we can also compare the value of numerator and denominator separately.

- For the numerator, we have

$$\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc} = \beta^2(\boldsymbol{z} - \alpha\mathbf{1})^\top \boldsymbol{L}(\boldsymbol{z} - \alpha\mathbf{1}) = \beta^2 \boldsymbol{z}^\top \boldsymbol{L}\boldsymbol{z}.$$

- For the denominator, we have

$$\begin{aligned}
\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc} &= \beta^2(\boldsymbol{z} - \alpha\mathbf{1})^\top \boldsymbol{D}(\boldsymbol{z} - \alpha\mathbf{1}) \\
&= \beta^2(\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z} - \alpha\mathbf{1}^\top \boldsymbol{D}\boldsymbol{z} - \alpha\boldsymbol{z}^\top \boldsymbol{D}\mathbf{1} + \alpha^2\mathbf{1}^\top \boldsymbol{D}\mathbf{1}) \\
&= \beta^2(\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z} - 2\alpha\boldsymbol{z}^\top \boldsymbol{D}\mathbf{1} + \alpha^2\mathbf{1}^\top \boldsymbol{D}\mathbf{1}) && \mathbf{1}^\top \boldsymbol{D}\boldsymbol{z} = (\mathbf{1}^\top \boldsymbol{D}\boldsymbol{z})^\top \text{ since it is a scalar} \\
&= \beta^2(\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z} - 2\alpha\boldsymbol{z}^\top \boldsymbol{d} + \alpha^2\mathbf{1}^\top \boldsymbol{D}\mathbf{1}) && \boldsymbol{D}\mathbf{1} = \boldsymbol{d} \\
&= \beta^2(\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z} + \alpha^2\mathbf{1}^\top \boldsymbol{D}\mathbf{1}) && \boldsymbol{z} \perp \boldsymbol{d} \\
&\geq \beta^2\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z} && \alpha^2\mathbf{1}^\top \boldsymbol{D}\mathbf{1} \geq 0
\end{aligned}$$

Combining these two, we can obtain that $\frac{\boldsymbol{z}_{sc}^\top \boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top \boldsymbol{D}\boldsymbol{z}_{sc}} \leq \frac{\beta^2\boldsymbol{z}^\top \boldsymbol{L}\boldsymbol{z}}{\beta^2\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z}} = \frac{\boldsymbol{z}^\top \boldsymbol{L}\boldsymbol{z}}{\boldsymbol{z}^\top \boldsymbol{D}\boldsymbol{z}}$.

3. Recall the probability density function

$$p(t) = \begin{cases} 2|t| & t \in [\boldsymbol{z}_{sc}(1), \boldsymbol{z}_{sc}(n)] \\ 0 & \text{o.w.} \end{cases}$$

Noted that after centering, we have $\boldsymbol{z}_{sc}(1) \leq 0$ and $\boldsymbol{z}_{sc}(n) \geq 0$. Moreover, we ensure $\boldsymbol{z}_{sc}(1)^2 + \boldsymbol{z}_{sc}(n)^2 = 1$ after scaling.

$$\begin{aligned} \int_\tau \mathbb{P}[\tau = \ell] d\tau &= \int_{\boldsymbol{z}_{sc}(1)}^0 -2\tau d\tau + \int_0^{\boldsymbol{z}_{sc}(n)} 2\tau\tau \\ &= -\tau^2 \Big|_{\boldsymbol{z}_{sc}(1)}^0 + \tau^2 \Big|_0^{\boldsymbol{z}_{sc}(n)} \\ &= \boldsymbol{z}_{sc}(1)^2 + \boldsymbol{z}_{sc}(n)^2 \\ &= 1. \end{aligned}$$

4. According to the lecture notes, we have already shown that

$$\mathbb{E}_\tau[|E(S_\tau, V \setminus S_\tau)|] = \sum_{\substack{\{i,j\} \in E \\ \boldsymbol{z}_{sc}(i) \leq \boldsymbol{z}_{sc}(j)}} \operatorname{sgn}(j) \boldsymbol{z}_{sc}(j)^2 - \operatorname{sgn}(i) \boldsymbol{z}_{sc}(i)^2.$$

It remains to prove that for any $\{i, j\} \in E$ s.t. $\boldsymbol{z}_{sc}(i) \leq \boldsymbol{z}_{sc}(j)$, we have

$$\operatorname{sgn}(j) \boldsymbol{z}_{sc}(j)^2 - \operatorname{sgn}(i) \boldsymbol{z}_{sc}(i)^2 \leq |\boldsymbol{z}_{sc}(i) - \boldsymbol{z}_{sc}(j)| \cdot (|\boldsymbol{z}_{sc}(i)| + |\boldsymbol{z}_{sc}(j)|).$$

We distinguish by cases:

