Lecture 20 — March 20, 2013

Prof. Nick Harvey *Scribe: Alexandre Fréchette*

We study the MAXIMUM CUT PROBLEM with approximation in mind, and naturally we provide a spectral graph theoretic approach.

1 The Maximum Cut Problem

Definition 1.1. Given a (undirected) graph $G = (V, E)$, the *maximum cut* $\delta(U)$ for $U \subseteq V$ is the cut with maximal value $|\delta(U)|$. The MAXIMUM CUT PROBLEM consists of finding a maximal cut.

We let $\mathsf{MaxCut}(G) = \max\{|\delta(U)| : U \subseteq V\}$ be the value of the maximum cut in G, and $\mathsf{MaxCut}' = \frac{\mathsf{MaxCut}(G)}{|E|}$ be the normalized version (note that both normalized and unnormalized maximum cut values are induced by the same set of nodes — so we can interchange them in our search for an actual maximum cut).

The MAXIMUM CUT PROBLEM is **NP**-hard. For a long time, the randomized algorithm consisting of uniformly choosing a cut was state-of-the-art with its $\frac{1}{2}$-approximation factor[1] The simplicity of the algorithm and our inability to find a better solution were unsettling. Linear programming unfortunately didn't seem to help. However, a (novel) technique called semi-definite programming provided a 0.878-approximation algorithm [Goemans-Williamson '94], which is optimal assuming the Unique Games Conjecture.

2 A Spectral Approximation Algorithm

Today, we study a result of [Trevison '09] giving a spectral approach to the MAXIMUM CUT PROBLEM with a 0.531-approximation factor. We will present a slightly simplified result with 0.5292-approximation factor.

To get the general idea, consider d-regular graphs. As we discussed earlier, the smallest non-zero eigenvalue of the Laplacian determines the spectral expansion. But what does the *maximum* eigenvalue tell us? It is known that $\lambda_{\max}(\mathbf{L}_G) = 2d$ if and only if G is bipartite, that is the maximum cut of G contains every edge! Hence, the idea is that maybe $\lambda_{\max}(\mathbf{L}_G) \approx 2d$ implies G has a big cut.

In more detail, we would like to show that

$$\mathsf{MaxCut}'(G) \leq \frac{1}{2}\frac{\lambda_{\max}(\mathbf{L}_G)}{d}$$

and if $\lambda_{\max}(\mathbf{L}_G) = 2d(1-\varepsilon)$ for some $0 \leq \varepsilon \leq 1$, then we can find a cut that cuts $1 - O(\sqrt{\varepsilon})$ fraction of the edges. To do so, we proceed in three distinct steps:

[1] Let U^* induce the optimal max cut and U be chosen uniformly at random. Then, $\mathbb{E}\{\delta(U)\} = \sum_{e\in E}\mathbb{P}\{e \in \delta(U)\} \geq \sum_{e\in\delta(U^*)}\mathbb{P}\{e \in \delta(U)\} = \sum_{e\in\delta(U^*)}\frac{1}{2} = \frac{1}{2}\delta(U^*) = \frac{1}{2}\mathsf{MaxCut}(G)$.

- formalizing the relationship between $\mathsf{MaxCut}'(G)$ and the eigenvalues of L_G,
- show how to partition a graph, cutting lots of edges, given an eigenvector for the maximum eigenvalue, and
- assemble an algorithm from our observations.

2.1 Graph Spectrum and Maximum Cuts

Fix a (undirected) graph $G = (V, E)$, and as usual assume $V = [n]$. Let $\mathbf{A}$ be the adjacency matrix of G, and $\mathbf{D}$ the diagonal matrix with the degree of the vertices on the diagonal. Recall that the Laplacian is $\mathbf{L}_G = \mathbf{D} - \mathbf{A}$. Today it will be more convenient to work with the matrix $\mathbf{S}_G = \mathbf{D} + \mathbf{A}$, which Gary Miller calls the "sum Laplacian". The reason that $\mathbf{S}_G$ is more convenient is that, in the d-regular case, the quantity $1 - \lambda_{\max}(\mathbf{L}_G)/2d$ is simply $\lambda_{\min}(\mathbf{S}_G)/2d$.

