Low Conductance Cuts

Kent Quanrud

October 28, 2020

1 Conductance

Let $G = (V, E)$ be an undirected graph with m edges and n vertices, and positive edge weights $w : E \to \mathbb{R}_{>0}$. Let $A : \mathbb{R}^V \to \mathbb{R}^V$ be the weighted adjacency, $d = A\mathbb{1} \in \mathbb{R}^V_{>0}$ the weighted degress, and $D = \text{diag}(D)$. Let $L = \sum_{e \in E} w(e) L_e = D - A : \mathbb{R}^V \to \mathbb{R}^V$ be the Laplacian of G.

Recall that the *sparsity* of a cut $\partial(S)$ (where $S \subset V$) was the ratio

$$\Psi(S) = \frac{\sum_{e \in \partial(S)} w(e)}{\min\{|S|, |\bar{S}|\}},$$

where $\bar{S} = V \setminus S$. The sparsity of the graph G is define as the sparsity of the sparsest cut,

$$\Psi(G) = \min_{S \subset V} \Psi(S).$$

The sparsity arises in the dual of multicommodity flow and has many applications in graph algorithms. The sparsest cut is used as a primitive for divide and conquer algorithms in graphs. In a previous discussion we pointed out the connection between the sparsity of a graph and the Laplacian. In particular, for any nonempty set $S \subset V$ with at most $n/2$ vertices, we have

$$\Psi(S) = \frac{\sum_{e \in \partial(S)} w(e)}{|S|} = \frac{\langle \mathbb{1}_S, L\mathbb{1}_S \rangle}{\langle \mathbb{1}_S, \mathbb{1}_S \rangle}, \tag{1}$$

where $\mathbb{1}_S$ is the $\{0, 1\}$-indicator vector for S. Recall also that the Rayleight quotients $\langle x, Lx \rangle / \langle x, x \rangle$ were strongly linked to the eigenvalues of L. Continuing in this direction, we showed that if we let λ denotes the second smallest eigenvalue of L, then

$$\lambda \leq n\Psi(G).$$

Today we will discuss similar and also important notion called *conductance*. For a set of vertices S, the **volume** of S, denoted $\textbf{vol}(S)$, is the sum of degrees of vertices in S:

$$\text{vol}(S) = \sum_{v \in S} d(v).$$

The **conductance** of a set S, denoted $\boldsymbol{\Phi(S)}$, is defined as

$$\Phi(S) = \frac{\sum_{e \in \partial(S)} w(e)}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}}$$

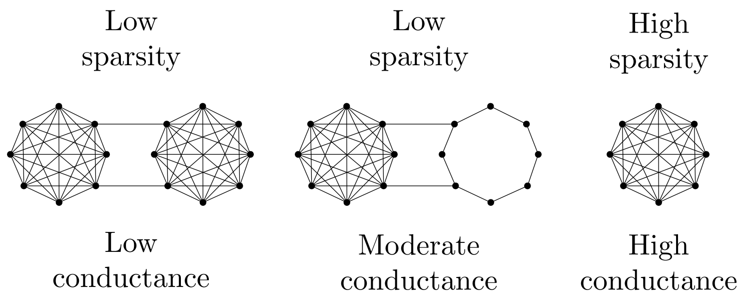


Figure 1: Examples of graphs with varying levels of sparsity and conductance.

Note that $\Phi(S)$ is always positive (for a connected graph) and at most 1. There is a clear resemblance between conductance and sparsity except here the vertices in the denominator are weighted by their degree. Similary to sparsity, we define the conductance of a graph as the minimum conductance of any cut:

$$\Phi(G) = \min_{\emptyset \subsetneq S \subsetneq V} \Phi(S).$$

Like sparsity, conductance is also useful for divide and conquer. The sparsest cut is more suited for divide and conquer on vertices, while conductance, where vertices are weighted by their degree, is more conducive to divide and conquer on edges. Recall that sparsity was naturally motivated by its connection to multicommodity flow. On the other hand, conductance is strongly connected to random walks. Indeed, for any set S, the stationary distribution is in S with probability proportional to $\mathrm{vol}(S)$. To continue this analogy, the conductance of a (small) set S models the amount of probability mass that enters and leaves S in each step at the stationary distribution. Figure 1 gives some examples of graphs with different levels of sparsity and conductance.

