Lecture 20: Graph Cut Sparsifiers

Lecturer: *Huacheng Yu*

1 Cut Sparsifiers

Definition 1. *Let $G = (V, E)$ be an unweighted undirected graph. Then, (possibly weighted) graph $G' = (V, E')$ is an ε-cut sparsifier of G if it preserves approximately all cuts in G. That is, for all $S \subseteq V$,*

$$|E_{G'}(S, \overline{S})| \in (1 \pm \varepsilon)|E_G(S, \overline{S})|$$

Of course, G is a cut sparsifier of itself. We are looking for a sparse G' that has way less edges than G. In fact, you can show you need to have a weighted sparsifier in order to get any nontrivial sparsification. Because of cut-flow correspondence, we can get an approximate max-flow by running a max-flow algorithm on the cut sparsifier. Thus, if we could find a cut sparsifiers, this will speed up our algorithms.

Theorem 1. *For any unweighted G, there is a cut sparsifier G' with $\tilde{O}\left(\frac{n}{\varepsilon^2}\right)$. Moreover, G' is a weighted subgraph of G.*

There are very efficient (near-linear-time) algorithms, but the algorithm we present here will be just polynomial time (so no max-flow speedups yet). The main idea is that we will sample some of the edges at random to keep, and discard the rest. However, uniform sampling will not achieve this goal. For example, consider the graph in Figure 1: made up of two dense subgraphs connected by a single edge.

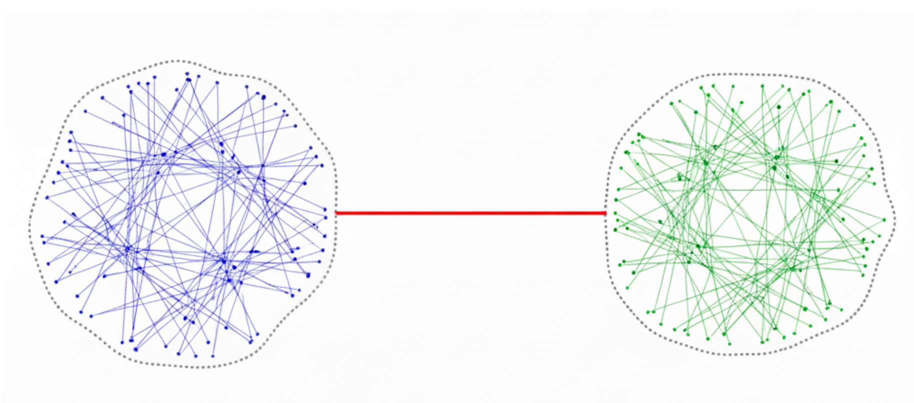

Figure 1: An adversarial case for uniform sampling.

Now, if you remove the "bridge," the cut value between the dense parts will drop from 1 to 0, and so will not be an ε-approximation to the original cuts. Thus, this bridge is somehow more *important* to preserving cuts in the graph, so we should sample it with higher probability!

If there's a small cut in the graph, we will remove the cut (like in the above bad case), then recurse on the remaining part. If this graph has no small cut, then this problem is alleviated and we can just uniformly sample. Here is the exact details.
`SPARSIFY`(G)

1. If G has a cut $(S, \overline{S})$ of size $\leq \frac{c}{\varepsilon^2} \log^3 n$
   (a) We recursively compute sparsifiers G'_S for $G[S]$ and $G'_{\overline{S}}$ for $G[\overline{S}]$.
   (b) We return $G' := G'_S \cup G'_{\overline{S}} \cup E_G(S, \overline{S})$
2. Otherwise:
   (a) Sample every edges of G with probability $1/2$, obtain H.
   (b) Recursively compute a sparsifier H' for H.
   (c) Return $G' = H'$, but with all the edge weights multiplied by 2.

Theorem 2. *G' returned by `SPARSIFY`(G) has $O\left(\frac{n}{\varepsilon^2} \log^3 n\right)$ edges.*

Proof. The edges we explicitly add to G' are the small cuts of size $O\left(\frac{1}{\varepsilon^2} \log^3 n\right)$. By definition of a cut, every time we add such a cut, the graph is partitioned into two nonempty parts and we recurse. However, the graph has n vertices, so this can only happen at most $n-1$ times, giving the required count. □

Theorem 3. *G' returned by `SPARSIFY`(G) is a ε-sparsifier with probability $1 - O\left(n^{-2}\right)$.*

Proof. We will argue each step of the process incurs a small error to the cut size.