Claim 2.1. If $\frac{|\delta(U)|}{|E|} = 1 - \frac{\varepsilon}{2}$ for some $\varepsilon > 0$, then if $\mathbf{x} = \mathbf{1} - 2\chi(U)$,

$$\mathbf{x}^\top \mathbf{S}_G \mathbf{x} = \varepsilon \mathbf{x}^\top \mathbf{D} \mathbf{x}.$$

Proof. First, note that $\mathbf{x}_i = \begin{cases} -1 & \text{if } i \in U \\ 1 & \text{if } i \notin U \end{cases}$. Hence,

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = 2\Big(|\{e \notin \delta(U)\}| - |\{e \in \delta(U)\}|\Big) = 2\Big(\frac{\varepsilon}{2} - (1 - \frac{\varepsilon}{2})\Big)|E| = 2(\varepsilon - 1)|E| \tag{2.1}$$

and

$$\mathbf{x}^\top \mathbf{D} \mathbf{x} = \sum_{i \in V} \deg(i) = 2|E|. \tag{2.2}$$

So

$$\begin{aligned} \mathbf{x}^\top \mathbf{S}_G \mathbf{x} &= \mathbf{x}^\top (\mathbf{D} + \mathbf{A}) \mathbf{x} \\ &= \mathbf{x}^\top \mathbf{D} \mathbf{x} + \mathbf{x}^\top \mathbf{A} \mathbf{x} \\ &= 2|E| + 2(\varepsilon - 1)|E| \end{aligned}$$

by Equations (2.1) and (2.2),

$$\begin{aligned} &= \varepsilon 2|E| \\ \mathbf{x}^\top \mathbf{S}_G \mathbf{x} &= \varepsilon \mathbf{x}^\top \mathbf{D} \mathbf{x}. \end{aligned}$$

□

Corollary 2.2. If $\min_{\mathbf{x} \neq \mathbf{0}} \{\frac{\mathbf{x}^\top (\mathbf{D}+\mathbf{A})\mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}}\} \geq \varepsilon$ for some $\varepsilon > 0$, then $\mathsf{MaxCut}'(G) \leq 1 - \frac{\varepsilon}{2}$.

Proof. This is just the contrapositive of Claim 2.1. Indeed, let $\min_{\mathbf{x} \neq \mathbf{0}} \{\frac{\mathbf{x}^\top (\mathbf{D}+\mathbf{A})\mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}}\} \geq \varepsilon$, and suppose there was a cut with value larger than $1 - \frac{\varepsilon}{2}$. Then by Claim 2.1, the minimum of $\frac{\mathbf{x}^\top (\mathbf{D}+\mathbf{A})\mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}}$ would in fact be smaller than ε; contradiction. □

This sums up our formalized relationship between the maximum cut and the eigenvalue/quadratic form for an algebraic representation of G.

2.2 Building Cuts from Eigenvectors

We want to devise a way to cut our graph using eigenvectors. First, fix a partitioning of the vertex set of our graph in three sets L (left), R (right) and D (deferred) — see Figure 1.

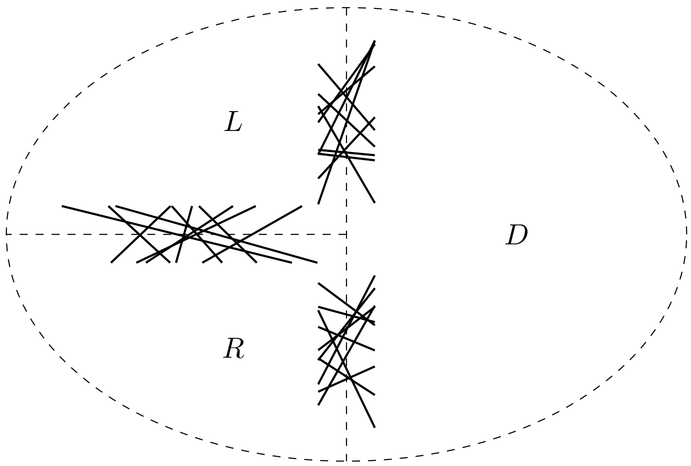

Figure 1: Partitioning a graph in three pieces L, R and D. Only the edges crossing the partitions are shown.

Given two disjoint sets of vertices $S, T \subseteq V$, we let $\delta(S, T) = \{e = ij \in E : i \in S, j \in T \text{ or } i \in T, j \in S\}$. Recall also that, for a set of vertices $S \subseteq V$, $E[S] = \{e = ij \in E : i \in S, j \in S\}$.

Fix our partitioning L,R,D of the vertex set. Then we categorize the edges as follows (see Figure 2):

- $\mathsf{cut} = \delta(L, R)$ are the edges we definitely cut,
- $\mathsf{uncut} = E[L] \cup E[R]$ are edges we do not cut,
- $\mathsf{cross} = \delta(L, D) \cup \delta(R, D)$ are edges we may or may not cut,
- $\mathsf{defer} = E[D]$ are edges we recursively evaluate to check if we cut or not.