We would like to express conductance in algebraic terms, similar to sparsity in (1). While the numerator in (1) seems appropriate, the denominator does not capture the volume. Instead, consider the following quotient:

$$\frac{\langle x, Lx \rangle}{\langle x, Dx \rangle} \text{ where } x \in \mathbb{R}^V. \tag{2}$$

For any set S with at most half the total volume, we have

$$\Phi(S) = \frac{\sum_{e \in \partial(S)} w(e)}{\mathrm{vol}(S)} = \frac{\langle \mathbb{1}_S, L\mathbb{1}_S \rangle}{\langle \mathbb{1}_S, D\mathbb{1}_S \rangle}.$$

That said, the quotient (2) does not have a direct connection to the Laplacian L in the same way as sparsity did. However, it is connected to the *normalized Laplacian*, which is the map $M : \mathbb{R}^V \to \mathbb{R}^V$ defined by

$$M = D^{-1/2} L D^{-1/2}.$$

For any vector x, letting $y = D^{1/2}x$, we have

$$\frac{\langle x, Lx \rangle}{\langle x, Dx \rangle} = \frac{\langle y, My \rangle}{\langle y, y \rangle}.$$

Since the normalized Laplacian M is also symmetric, the RHS models the eigenvalues of M. In today's discussion, we will study the eigenvalues of M and relate it to the condutance of the graph.

We first point out that there are some *similarities* to other matrices that we have studied. Let $R = AD^{-1} : \mathbb{R}^V \to \mathbb{R}^V$ denote the random walk map. Recall that the *normalized random walk matrix* Q was defined as

$$Q = D^{-1/2} R D^{1/2} = D^{-1/2} A D^{-1/2}.$$

To draw the connection to M, if we expand $L = D - A$, then we have

$$M = D^{-1/2}(D - A)D^{-1/2} = I - Q = D^{-1/2}(I - R)D^{1/2}.$$

Theorem 1. *Let $G = (V, E)$ be an undirected graph with m edges and n vertices, and positive edge weights $w : E \to \mathbb{R}_{>0}$. Let $M : \mathbb{R}^V \to \mathbb{R}^V$ be the normalized Laplacian and $R : \mathbb{R}^V \to \mathbb{R}^V$ the random walk matrix. Then M is similar to $I - R$, and (equivalently) $I - M$ is similar to R.*

Recall that similarity preserves eigenvalues. Since R has its eigenvalues in $[-1, 1]$ and 1 with multiplicity 1, M has its eigenvalues in $[0, 2]$ and eigenvalue 0 with multiplicity 1.

The goal today is to prove another theorem about the eigenvalues of M. This one relates the second smallest eigenvalue of M to the conductance of G. The inequality is called Cheeger's inequality due to an analogous bound by Jeff Cheeger for continuous manifolds.

Theorem 2. *Let $G = (V, E)$ be an undirected graph with m edges and n vertices, and positive edge weights $w : E \to \mathbb{R}_{>0}$. Let M be the normalized Laplacian of G. Let λ be the second smallest eigenvalue of M. Then*

$$\frac{\lambda}{2} \leq \Phi(G) \leq \sqrt{2\lambda}.$$

2 Fiedler's algorithm: the upper bound

In this section we present an algorithm proof of the upper bound, $\Phi(G) \leq \lambda$, due to Fiedler [Fie73]. The algorithm is fairly simple, based on previous discussions, the algorithm might be able to guess it.

Recall that

$$\lambda = \min_{y : \langle \sqrt{d}, y \rangle = 0} \frac{\langle y, My \rangle}{\langle y, y \rangle} = \min_{x : \langle d, x \rangle = 0} \frac{\langle x, Lx \rangle}{\langle x, Dx \rangle}.$$

Let $x \in \mathbb{R}^V$ with $\langle d, x \rangle = 0$ attain λ on the RHS. (We note that eigenvectors, thus x, can be computed.) x is orthogonal to d and, assuming that we have normalized such that $\langle x, Dx \rangle = 1$, x has a "fractional cut value" of $\langle x, Lx \rangle = \lambda$. Our goal is to "round" the "fractional cut" $x \in \mathbb{R}^V$ to a set S without loosing too much on the sparsity. How?

As an additional hint, we point out that a similar setup arose before for minimum (s, t)-cut and sparsest cut. In each case we had a "fractional cut" from the LP and wanted to produce a discrete one.

The correct answer is to output the *best cut along the line embedding* x – yet again! This is called *Fiedler's algorithm* and pseudocode is given in Figure 2.

