CS787: Advanced Algorithms	Scribe: Surya Teja Chavali
Lecture 24: Cut Sparsifiers and The Probabilistic Method	Date: April 25, 2019

24.1 Recap: Chernoff Bounds

Theorem 24.1.1 (Chernoff Bound) *Let $X_1, X_2, \ldots X_n$ be n independent random variables such that $X_i \in [0, B]$ with probability 1 for each i. Let $X = \sum_i X_i$ and $\mu = \mathbf{E}[X]$.*

- *For $\epsilon \in (0, 1)$,*
$$\mathbf{Pr}[|X - \mu| \geq \epsilon\mu] \leq \exp\left(-\frac{\epsilon^2\mu}{3B}\right)$$
- *For $\lambda > 0$,*
$$\mathbf{Pr}[|X - \mu| \geq \lambda] \leq \exp\left(-\frac{\lambda^2}{3\mu B}\right)$$

24.2 Cut Sparsifiers

The goal of *cut sparsification* is to take a large undirected graph G on n vertices and return a smaller, weighted graph H on the same vertex set, but with $O(n \operatorname{polylog}(n))$ edges such that the weights of all cuts are preserved.

Formally, given a weighted undirected graph $G(V, E)$, a cut sparsifier of G is defined as a weighted graph $H(V, E')$ such that for all partitions (U_1, U_2) of V,

$$(1 - \epsilon)w_G(U_1, U_2) \leq w_H(U_1, U_2) \leq (1 + \epsilon)w_G(U_1, U_2)$$

24.2.1 A simple example

Let $G = K_n$. Our initial idea is as follows: for each edge $e \in E$, pick e with probability p(which is to be set later), and set its weight to be $\frac{1}{p}$ in H; else delete it.

Consider a cut C of size μ. Let $X = w_H(C)$. $\mathbf{E}[X] = \mu$.

$$\mathbf{Pr}[|X - \mu| \geq \epsilon\mu] \leq \exp\left(-\frac{\epsilon^2\mu}{3}\right) \leq \exp\left(-\frac{\epsilon^2(n-1)}{3}\right)$$

There are exponentially many$\left(2^{n-1}\right)$ cuts in G. A union bound over these many terms would be essentially useless. The key issue is that the upper bound $\exp\left(-\frac{\epsilon^2(n-1)}{3}\right)$ is very loose, as there are only n cuts of size $n - 1$ in G. There are exponentially many cuts of much larger size. The lemma below shows this.

Lemma 24.2.1 (Karger) *If G is an undirected weighted graph on n vertices, with min cut of size k, the number of cuts of size $\leq \alpha k$ is at most $n^{2\alpha}$.*

Using Lemma 24.2.1, we will be able to produce a much tighter union bound. The probability of failure is bounded as below:

$$\sum_{\alpha=1,2,4,\ldots 2^{\log n}} \sum_{\substack{\text{Cuts } C \text{ such that} \\ \frac{\alpha}{2}n \le |C| \le \alpha n}} \mathbf{Pr}[C \text{ is not preserved in } H]$$

$$\le \sum_{\alpha} \sum_{C} \exp\left(-\frac{\epsilon^2 |C|}{3}\right)$$

$$\le \sum_{\alpha} n^{2\alpha} \exp\left(-\frac{\epsilon^2}{3}\frac{\alpha n}{2}\right)$$

Put $p = \frac{18}{\epsilon^2}\frac{\log n}{n}$, so that the term in the exponent becomes $(-3\alpha \log n)$.

Then, the sum becomes

$$\sum_{\alpha} n^{-\alpha} \to 0$$

Further, $\mathbf{E}[|E'|] = p\frac{n^2}{2} = O\left(\frac{n \log n}{\epsilon^2}\right)$

24.2.2 Generalizing the Example

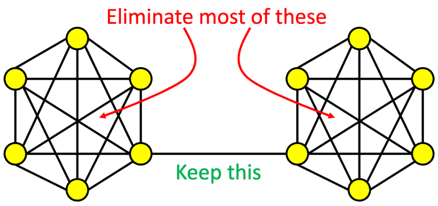

Figure 24.2.1: An example where we cannot drop all edges with the same probability

The above analysis works well in the context of the symmetric K_n. An example where dropping all edges with the same probability is shown above. While the large cuts in the above graph are preserved with good probability, the small cuts are not. Specifically, the bridge between the two cliques is deleted with large probability.

This suggests that edges which belong to small cuts must be preserved with large probability. For $e \in E$, let k_e denote the size of the smallest cut containing e. This leads to the following algorithm, due to Benczur and Karger:

$\forall \quad e \in E$

- Put $w_H(e) = 0$
- Independently ρ times
  - with probability $\frac{1}{k_e}$, add $\frac{k_e}{\rho}$ to $w_H(e)$

Notice that the bridge in the previous pathological example is preserved with probability 1 under this algorithm.