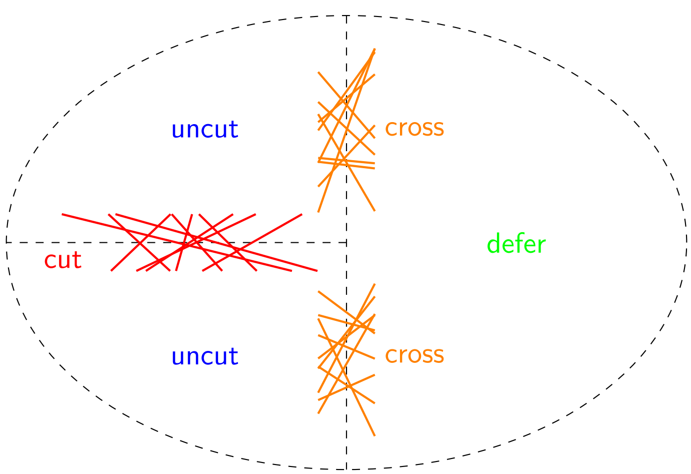

Figure 2: The categorization of the edges given an L,R,D partition.

The idea is that we will identify L and R, and then recursively find (L', R') in the remainder D. We then output $(L \cup L', R \cup R')$ or $(L \cup R', R \cup L')$ as our cut in the graph, whichever is better. Note that one of these choices cuts at least half of the cross edges, and all the cut edges. Hence, we want $|\mathsf{cut}|$ and $|\mathsf{cross}|$ to be large.

Let's see how we can identify L and R.

Lemma 2.3. Given a vector $\mathbf{x}$ such that $\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}} \leq \varepsilon$ for some $\varepsilon > 0$, we can find a vector $\mathbf{y} \in \{-1, 0, 1\}^V$ such that

$$\frac{\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{y}_i + \mathbf{y}_j|}{\sum_{i \in V} \mathbf{D}_{ii} |\mathbf{y}_i|} \leq \sqrt{8\varepsilon}.$$

Now suppose we set $L = \{i : \mathbf{y}_i = 1\}$, $R = \{i : \mathbf{y}_i = -1\}$ and $D = \{i : \mathbf{y}_i = 0\}$. Then in the numerator $\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{y}_i + \mathbf{y}_j|$ of Lemma 2.3's result, a cut edge does not contribute, an uncut edge contributes 4 and a cross edge contributes 2. So $\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{y}_i + \mathbf{y}_j| = 4|\mathsf{uncut}| + 2|\mathsf{cross}|$. As for the denominator, we get the sum of the degrees in $L \cup R$, which is $2(|\mathsf{cut}| + |\mathsf{uncut}|) + |\mathsf{cross}|$. Hence, Lemma 2.3 tells us that

$$\frac{4|\mathsf{uncut}| + 2|\mathsf{cross}|}{2|\mathsf{cut}| + 2|\mathsf{uncut}| + |\mathsf{cross}|} \leq \sqrt{8\varepsilon}. \tag{2.3}$$

But for the left hand side ratio to be that small (i.e. less than $\sqrt{8\varepsilon}$), we need $|\mathsf{cut}|$ to be large, which is to our advantage. Lemma 2.3 is exactly the rounding mechanism we need for our algorithm.

Claim 2.4. Fix $\varepsilon > 0$ and some $\mathbf{x} \in \mathbb{R}^V$. Then

$$\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}} \leq \varepsilon$$

if and only if

$$\frac{1}{2} \sum_{i,j \in V} \mathbf{A}_{ij} (\mathbf{x}_i + \mathbf{x}_j)^2 \leq \varepsilon \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2 = \frac{\varepsilon \sum_{i,j \in V} \mathbf{A}_{ij} (\mathbf{x}_i^2 + \mathbf{x}_j^2)}{2}.$$