As with (s, t)-cut and sparsest cut before, the analysis is probabilistic. We will show that that a randomized variant returns a satisfying cut with nonzero probability. For an appropriate distribution of random cuts alone the line embedding x, we show that a *random cut* is good in expectation. More precisely, for $t \in \mathbb{R}$, let $S_t = \{v \in V : x_v \geq t\}$. We will find a distribution over

```
Fiedler(G = (V, E))

  1. let x minimize ⟨x,Lx⟩/⟨x,Dx⟩ s.t. ⟨x,d⟩ = 0 and x ≠ 𝟘
  2. for t ∈ ℝ, let S_t = {v ∈ V : x_v ≥ t}
  3. let t ∈ ℝ minimize ⟨S_t,LS_t⟩ / min{∂(S_t),∂(S̄_t)}
  4. return S_t
```

Figure 2: Fiedler's algorithm for low-conductance cuts.

$t \in \mathbb{R}$ such that a random cut S_t has nonzero chance of good sparsity. The distribution will be more involved then the simple distribution for (s,t)-cut, which was simply the uniform distribution.

We define the distribution is defined as follows. First, let $m \in \mathbb{R}^n$ be the "median point" such that

$$\sum_{v: x_v \le m} d_v \ge \frac{1}{2} \text{ and } \sum_{v: x_v \ge m} d_v \ge \frac{1}{2}.$$

Let $t_0 = \min_v x_v$ be the smallest coordinate value in x and let $t_1 = \max_v x_v$ be the largest coordinate value. Scaling x if necessary, we assume that $(m - t_0)^2 + (t_1 - m)^2 = 1$. The distribution over $t \in [t_0, t_1]$ is defined by

$$\mathrm{P}[\alpha \le t \le \beta] = (m - \alpha)^2 + (m - \beta)^2$$

for $t_0 \le \alpha \le m \le \beta \le t_1$.[1] For $t \in \mathbb{R}$, let $S_t = \{v : x_v \le t\}$. We will show that with nonzero probability, $\Phi(\partial(S_t)) \le \sqrt{2\lambda}$.

2.1 High level proof

In this section, we give a high-level proof of correctness of Fiedler's algorithm. In particular, we identify three key lemma's, and then use them to prove the overall theorem. These lemma's are proved in subsequent sections.

The overall structure is similar to our previous discussion on sparsest cut. There we had a ratio of random terms and we analyzed the expected value of the numerator and denominator separately. These were then combined to show that there exists a good cut. Here we start with a lemma addressing the number; i.e., the expected weight of edges cut by S_t.

Lemma 3. $\mathrm{E}\left[\sum_{e \in \partial(S_t)} w(e)\right] \le \sqrt{2\langle x, Lx\rangle \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle}$.

The next lemma addresses the denominator; i.e., the volume on the smaller side of the cut.

Lemma 4. $\mathrm{E}\left[\min\left\{\mathrm{vol}(S_t), \mathrm{vol}\left(\bar{S}_t\right)\right\}\right] = \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle$.

[1] We note that specifying the probability of t being in any given closed interval suffices to describe a continuous distribution $t \in \mathbb{R}$. Note that the rescaling sets $\mathrm{P}[t_0 \le t \le t_1] = 1$

We now have bounds on the expecation of the numerator and denominator of $\Phi(S_t)$. As we showed in our discussion of sparsest cut, this implies there *exists* a cut S_t whose conductance matches these expected values.

Lemma 5. $\mathrm{P}\left[\Phi(\partial(S_t)) \leq \frac{\mathrm{E}\left[\sum_{e\in\partial(S_t)} w(e)\right]}{\mathrm{E}\left[\min\left\{\mathrm{vol}(S_t),\mathrm{vol}\left(\bar{S}_t\right)\right\}\right]}\right] > 0.$

Finally, to relate $\partial(S_t)$ to the eigenvalue, we need to address the fact that $x - m\mathbb{1}$ is not orthogonal to d. But the following lemma shows that the affine shift by $m\mathbb{1}$ only decreases the Rayleigh quotient, as follows. This is simply because translation by $\mathbb{1}$ does not effect the numerator, and can only increase the denominator.

Lemma 6. *Let* $\langle x, d\rangle = 0$. *For any* $\alpha \in \mathbb{R}$,

$$\frac{\langle x + \alpha\mathbb{1}, L(x + \alpha\mathbb{1})\rangle}{\langle x + \alpha\mathbb{1}, D(x + \alpha\mathbb{1})\rangle} \leq \frac{\langle x, Lx\rangle}{\langle x, Dx\rangle}.$$

Proof. We have $\langle x, Lx\rangle = \langle x + \alpha\mathbb{1}, L(x + \alpha\mathbb{1})\rangle$ because $\mathbb{1} \in \ker(L)$. On the other hand, the function $f(\alpha) = \langle x + \alpha\mathbb{1}, D(x + \alpha\mathbb{1})\rangle$ is convex with derivative $2\langle d, x + \alpha\mathbb{1}\rangle = 2\alpha\langle d, \mathbb{1}\rangle$, so it is minimized at $\alpha = 0$. ∎

All put together, we obtain the following.

