Lecture 12
Scribe: Anupam
Last revised

Lecture 12

1 The hard side of Cheeger's inequality

We introduced the notions of edge expansion and the spectral gap and proved the easy direction of Cheeger's inequality connecting the edge expansion to the spectral gap. It remains to prove the hard direction of the Cheeger bound,

$$\frac{1-\lambda_2}{2} \leq h(G) \leq \sqrt{2(1-\lambda_2)} \tag{1}$$

Easy direction recap: The left side of Cheeger's inequality was proved by observing that the scaled characteristic vector of a cut $(S, \overline{S})$ is orthogonal to $\vec{1}$ and has Rayleigh quotient at least $1 - 2h(S)$, thus yielding a lower bound on λ_2.

Hard direction overview: The proof of the hard side of the inequality does the reverse, it is a spectral partitioning algorithm that constructs a sparse cut starting from a vector v with high Rayleigh quotient.

1. Sort the coordinates of v in ascending order to obtain the sequence $v_1 \leq v_2 \leq \cdots \leq v_n$.
2. A sweep cut is a cut separating the first k vertices in the sorted order from the remaining $n-k$ vertices. Compute the expansion of all the sweep cuts and output the sweep cut with the minimum edge expansion. Denote the expansion of the sparsest sweep cut by h_s.

For simplicity we will analyze the spectral partitioning algorithm with the input v being the second eigenvector. It is possible to perform a more general analysis, but analyzing the algorithm on v_2 suffices to prove Cheeger's inequality.

Claim 1
If h_s is the output of the spectral partitioning algorithm with input $v = v_2$ then,

$$h_s \leq \sqrt{2(1-\lambda_2)} \tag{2}$$

Remark: The claim shows that h_s approximates $h(G)$ up to a quadratic factor as $h_s \leq \sqrt{2(1-\lambda_2)} = \sqrt{4\frac{(1-\lambda_2)}{2}} \leq 2\sqrt{h(G)}$ by the easy side of Cheeger's inequality.

$$h(G) \leq h_s \leq 2\sqrt{h(G)} \tag{3}$$

Proof: We prove the slightly weaker statement $h_s \leq \sqrt{16(1-\lambda_2)}$, a more careful analysis yields the correct constant.

The coordinates of the second eigenvector v_2 are denoted by e_i, as v_2 is normalized to have unit length we have $\sum e_i^2 = |v_2|^2 = 1$. The proof strategy is to apply one step of the lazy random walk $\frac{M+I}{2}$ to v_2 and show that its squared Euclidean norm decreases by a factor related to h_s,

$$\left|\frac{M+I}{2}v_2\right|^2 = \left(\frac{1+\lambda_2}{2}\right)^2 \leq 1 - \frac{h_s^2}{16} \tag{4}$$

Completing the square on the right hand side of the above inequality we have,

$$\left(\frac{1+\lambda_2}{2}\right)^2 \leq 1 - \frac{h_s^2}{16} \leq \left(1 - \frac{h_s^2}{32}\right)^2$$

Taking square roots and rearranging, we have the statement in the claim,

$$\frac{1+\lambda_2}{2} \leq 1 - \frac{h_s^2}{32} \Rightarrow (1-\lambda_2) \geq \frac{h_s^2}{16} \tag{5}$$

Why does the squared Euclidean norm of v_2 decrease when one step of the lazy random walk $\frac{M+I}{2}$ is applied? The following observation is central for quantifying the reduction in the squared norm.

Observation: If the tuple (x, y) is replaced by the average tuple $((x+y)/2, (x+y)/2)$ the squared Euclidean norm decreases from (x^2+y^2) to $2.(\frac{x+y}{2})^2 = (x+y)^2/2$. The net decrease in the squared Euclidean norm over the averaging operation is $x^2+y^2-(x+y)^2/2 = (x-y)^2/2$.