Proof. Fix an edge $e = ij$, and let $\mathbf{z}_{ij} = \mathbf{e}_i + \mathbf{e}_j$ (where $\mathbf{e}_i \in \mathbb{R}^V$ is the i-th standard basis), and $\mathbf{Z}_e = \mathbf{z}_{ij} \mathbf{z}_{ij}^\top$. Then, as with the graph Laplacian, we get that $\mathbf{S}_G = \sum_{e=ij \in E} \mathbf{Z}_e$. Now note that $\mathbf{x}^\top \mathbf{Z}_e \mathbf{x} = \mathbf{x}_i^2 + 2\mathbf{x}_i \mathbf{x}_j + \mathbf{x}_j^2 = (\mathbf{x}_i + \mathbf{x}_j)^2$. Hence,

$$\mathbf{x}^\top \mathbf{S}_G \mathbf{x} = \mathbf{x}^\top \Big(\sum_{e=ij \in E} \mathbf{Z}_e \Big) \mathbf{x} = \sum_{e=ij \in E} \mathbf{x}^\top \mathbf{Z}_e \mathbf{x} = \sum_{e=ij \in E} (\mathbf{x}_i + \mathbf{x}_j)^2 = \frac{1}{2} \sum_{i,j \in V} \mathbf{A}_{ij} (\mathbf{x}_i + \mathbf{x}_j)^2. \tag{2.4}$$

Moreover,

$$\mathbf{x}^\top \mathbf{D} \mathbf{x} = \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2 \tag{2.5}$$

by definition of $\mathbf{D}$.

Hence,

$$\begin{aligned} \frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}} \leq \varepsilon \quad &\Leftrightarrow \quad \mathbf{x}^\top \mathbf{S}_G \mathbf{x} \leq \varepsilon \mathbf{x}^\top \mathbf{D} \mathbf{x} \\ &\Leftrightarrow \quad \frac{1}{2} \sum_{i,j \in V} \mathbf{A}_{ij} (\mathbf{x}_i + \mathbf{x}_j)^2 \leq \varepsilon \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2 \end{aligned}$$

by Equations (2.4) and (2.5). Finally,

$$\sum_{i\in V}\mathbf{D}_{ii}\mathbf{x}_i^2=\frac{1}{2}\sum_{i,j\in V}\mathbf{A}_{ij}(\mathbf{x}_i^2+\mathbf{x}_j^2) \tag{2.6}$$

by double counting — each edge is accounted for by two vertices. □

We are now ready to prove our main lemma.

Proof of Lemma 2.3. Suppose we have a vector $\mathbf{x}$ such that $\frac{\mathbf{x}^\top \mathbf{S}_G\mathbf{x}}{\mathbf{x}^\top \mathbf{D}\mathbf{x}}\leq\varepsilon$ for some $\varepsilon>0$. Note that any scaling of $\mathbf{x}$ does not affect the above inequality, so we can assume that $|\mathbf{x}_i|\leq 1$ for all $i\in V$.

By Claim 2.4, $\mathbf{x}$ satisfies

$$\sum_{i,j\in V}\mathbf{A}_{ij}(\mathbf{x}_i+\mathbf{x}_j)^2\leq 2\varepsilon\sum_{i\in V}\mathbf{D}_{ii}\mathbf{x}_i^2 \tag{2.7}$$

and

$$\sum_{i,j\in V}\mathbf{A}_{ij}(\mathbf{x}_i+\mathbf{x}_j)^2\leq \varepsilon\sum_{i,j\in V}\mathbf{A}_{ij}(\mathbf{x}_i^2+\mathbf{x}_j^2).$$

Let us try to interpret this second inequality. It says that, on average over all edges $e=ij\in E$, $(\mathbf{x}_i+\mathbf{x}_j)^2$ is much smaller than $(\mathbf{x}_i^2+\mathbf{x}_j^2)$. In order for this to happen, this means $\mathbf{x}_i$ and $\mathbf{x}_j$ have opposite signs, and their magnitude are about the same, again on average with respect to all edges.

With this observation in hand, consider laying the vertices on the unit interval on the real line according to their $\mathbf{x}$ value - see Figure 3.

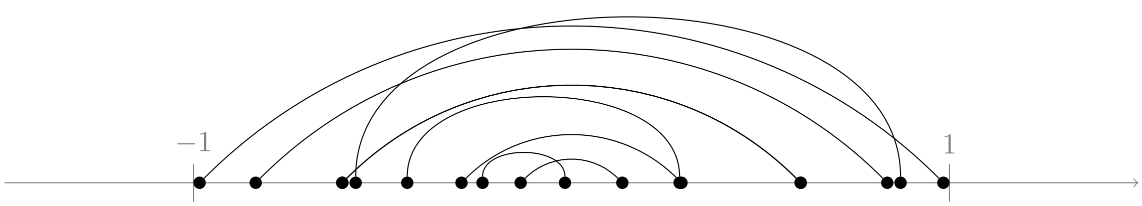

Figure 3: A depiction of vertices on the real line according to their $\mathbf{x}$ coordinate

To partition the graph, we will pick a threshold $t\in[0,1]$ then partition the real line into three parts as follows: $L=\{i:\mathbf{x}_i\leq -t\}$, $R=\{i:\mathbf{x}_i\geq t\}$ and $D=\{i:-t<\mathbf{x}_i<t\}$. See Figure 4.