Theorem 7. *In polynomial time, Fiedler's algorithm computes* S *such that* $\Phi(\partial(S)) \leq 2\sqrt{\Phi(G)}$.

Proof. We have

$$\mathrm{P}\left[\Phi(\partial(S_t)) \leq \sqrt{2\frac{\langle x, Lx\rangle}{\langle x, Dx\rangle}}\right] \overset{\text{(a)}}{\geq} \mathrm{P}\left[\Phi(\partial(S_t)) \leq \frac{\mathrm{E}\left[\sum_{e\in\partial(S_t)} w(e)\right]}{\mathrm{E}\left[\min\left\{\mathrm{vol}(S_t), \mathrm{vol}\left(\bar{S}_t\right)\right\}\right]}\right] \overset{\text{(b)}}{>} 0,$$

as desired. Here (a) is by Lemma 6 and (b) is by Lemma 5. ∎

2.2 Expected size of the cut

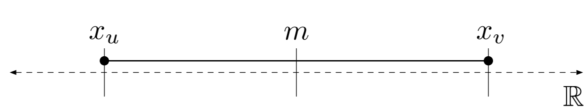
We first analyze the expected size of the cut random. The first probability gives an upper bound on the problem of an edge $e = \{u, v\}$, in terms of the coordinates x_u and x_v.

Lemma 8. *For an edge* $e = \{u, v\} \in E$,

$$\mathrm{P}[e \in \partial(S_t)] \leq |x_u - x_v|(|x_u - m| + |x_v - m|).$$

Proof. We have three different cases, depending on where x_u and x_v lie relative to m. We assume without loss of generality that $x_u \leq x_v$.

1. Suppose $x_u \leq m \leq x_v$.

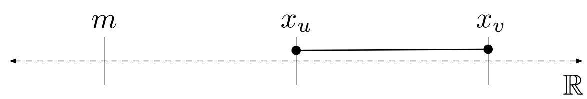


Then

$$\mathrm{P}[e \in \partial(S_t)] = \mathrm{P}[x_u \le t \le x_v] = (m - x_u)^2 + (m - x_v)^2 \overset{(a)}{\le} |x_u - x_v|(|m - x_u| + |m - x_v|).$$

where (a) observes that $|m - x_u| \le |x_u - x_v|$ and $|m - x_v| \le |x_u - x_v|$.

2. Suppose $m \le x_u \le x_v$.

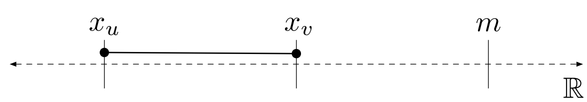


Then

$$\begin{aligned}\mathrm{P}[e \in \partial(S_t)] &= (m - x_v)^2 - (m - x_u)^2 \\ &= (|x_v - m| + |x_u - m|)(|x_v - m| - |x_u - m|) \\ &\overset{(b)}{=} |x_v - x_u|(|x_v - m| + |x_u - m|),\end{aligned}$$

where (b) observes that $|x_v - m| = x_v - m$, $|x_u - m| = x_u - m$, and $x_v - x_u = |x_v - x_u|$.

3. Suppose $x_u \le x_v \le m$.



Then

$$\begin{aligned}\mathrm{P}[e \in \partial(S_t)] &= (m - x_u)^2 - (m - x_v)^2 \\ &= (|x_u - m| + |x_v - m|)(|x_u - m| - |x_v - m|) \\ &\overset{(c)}{=} |x_v - x_u|(|x_v - m| + |x_u - m|),\end{aligned}$$

where (c) observes that $|x_v - m| = m - x_v$, $|x_u - m| = m - x_u$, and $x_v - x_u = |x_v - x_u|$.

■

The next lemma bounds the expected weight of the cut.

Lemma 3. $\mathrm{E}\left[\sum_{e \in \partial(S_t)} w(e)\right] \le \sqrt{2\langle x, Lx\rangle \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle}$.