Note: In the above,

- The parameter ρ will be set later.
- In the result for K_n, the probability contained a (poly-)logarithmic factor; this is the role played by ρ. Instead of increasing weight all at once, we are doing this in many steps.
- The quantity k_e can be computed by solving a min-cut(max-flow) problem between the end points of e.

24.2.3 Analyzing the Algorithm

Fact 24.2.2 *For all cuts C, $\mathbf{E}[w_H(C)] = w_G(C)$.*

Fact 24.2.3 *For all cuts C, $\mathbf{E}[|\{e : w_H(e) > 0\}|] \leq n\rho$.*

Proof:

$$\mathbf{E}[|\{e : w_H(e) > 0\}|] = \sum_{e \in E} \frac{\rho}{k_e}$$

Claim 24.2.4 $\sum\limits_{e \in E} \frac{1}{k_e} \leq n$

Proof: By Induction. Consider a graph G. Consider C, the min-cut of G of size k^*, dividing G into G_1 and G_2 of size n_1 and n_2 respectively.

For $e \in G_1$, $\frac{1}{k_e(G_1)} \geq \frac{1}{k_e(G)}$ because the size of the min cut can only increase in a supergraph.

Hence,

$$\begin{aligned}
&\sum_{e \in E} \frac{1}{k_e(G)} \\
&= \sum_{e \in E(G_1)} \frac{1}{k_e(G)} + \sum_{e \in E(G_2)} \frac{1}{k_e(G)} + \sum_{e \in C} \frac{1}{k_e(G)} \\
&\leq \sum_{e \in E(G_1)} \frac{1}{k_e(G_1)} + \sum_{e \in E(G_2)} \frac{1}{k_e(G_2)} + k^* \times \frac{1}{k^*} \\
&\leq n_1 - 1 + n_2 - 1 + 1 = n - 1
\end{aligned}$$

■

The statement follows from the claim above.

■

24.2.3.1 Bounding the probability of failure

In this section, we will analyze the probability of failure. Define

$$E_i = \{e : k_e \in (2^{i-1}, 2^i]\}, \quad i = 0, 1, \ldots 2\log n$$

For each cut C, consider the *projection* $P = |C \cap E_i|$. Note that, for any projection, there can be many cuts which generated it. Our strategy will be to bound the number of projections, and the probability of failure of any cut for a given projection. We will consider the smallest cut with a given projection for this purpose. Define

$$\mathrm{sm}(P) = \min\{|C| : C \cap E_i = P\}$$

We shall use the following lemma, similar in spirit to Karger's Lemma, in the proof below.

Lemma 24.2.5 (unproven) *For E_i as defined,*

$$\left|\{P : sm(P) \le \alpha \cdot 2^i\}\right| \le n^{2\alpha}$$

We now proceed to bound the probabity of failure.

$$\mathbf{Pr}\left[||w_H(P)| - \mathbf{E}[|w_H(P)|]| > \frac{\epsilon \cdot \mathrm{sm}(P)}{2\log n}\right] \le \exp\left(-\frac{1}{3}\frac{\epsilon^2 \cdot (\mathrm{sm}(P))^2}{|P|\left(\frac{2^i}{\rho}\right)}\frac{1}{4\log^2 n}\right)$$

$$\le \exp\left(-\frac{1}{12}\frac{\epsilon^2 \cdot \rho \cdot \mathrm{sm}(P)}{2^i \cdot \log^2 n}\right)$$

$$\mathbf{Pr}[\text{Failure}] \le \sum_{\text{classes } i} \sum_{\text{projections } P} \mathbf{Pr}[P \text{ deviates from its expectation}]$$

$$\le \sum_{i=1}^{2\log n} \sum_{\alpha} \sum_{\{P : \mathrm{sm}(P) \in \left(\frac{\alpha}{2}, \alpha\right) 2^i\}} \mathbf{Pr}[P \text{ deviates from its expectation}]$$

$$\le \sum_i \sum_\alpha n^{2\alpha} \exp\left(-\frac{1}{12}\frac{\epsilon^2 \cdot \rho \cdot 2^{i-1}\alpha}{2^i \cdot \log^2 n}\right)$$

From the lemma above

Put $\rho = \dfrac{72\log^3 n}{\epsilon^2}$, then, each term in the sum becomes $n^{-\alpha}$,

and the sum becomes $O\left(\dfrac{1}{\mathrm{poly}\, n}\right)$

24.3 The Probabilisitic Method

In the next lecture, we will discuss the *Probabilistic Method*, a powerful tool to construct deterministic structures using randomness. In essence, the Probabilistic Method defines an appropriate random process to generate a given structure, and shows that the desired event on that structure happens with a non-zero probability. This causes us to conclude that a realization of that desired event is possible.

Specifically, we will construct an exponentially large set of vectors which are approximately mutually orthogonal.