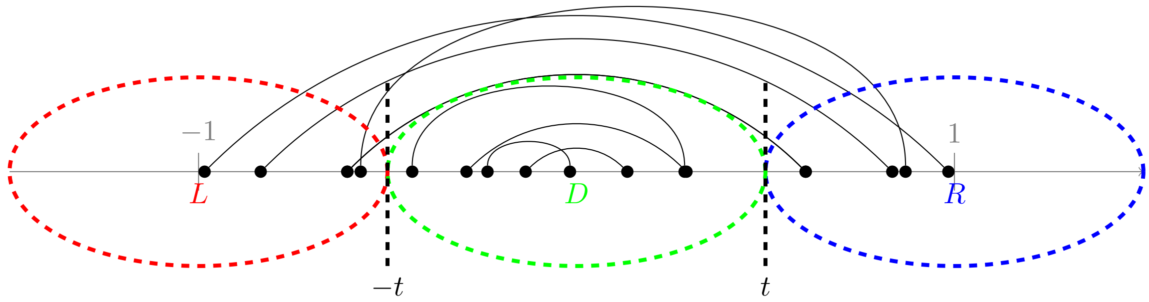

Figure 4: A tri-partition of the real line cutting many edges.

Because of how the $\mathbf{x}_i$'s are (on average) balanced for each edge, we should expect to cut many edges with this L, R, D partition.

But how should we choose t? You might guess that we will choose it randomly. Indeed we will, but the distribution is not what you might expect. We choose $u \sim U[0,1]$ uniformly at random from the unit interval, then set $t = \sqrt{u}$. The L, R, D partition is then encoded in the random vector $\mathbf{Y} \in \{-1, 0, 1\}^V$ defined by

$$\mathbf{Y}_i = \begin{cases} -1 & \text{if } \mathbf{x}_i \leq -t \\ 0 & \text{if } |\mathbf{x}_i| < t \\ 1 & \text{if } \mathbf{x}_i \geq t \end{cases}.$$

We prove that $\mathbf{Y}$ satisfies the conditions of the lemma on expectation, and thus there exists some (non-random) $\mathbf{y}$ satisfying the lemma.

We want to show that

$$\mathbb{E}\{\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{Y}_i + \mathbf{Y}_j|\} \leq \sqrt{8\varepsilon} \cdot \mathbb{E}\{\sum_{i \in V} \mathbf{D}_{ii} |\mathbf{Y}_i|\}. \tag{2.8}$$

We show in Claim 2.5 that

$$\mathbb{E}\{|\mathbf{Y}_i|\} = \mathbf{x}_i^2 \tag{2.9}$$

and

$$\mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} \leq |\mathbf{x}_i + \mathbf{x}_j|(|\mathbf{x}_i| + |\mathbf{x}_j|). \tag{2.10}$$

We use those two properties of $\mathbf{Y}$ to get our conclusion:

$$\begin{aligned} \mathbb{E}\Big\{\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{Y}_i + \mathbf{Y}_j|\Big\} &= \sum_{i,j \in V} \mathbf{A}_{ij} \mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} \\ &\leq \sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{x}_i + \mathbf{x}_j|(|\mathbf{x}_i| + |\mathbf{x}_j|) \end{aligned}$$

by Equation (2.10),

$$\leq \sqrt{\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{x}_i + \mathbf{x}_j|^2} \sqrt{\sum_{i,j \in V} \mathbf{A}_{ij} (|\mathbf{x}_i| + |\mathbf{x}_j|)^2}$$

by the Cauchy-Schwartz inequality. We can then apply our hypothesis Equation (2.7) to the left sum, and use the fact that $(a+b)^2 \leq 2(a^2+b^2)$ for $a, b \in \mathbb{R}$ on the summands of the right sum to get

$$\begin{aligned} &\leq \sqrt{2\varepsilon \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2} \sqrt{2 \sum_{i,j \in V} \mathbf{A}_{ij} (|\mathbf{x}_i|^2 + |\mathbf{x}_j|^2)} \\ &= \sqrt{2\varepsilon \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2} \sqrt{4 \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2} \end{aligned}$$