Proof. We have,

$$\begin{aligned}\mathrm{E}\left[\sum_{e \in \partial(S_t)} w(e)\right] &= \sum_{\{u,v\} \in E} w(e)\mathrm{P}[e \in \partial(S)] \\ &\le \sum_{\{u,v\} \in E} w(e)|x_v - x_u|(|x_v - m| + |x_u - m|) \\ &\overset{(a)}{\le} \sqrt{\sum_{\{u,v\} \in E} w(e)(x_v - x_u)^2}\sqrt{\sum_{\{u,v\} \in E} w(e)(|x_v - m| + |x_u - m|)^2}.\end{aligned}$$

by (a) Cauchy-Schwartz. For the first term, we have

$$\sum_{\{u,v\}\in E} w(e)(x_v - x_u)^2 = \langle x, Lx\rangle.$$

For the second term, we have

$$\begin{aligned}\sum_{\{u,v\}\in E} w(e)(|x_v - m| + |x_u - m|)^2 &\overset{\text{(b)}}{\leq} 2\sum_{\{u,v\}\in E} w(e)\Big((x_v - m)^2 + (x_u - m)^2\Big)\\ &= 2\sum_u \mathrm{vol}(u)(x_u - m)^2\\ &= \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle,\end{aligned}$$

where (b) applies the inequality $2ab \leq a^2 + b^2$. ■

2.3 Expected volume

Lemma 4. $\mathrm{E}\big[\min\{\mathrm{vol}(S_t), \mathrm{vol}(\bar{S}_t)\}\big] = \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle$.

Proof. If $t \geq m$, then $\mathrm{vol}(S_t) \leq \mathrm{vol}(\bar{S}_t)$, and if $t \leq m$, then $\mathrm{vol}(\bar{S}_t) \leq \mathrm{vol}(S_t)$. We have $\mathrm{P}[t \geq m] = (t_1 - m)^2$ and

$$\mathrm{E}[\mathrm{vol}(S_t) \,|\, t \geq m] = \frac{1}{(u - m)^2}\sum_{x_v \geq m} d_v^2 (x_v - m)^2.$$

Similarly, we have $\mathrm{P}[t \leq m] = (m - t_0)^2$ and

$$\mathrm{E}\big[\mathrm{vol}(\bar{S}_t) \,\big|\, t \leq m\big] = \frac{1}{(m - t_0)^2}\sum_{x_v \leq m} d_v^2 (x_v - m)^2.$$

By conditioning on whether t is $\geq$ or $\leq m$, we have

$$\begin{aligned}\mathrm{E}\big[\min\{\mathrm{vol}(S_t), \mathrm{vol}(\bar{S}_t)\}\big] &= \mathrm{E}[\mathrm{vol}(S_t) \,|\, t \geq m]\mathrm{P}[t \geq m] + \mathrm{E}[\mathrm{vol}(S_t) \,|\, t \leq m]\mathrm{P}[t \leq m]\\ &= \sum_{x_v} d_v^2 (x_v - m)^2 = \langle x - m\mathbb{1}, D(x - m\mathbb{1})\rangle,\end{aligned}$$

as desired. ■

3 The lower bound, as an exercise

Exercise 1. Let $S \subseteq V$ induced the minimum conductance cut; i.e., $\mathrm{vol}(S) \leq \mathrm{vol}(V)/2$ and $\Phi(G) = \Phi(S)$. Consider the vector $x = D^{1/2}\mathbb{1}_S$ and let $y = D^{1/2}\mathbb{1}_{\bar{S}}$.

1. Show that

$$\frac{\langle x, Mx\rangle}{\langle x, x\rangle} = \frac{\langle S, LS\rangle}{\langle S, DS\rangle}, \ \frac{\langle y, My\rangle}{\langle y, y\rangle} = \frac{\langle \bar{S}, L\bar{S}\rangle}{\langle \bar{S}, D\bar{S}\rangle}, \text{ and } \langle x, y\rangle = 0.$$

2. Show that for any $\alpha, \beta \neq 0$, we have

$$\langle \alpha x + \beta y, M(\alpha x + \beta y)\rangle \leq 2\Phi(G).$$

3. Argue that one can choose $\alpha, \beta \neq 0$ such that $\langle d, \alpha x + \beta y\rangle = 0$.

4. Finally, prove that the second smallest eigenvector of M is $\leq 2\Phi(G)$.

References

[Fie73] Miroslav Fiedler. "Algebraic connectivity of graphs". In: *Czechoslovak Mathematical Journal* 23.2 (1973), pp. 298–305.

[Tre17] Luca Trevisan. "Lecture notes on graph partitioning, expanders, and spectral methods". Available at `http://www.eecs.berkeley.edu/~luca/books/expanders-2016.pdf`. 2017.

[Spi19] Daniel A. Spielman. "Spectral and Algebraic Graph Theory". Available at `http://cs-www.cs.yale.edu/homes/spielman/sagt`. 2019.