Bounding the decrease in squared norm: The squared norm of v_2 can be expressed as a sum over the edges using the fact that G is a d regular graph.

$$1 = |v_2|^2 = \sum_{i \in V} e_i^2 = \frac{1}{d} \sum_{(i,j) \in E} (e_i^2 + e_j^2) \tag{6}$$

It is useful to think of the evolution of the lazy random walk matrix $\frac{M+I}{2}v_2$ from step 0 to step 1 as a composition of the following operations: (i) At step 0 replace the e_i at the vertex by tuples (e_i, e_j) over edges as in equation (6). (ii) At step 1/2 average across the edges replacing the tuple (e_i, e_j) by the average tuple $(\frac{e_i+e_j}{2}, \frac{e_i+e_j}{2})$. (iii) At step 1 average across vertices replacing the tuples for edges incident at vertex i by the average $\frac{\sum_{j\sim i} e_i+e_j}{2d}$.

The drop Δ in the squared norm for step (ii) of averaging across the edges can be quantified using the main observation,

$$\Delta = \frac{1}{d} \sum_{i \sim j} \frac{(e_i - e_j)^2}{2} \tag{7}$$

It suffices to consider the drop Δ to establish Cheeger's bound, all we need about step (iii) is that the squared norm does not increase if we average across vertices. Averaging always decreases the squared norm as the sum of the squares of d numbers having a fixed sum is minimum when all the numbers are equal. The squared norm of $\frac{M+I}{2}v_2$ is therefore less than $1 - \Delta$,

$$\left|\frac{M+I}{2}v_2\right|^2 \leq 1 - \frac{1}{2d} \sum_{i \sim j} (e_i - e_j)^2 \tag{8}$$

The remaining part of the proof is to show that $\Delta = \frac{1}{2d} \sum_{(i,j)\in E}(e_i - e_j)^2 \geq \frac{h_s^2}{16}$.

First Attempt: In order to write Δ as a sum over sweep cuts we could write $(e_i - e_j)^2$ as the telescoping sum $((e_i - e_{i+1}) + \cdots + (e_{j-1} - e_j))^2 \geq \sum_{i \leq k < j}(e_k - e_{k+1})^2$. This approach yields the bound $\Delta \geq \sum E(k, V \setminus k)(e_k - e_{k+1})^2$ in terms of the number of edges crossing the k-th sweep cut. However, the telescoping sum bound is not tight and loses a factor of the length of the edge $|j - i|$, losing a factor of n for long edges. There are several long edges crossing the sweep cuts, the number of edges of length at most k can be at most kd due to d regularity.

We need a bound that is not sensitive to the subdivision of edges across sweep cuts, the following somewhat 'magical' application of the Cauchy Schwarz inequality $(\sum a_i^2).(\sum b_i)^2 \geq (\sum a_i b_i)^2$ achieves this,

$$\begin{aligned} \frac{1}{2d} \sum_{(i,j)\in E} (e_i - e_j)^2 . \frac{\sum_{(i,j)\in E}(e_i + e_j)^2}{\sum_{(i,j)\in E}(e_i + e_j)^2} &\geq \frac{1}{4d^2} \left(\sum_{(i,j)\in E} (e_i - e_j)(e_i + e_j) \right)^2 \\ &= \left(\frac{\sum_{(i,j)\in E} e_i^2 - e_j^2}{2d} \right)^2 \end{aligned} \tag{9}$$

The denominator was bounded as follows, $\sum_{(i,j)}(e_i + e_j)^2 \leq \sum_{(i,j)} 2e_i^2 + 2e_j^2 = 2d \sum_i e_i^2 = 2d$ using the d regularity of the graph. The right hand side of (9) can now be expressed in terms of the sweep cuts as follows,

$$\begin{aligned} \left(\frac{\sum_{(i,j)\in E} e_i^2 - e_j^2}{2d} \right)^2 &\geq \left(\frac{\sum_{k\in[n/2]} dkh_s(e_k^2 - e_{k+1}^2)}{2d} \right)^2 \\ &= \frac{1}{4} \left(\sum_{k\in[n/2]} kh_s e_k^2 - (k-1)h_s e_k^2 \right) \\ &= \frac{1}{4} \left(\sum_{k\in[n/2]} h_s e_k^2 \right)^2 \geq \frac{h_s^2}{16} \end{aligned} \tag{10}$$

Remark: The definition of edge expansion applies to the smaller side of the cut do for all $k \leq n/2$ there must be at least kdh_s edges crossing the k-th sweep cut. Wlog we assume that $\sum_{i\in[n/2]} e_k^2$ is greater than 1/2, else we can change signs and work with $-v_2$. □