applying Equation (2.6),

$$= \sqrt{8\varepsilon} \sum_{i \in V} \mathbf{D}_{ii} \mathbf{x}_i^2$$

$$\mathbb{E}\{\sum_{i,j \in V} \mathbf{A}_{ij} |\mathbf{Y}_i + \mathbf{Y}_j|\} \leq \sqrt{8\varepsilon} \sum_{i \in V} \mathbf{D}_{ii} \mathbb{E}\{|\mathbf{Y}_i|\}$$

from Equation (2.9). □

Claim 2.5. Given $\mathbf{x} \in [-1,1]^V$, let $u \sim U[0,1]$ and $\mathbf{Y} \in \{-1,0,1\}^V$ be such that

$$\mathbf{Y}_i = \begin{cases} -1 & \text{if } \mathbf{x}_i \leq -\sqrt{u} \\ 0 & \text{if } |\mathbf{x}_i| < \sqrt{u} \\ 1 & \text{if } \mathbf{x}_i \geq \sqrt{u} \end{cases}.$$

Then

$$\mathbb{E}\{|\mathbf{Y}_i|\} = \mathbf{x}_i^2$$

for all $i \in V$, and

$$\mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} \leq |\mathbf{x}_i + \mathbf{x}_j|(|\mathbf{x}_i| + |\mathbf{x}_j|)$$

for all $i, j \in V$.

Proof. The first result follows directly from definition:

$$\mathbb{E}\{|\mathbf{Y}_i|\} = \mathbb{P}\{|\mathbf{x}_i| \geq \sqrt{u}\} = \mathbb{P}\{|\mathbf{x}_i|^2 \geq u\} = \mathbf{x}_i^2$$

For the second result, we use a case analysis. Without loss of generality, we may assume that $|\mathbf{x}_i| \leq |\mathbf{x}_j|$.

- Suppose that x_i and x_j have opposite signs, or that $x_i = 0$. Then $\mathbf{Y}_i$ and $\mathbf{Y}_j$ cannot have the same sign, so it is not possible that $|\mathbf{Y}_i + \mathbf{Y}_j| = 2$. Thus

  $$\mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} = \mathbb{P}\{|\mathbf{Y}_i + \mathbf{Y}_j| = 1\}.$$

  Because we assume that $|\mathbf{x}_i| \leq |\mathbf{x}_j|$, the event $\{|\mathbf{Y}_i + \mathbf{Y}_j| = 1\}$ occurs precisely when $|\mathbf{x}_i| < \sqrt{u} \leq |\mathbf{x}_j|$. Thus

  $$\mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} \;=\; |\mathbf{x}_j|^2 - |\mathbf{x}_i|^2 \;=\; (|\mathbf{x}_j| - |\mathbf{x}_i|)(|\mathbf{x}_i| + |\mathbf{x}_j|).$$

  Finally, note that $|\mathbf{x}_i + \mathbf{x}_j| = |\mathbf{x}_j| - |\mathbf{x}_i|$ due to our assumptions that $|\mathbf{x}_i| \leq |\mathbf{x}_j|$ and that x_i and x_j have opposite signs (or $x_i = 0$). So the claim is proven in this case.

- Suppose x_i is non-zero and that x_i and x_j have the same sign. Because the claim is invariant under simultaneously negating x_i and x_j, it suffices to consider the case $0 < x_i \le x_j$. By definition of $\mathbf{Y}$, we have

$$|\mathbf{Y}_i + \mathbf{Y}_j| = \begin{cases} 0 & \text{when } x_j < \sqrt{u} \\ 1 & \text{when } x_i < \sqrt{u} \le x_j \\ 2 & \text{when } \sqrt{u} \le x_i \end{cases}$$

Hence,

$$\begin{aligned} \mathbb{E}\{|\mathbf{Y}_i + \mathbf{Y}_j|\} &= \mathbb{P}\{\mathbf{x}_i^2 < u \le \mathbf{x}_j^2\} + 2\mathbb{P}\{u \le \mathbf{x}_i^2\} \\ &= (\mathbf{x}_j^2 - \mathbf{x}_i^2) + 2\mathbf{x}_i^2 \;=\; \mathbf{x}_j^2 + \mathbf{x}_i^2 \\ &\le (\mathbf{x}_i + \mathbf{x}_j)^2 \;=\; |\mathbf{x}_i + \mathbf{x}_j|(|\mathbf{x}_i| + |\mathbf{x}_j|). \end{aligned}$$

So the claim is proven in this case too.