1. Suppose G has a small cut $(S, \overline{S})$ (i.e. case 1). Consider any cut $(T, \overline{T})$. Then, we can write

$$\begin{aligned} |E(T, \overline{T})| &= |E(S \cap T, S \cap \overline{T})| + |E(S \cap T, \overline{S} \cap \overline{T})| \\ &\quad + |E(\overline{S} \cap T, S \cap \overline{T})| + |E(\overline{S} \cap T, \overline{S} \cap \overline{T})| \end{aligned}$$

   Note that $(S \cap T, S \cap \overline{T})$ is a cut in $G[S]$ and $(\overline{S} \cap T, \overline{S} \cap \overline{T})$ is a cut in $G[\overline{S}]$. Thus, by recursion, we compute sparsifiers G'_S and $G'_{\overline{S}}$, which approximate $G[S]$ and $G[\overline{S}]$. Therefore, $|E_{G'_S}(S \cap T, S \cap \overline{T})| \in (1 \pm \varepsilon)|E_{G[S]}(S \cap T, S \cap \overline{T})|$ and $|E_{G'_{\overline{S}}}(\overline{S} \cap T, \overline{S} \cap \overline{T})| \in (1 \pm \varepsilon)|E_{G[\overline{S}]}(\overline{S} \cap T, \overline{S} \cap \overline{T})|$.

   For the other two terms, note that $E(S \cap T, \overline{S} \cap \overline{T}), E(\overline{S} \cap T, S \cap \overline{T}) \subseteq E(S, \overline{S})$. Since $E(S, \overline{S})$ is added to the sparsifier at the end, Furthermore, $|E_{G'}(S \cap T, \overline{S} \cap \overline{T})| = |E_G(S \cap T, \overline{S} \cap \overline{T})|$ and likewise for the other set of edges. Thus, the whole sum $|E(T, \overline{T})|$ is preserved to a $(1 \pm \varepsilon)$ factor.

2. Suppose we are in case 2. Then we sample every edge with probability $1/2$ assuming the min-cut size $C \geq \frac{c}{\varepsilon^2} \log^3 n$. Consider a min-cut $(S, \overline{S})$. By a Chernoff bound,

$$\begin{aligned} \Pr\left[\left|E_H\left(S, \overline{S}\right) \mid \notin \frac{1}{2}(1 \pm \eta)|E_G(S, \overline{S})\right|\right] &\leq \exp(-\Omega(\eta^2 |E_G(S, \overline{S})|)) \\ &\leq \exp\left(-\Omega\left(\eta^2 \frac{c}{\varepsilon^2} \log^3 n\right)\right) \end{aligned}$$

Let $\eta = \frac{\varepsilon}{\log n}$. Then, this probability is at most

$$\exp\left(-\Omega\left(\left(\frac{\varepsilon}{\log n}\right)^2 \frac{c}{\varepsilon^2}\log^3 n\right)\right) \leq n^{-\Omega(c)}$$

But, by cut counting (which we proved in HW 1), there are $\leq n^{2B}$ cuts of size at most BC. Thus, for a B-approximate min-cut, the same calculation yields that the cut is does not retain close to 1/2 of its edges with probability at most

$$\exp\left(-\Omega\left(\eta^2 \frac{c}{\varepsilon^2}\log^3 B\right)\right) \leq n^{-\Omega(cB)}$$

Then, by a union bound, if $c \gg 2$, then by setting $B = 1, 2, 4, \ldots, 2^i$ and considering all cuts of size in $(2^i C, 2^{i+1}C]$:

$$\Pr[\exists \text{ a cut whose size is not preserved to } (1 \pm \eta) \text{ factor}] \leq \sum_{i=0}^{\infty} n^{2(2^i)} n^{-\Omega(c(2^i))} \leq \frac{1}{n^2}$$

Thus, by choice of η, each sampling step preserves all cuts within $\left(1 \pm \frac{\varepsilon}{\log n}\right)$ error with high probability. Since there can only be $\log n$ sampling rounds, the total error is within $(1 \pm \varepsilon)$.