- If $\operatorname{sgn}(i) = \operatorname{sgn}(j)$, then

$$\begin{aligned} \operatorname{sgn}(j) \boldsymbol{z}_{sc}(j)^2 - \operatorname{sgn}(i) \boldsymbol{z}_{sc}(i)^2 &= |\boldsymbol{z}_{sc}(j)^2 - \boldsymbol{z}_{sc}(i)^2| \\ &= |(\boldsymbol{z}_{sc}(j) - \boldsymbol{z}_{sc}(i)) \cdot (\boldsymbol{z}_{sc}(j) + \boldsymbol{z}_{sc}(i))| \\ &= |\boldsymbol{z}_{sc}(i) - \boldsymbol{z}_{sc}(j)| \cdot |\boldsymbol{z}_{sc}(i) + \boldsymbol{z}_{sc}(j)| \\ &\leq |\boldsymbol{z}_{sc}(i) - \boldsymbol{z}_{sc}(j)| \cdot (|\boldsymbol{z}_{sc}(i)| + |\boldsymbol{z}_{sc}(j)|). \end{aligned}$$

- If $\operatorname{sgn}(i) \neq \operatorname{sgn}(j)$, which, more specifically, must be the case of $\operatorname{sgn}(i) = -1$ and $\operatorname{sgn}(j) = +1$ since $\boldsymbol{z}_{sc}(i) \leq \boldsymbol{z}_{sc}(j)$, then

$$\begin{aligned} \operatorname{sgn}(j) \boldsymbol{z}_{sc}(j)^2 - \operatorname{sgn}(i) \boldsymbol{z}_{sc}(i)^2 &= \boldsymbol{z}_{sc}(i)^2 + \boldsymbol{z}_{sc}(j)^2 \\ &= (-\boldsymbol{z}_{sc}(i))^2 + \boldsymbol{z}_{sc}(j)^2 \\ &\leq (-\boldsymbol{z}_{sc}(i) + \boldsymbol{z}_{sc}(j))^2 \\ &= (-\boldsymbol{z}_{sc}(i) + \boldsymbol{z}_{sc}(j)) \cdot (|\boldsymbol{z}_{sc}(i)| + |\boldsymbol{z}_{sc}(j)|) \\ &= |\boldsymbol{z}_{sc}(i) - \boldsymbol{z}_{sc}(j)| \cdot (|\boldsymbol{z}_{sc}(i)| + |\boldsymbol{z}_{sc}(j)|). \end{aligned}$$

5. Recall the definition of expectation of a discrete random variable, we have

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega),$$

where $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$.

Denote $X(\omega^*) = \max_{\omega \in \Omega} X(\omega)$, then

$$X(\omega^*) = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega^*) \geq \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) = \mathbb{E}[X].$$

6. (a) Rearranging and using Linearity of Expectation, we obtain

$$\mathbb{E}_\tau\left[\mathbf{1}_S^\top \boldsymbol{L}\mathbf{1}_S\right] \le \mathbb{E}_\tau\left[\sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top\boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top\boldsymbol{D}\boldsymbol{z}_{sc}}}\mathbf{1}_S^\top\boldsymbol{D}\mathbf{1}_S\right],$$

$$\mathbb{E}_\tau\left[\mathbf{1}_S^\top \boldsymbol{L}\mathbf{1}_S - \sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top\boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top\boldsymbol{D}\boldsymbol{z}_{sc}}}\mathbf{1}_S^\top\boldsymbol{D}\mathbf{1}_S\right] \le 0.$$

Now, we can apply the standard probabilistic method, and know there exists an $\boldsymbol{S}^*$ such that

$$\mathbf{1}_{S^*}^\top \boldsymbol{L}\mathbf{1}_{S^*} - \sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top\boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top\boldsymbol{D}\boldsymbol{z}_{sc}}}\mathbf{1}_{S^*}^\top\boldsymbol{D}\mathbf{1}_{S^*} \le 0,$$

thus

$$\frac{\mathbf{1}_{S^*}^\top \boldsymbol{L}\mathbf{1}_{S^*}}{\mathbf{1}_{S^*}^\top \boldsymbol{D}\mathbf{1}_{S^*}} \le \sqrt{2\cdot\frac{\boldsymbol{z}_{sc}^\top\boldsymbol{L}\boldsymbol{z}_{sc}}{\boldsymbol{z}_{sc}^\top\boldsymbol{D}\boldsymbol{z}_{sc}}}.$$

(b) Since X and Y are from the same distribution, according to the definition of expectation, we have

$$\mathbb{E}[X] = \sum_{\omega\in\Omega}\mathbb{P}(\omega)X(\omega),$$

$$\mathbb{E}[Y] = \sum_{\omega\in\Omega}\mathbb{P}(\omega)Y(\omega).$$

Then,

$$\frac{\mathbb{E}[X]}{\mathbb{E}[Y]} = \frac{\sum_{\omega\in\Omega}\mathbb{P}(\omega)X(\omega)}{\sum_{\omega\in\Omega}\mathbb{P}(\omega)Y(\omega)} \ge \min_{\omega\in\Omega}\frac{\mathbb{P}(\omega)X(\omega)}{\mathbb{P}(\omega)Y(\omega)} = \min_{\omega\in\Omega}\frac{X(\omega)}{Y(\omega)} := \frac{X(\omega^*)}{Y(\omega^*)},$$

where the inequality comes from $\min_i a_i/b_i \le \sum_i a_i/\sum_i b_i,\ a_i, b_i \ge 0$.