□

2.3 An Approximation Algorithm for the Maximum Cut Problem

As previously noted, given a $\mathbf{x}$ such that $\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D} \mathbf{x}} \le \varepsilon$ for some $\varepsilon > 0$, Lemma 2.3 gives us an L, R, D partition of our graph where the edges satisfy Equation (2.3). Let's work out this last equation to see what we can get out of it:

$$\begin{aligned} &\frac{4|\mathsf{uncut}| + 2|\mathsf{cross}|}{2|\mathsf{cut}| + 2|\mathsf{uncut}| + |\mathsf{cross}|} \le \sqrt{8\varepsilon} \\ \Rightarrow\ &\frac{|\mathsf{uncut}| + \frac{1}{2}|\mathsf{cross}|}{|\mathsf{cut}| + |\mathsf{uncut}| + \frac{1}{2}|\mathsf{cross}|} \le \sqrt{2\varepsilon} \\ \Rightarrow\ &\frac{|\mathsf{cut}|}{|\mathsf{cut}| + |\mathsf{uncut}| + \frac{1}{2}|\mathsf{cross}|} \ge 1 - \sqrt{2\varepsilon} \\ \Rightarrow\ &\frac{|\mathsf{cut}| + \frac{1}{2}|\mathsf{cross}|}{|\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}|} \ge 1 - \sqrt{2\varepsilon}. \end{aligned} \tag{2.11}$$

Note that the numerator $|\mathsf{cut}| + \frac{1}{2}|\mathsf{cross}|$ is the guaranteed minimum value of the best cut between $(L \cup L'), (R \cup R')$ and $(L \cup R'), (R \cup L')$ (where L' and L' are recursively found in D). Indeed, such cut cuts all the cut edges, and at least half the cross edges. Moreover, the denominator $|\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}|$ is the best contribution our two cuts could have, cutting all possible edges (but the deferred edges). Hence, Equation (2.3) tells us that our minimum contribution is not so bad - this is what we leverage.

Consider the following algorithm:

Algorithm 2.1 Spectral Maximum Cut Algorithm

Require: A graph $G = (V, E)$.
Ensure: A cut in G cut of value at least $0.5292\,\mathsf{MaxCut}(G)$.

1: $\tau \leftarrow 0.1107$
2: $\alpha \leftarrow 0.5292$
3: $\varepsilon \leftarrow \min\{\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D}\mathbf{x}} : \mathbf{x} \neq 0\}$
4: $\mathbf{x} \leftarrow \arg\min\{\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D}\mathbf{x}} : \mathbf{x} \neq 0\}$
5: **if** $\varepsilon \geq \tau$ **then**
6:     **return** a random cut in G.
7: **else if** $\varepsilon < \tau$ **then**
8:     Use Lemma 2.3 to get L, R, D partition and corresponding edges cut, uncut and cross.
9:     {Iteratively find best cut in deferred edges.}
10:     $(L', R') \leftarrow$ Spectral Maximum Cut Algorithm$(G[D])$.
11:     {Return the best of two possible cuts.}
12:     **return** $\arg\max\{|\delta(L \cup L', R \cup R')|, |\delta(L \cup R', R \cup L')|\}$
13: **end if**

Before analysing the actual quality of the solution outputted by Algorithm 2.1 (and confirming its guarantee), let us address some algorithmic issues.

First, note that line 8 in the algorithm uses Lemma 2.3 to find a partition, but the lemma is about a random partition. As usual, we can compensate for the randomness by amplifying and repeating the algorithm many times, taking the best cut encountered. The same goes for the random cut outputted at line 6.

Second, how can we efficiently solve the optimization problem $\min\{\frac{\mathbf{x}^\top \mathbf{S}_G \mathbf{x}}{\mathbf{x}^\top \mathbf{D}\mathbf{x}} : \mathbf{x} \neq 0\}$ of lines 3 and 4 in Algorithm 2.1? Well consider the change of variable $\mathbf{z} = \mathbf{D}^{\frac{1}{2}}\mathbf{x}$, which is non-degenerate since $\mathbf{D}$ is diagonal with strictly positive diagonal entries, thus invertible. Then the problem becomes

$$\min\Big\{\frac{\mathbf{z}^\top \mathbf{D}^{-\frac{\top}{2}} \mathbf{S}_G \mathbf{D}^{-\frac{1}{2}} \mathbf{z}}{\mathbf{z}^\top \mathbf{z}} : \mathbf{z} \neq 0\Big\}.$$

Since we can always rescale solutions and keep same objective value (given a solution $\mathbf{z}^*$, use $\frac{\mathbf{z}^*}{\sqrt{\mathbf{z}^{*\top}\mathbf{z}^*}}$), we can restrict the feasible space to unit norm vectors (which is equivalent to $\mathbf{z}^\top \mathbf{z} = 1$):

$$\min\big\{\mathbf{z}^\top \mathbf{D}^{-\frac{\top}{2}} \mathbf{S}_G \mathbf{D}^{-\frac{1}{2}} \mathbf{z} : \mathbf{z} \neq 0, \|\mathbf{z}\| = 1\big\}.$$

The solution of the latter problem is simply the smallest eigenvalue (and corresponding eigenvector) of $\mathbf{D}^{-\frac{\top}{2}} \mathbf{S}_G \mathbf{D}^{-\frac{1}{2}}$.