$\square$

2 Spectral Sparsifiers

Definition 2. *Consider graph $G = (V, E)$ with $n = |V|$. Let $e = (i, j)$ be an edge. Then we define the edge-Laplacian $L_e \in \mathbb{R}^{n \times n}$.*

$$L_e = \begin{array}{c} \\ i \\ \\ \vdots \\ \\ j \end{array} \begin{array}{c} \begin{array}{ccccc} i & & \ldots & & j \end{array} \\ \begin{pmatrix} 1 & 0 & \ldots & 0 & -1 \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ -1 & 0 & \ldots & 0 & 1 \end{pmatrix} \end{array}$$

Then, the ***Laplacian*** *of G, $L_G = \sum_{e \in G} L_e$. In other words*

$$(L_G)_{ij} = \begin{cases} \deg(i) & i = j \\ -1 & (i,j) \in E \\ 0 & \textit{otherwise} \end{cases}$$

so, we can write $L_G = D - A$, where D is a diagonal matrix of the degrees of each vertex.

Let $x \in \{0,1\}^n$ be the indicator vector of a cut $S \subseteq V$. Then

$$x^\top L_G x = x^\top \sum_{e \in G} L_e x = \sum_{e \in G} x^\top L_e x = \sum_{e=(i,j) \in G} (x_i^2 + x_j^2 - 2x_i x_j) = \sum_{e=(i,j) \in G} (x_i - x_j)^2$$

But notice the thing being summed over is just the indicator that for an edge $x_i \neq x_j$, i.e. edge is cut. Thus,

$$x^\top L_G x = |E_G(S, \overline{S})|$$

Hence, if we want to preserve all cuts, it's enough to preserve this quadratic form for all $0-1$ vectors.

Definition 3. *A graph G' with edge weights w_e is an ε-cut sparsifier of G if, defining the weighted Laplacian $L_{G'} = \sum_{e \in G'} w_e L_e$ and for all $x \in \{0,1\}^n$,*

$$x^\top L_{G'} x \in (1 \pm \varepsilon) x^\top L_G x$$

Definition 4. *A graph G' is a ε-spectral sparsifier of G if the above condition holds for all $x \in \mathbb{R}^n$.*

There's another equivalent condition. Let $y = L_G^{1/2} x$ (you can show the Laplacian is PSD, so it has a square root), so

$$\begin{aligned} \forall x \in \mathbb{R}^n,\ x^\top L_{G'} x \in (1 \pm \varepsilon) x^\top L_G x &\Leftrightarrow \forall y \in \mathbb{R}^n,\ y^\top L_G^{-1/2} L_{G'} L_G^{-1/2} y \in (1 \pm \varepsilon) \|y\|_2^2 \\ &\Leftrightarrow \left\| \sum_{e \in G'} w_e L_G^{-1/2} L_e L_G^{-1/2} - I \right\|_{\mathrm{op}} \leq \varepsilon \end{aligned}$$

Note that if we have $G = G'$, then $\sum_{e \in G} L_G^{-1/2} L_e L_G^{-1/2} = I$. Thus, you can think of it as we have a bunch of terms that sum to I, then we take a small subset of them and reweight them to get something close to I. Then, it turns out a special type of nonuniform sampling works.

`SPECTRAL-SPARSIFY`(G)

1. For every edge e, define $p_e = \|L_G^{-1/2} L_e L_G^{-1/2}\|_{\mathrm{op}}$.
2. For all edges, sample e with probability p_e. If sampled, set $w_e = \frac{1}{p_e}$.
3. Repeat $\tilde{O}\left(\frac{1}{\varepsilon^2}\right)$ times and add up the weights.

To prove that this preserves the largest eigenvalue of the sum (which is also the average of the sparsified sum), we can use a tool from random matrix theory called the matrix Chernoff bound. It directly gives the result.

To explain this strange quantity, there is a physical intuition. If you view the graph as a set of resistors of 1Ω at each edge, then $p_{(i,j)}$ is the effective resistance between i and j. p_e is also the probability that a uniformly random spanning tree T contains the edge e. The second interpretation implies $\sum_e p_e = n - 1$, so we only get expected n edges from each stage.