Finally, let's analyze the quality of the cut returned by Algorithm 2.1. In the first case, $\varepsilon \geq \tau$. By Corollary 2.2, this means $\mathsf{MaxCut}(G) \leq (1 - \frac{\varepsilon}{2})|E|$. But, as we have observed in the introduction, a random cut in expectation cuts at least $\frac{1}{2}|E|$ edges. Hence, our cut outputted as an approximation factor of at least

$$\frac{\frac{1}{2}|E|}{(1 - \frac{\varepsilon}{2})|E|} = \frac{1}{2 - \varepsilon} \geq \frac{1}{2 - \tau} \geq 0.5292 = \alpha.$$

For the second case, we have $\varepsilon < \tau$ and our edge sets cut, uncut, cross and defer induced by our tri-partition. Note that

$$\begin{aligned}\mathsf{MaxCut}(G) &\leq \mathsf{MaxCut}((V, \mathsf{cut} \cup \mathsf{uncut} \cup \mathsf{cross})) + \mathsf{MaxCut}((V, \mathsf{defer}))\\ &\leq |\mathsf{cut} \cup \mathsf{uncut} \cup \mathsf{cross}| + \mathsf{MaxCut}((V, \mathsf{defer}))\end{aligned}$$

where recall that for $F \subseteq E$, $\mathsf{MaxCut}((V, F))$ is the value of the maximum cut in the subgraph (V, F) of G - so $\mathsf{MaxCut}((V, \mathsf{defer})$ is the maximum cut in the deferred edges section of the graph. Thus,

$$\mathsf{MaxCut}(G) \leq |\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}| + \mathsf{MaxCut}((V, \mathsf{defer}))$$

Equation (2.11) guarantees that

$$\Rightarrow \frac{|\mathsf{cut}| + \frac{1}{2}|\mathsf{cross}|}{|\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}|} \geq 1 - \sqrt{2\varepsilon} \geq 1 - \sqrt{2\tau}.$$

Hence, we inductively[2] find L', R' in the deferred edges and output the best cut out of $(L \cup L', R \cup R')$ and $(L \cup R', R \cup L')$, whichever is better and cut at least half the cross edges. Hence, our outputted cut as value at least

$$|\mathsf{cut}| + \frac{|\mathsf{cross}|}{2} + \alpha \mathsf{MaxCut}((V, \mathsf{defer}))$$

by induction, where recall $\alpha = 0.5292$. Piecing everything together to get our ratio, we have

$$\begin{aligned}\frac{|\mathsf{cut}| + \frac{|\mathsf{cross}|}{2} + \alpha \mathsf{MaxCut}((V, \mathsf{defer}))}{\mathsf{MaxCut}(G)} &\geq \frac{|\mathsf{cut}| + \frac{|\mathsf{cross}|}{2} + \alpha \mathsf{MaxCut}((V, \mathsf{defer}))}{|\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}| + \mathsf{MaxCut}((V, \mathsf{defer}))}\\ &\geq \min\left\{\frac{|\mathsf{cut}| + \frac{|\mathsf{cross}|}{2}}{|\mathsf{cut}| + |\mathsf{uncut}| + |\mathsf{cross}|}, \frac{\alpha \mathsf{MaxCut}((V, \mathsf{defer}))}{\mathsf{MaxCut}((V, \mathsf{defer}))}\right\}\end{aligned}$$

since $\frac{a+b}{c+d} \geq \min\{\frac{a}{c}, \frac{b}{d}\}$ for $a, b, c, d > 0$,

$$\begin{aligned} &\geq \min\{1 - \sqrt{2\tau}, \alpha\}\\ \frac{|\mathsf{cut}| + \frac{|\mathsf{cross}|}{2} + \alpha \mathsf{MaxCut}((V, \mathsf{defer}))}{\mathsf{MaxCut}(G)} &\geq \alpha = 0.5292.\end{aligned}$$

So our outputted cut in any case has an approximation ratio of at least $\alpha = 0.5292$.

[2]We are not addressing the formal proof by induction of the result.