CHAPTER 7

# Graph sparsification

In the previous lecture, we considered a probabilistic algorithm for finding minimum cuts, based on contracting randomly chosen edges. In today's lecture, we will discuss a method called graph sparsification, which will allow us to replace any graph $G$ by a sparse graph $G'$ in such a way that the sizes of all cuts are (approximately) preserved.

## 1. The number of cuts

In the previous lecture, we discussed Karger's min-cut algorithm, which subsequently contracts randomly chosen edges, until two vertices remain. If $C$ is any min-cut, the probability that it survives this procedure is at least $1/\binom{n}{2}$ by Theorem 6.9. This immediately yields the following result.

COROLLARY 7.1. *The number of min-cuts in a graph on $n$ vertices is at most $\binom{n}{2}$.*

PROOF. Suppose that there are $M$ min-cuts, that are enumerated $C_1, C_2, \ldots, C_M$. The events $\{C_i \text{ survives}\}$ are pairwise disjoint, so from the union bound

$$\frac{M}{\binom{n}{2}} \leq \sum_{i=1}^{M} \mathbb{P}\left(C_i \text{ survives}\right) \leq 1, \tag{7.1}$$

from which the claim follows. □

Note that the $\binom{n}{2}$-bound is tight, as in the cycle $C_n$ any pair of edges defines a min-cut.

We can give a result that is more general than Corollary 7.1. For $\alpha \geq 1$, we define an $\alpha$-cut if its size is at most $\alpha$ times the size of a minimum cut. We can show that if $C$ is an $\alpha$-cut, then the probability that it survives Karger's min-cut procedure is at least

$$\frac{1}{\binom{n}{2\alpha} 2^{2\alpha}}. \tag{7.2}$$

This gives a bound similar to the bound in Corollary 7.1.

COROLLARY 7.2. *The number of $\alpha$-cuts in a graph on $n$ vertices is at most $\binom{n}{2\alpha} 2^{2\alpha}$.*

## 2. Sparsification

In this lecture, we will consider graph sparsification. This technique refers to the approximation of a graph by a (weighted) subgraph that is sparse (i.e. has less than a quadratic number of edges). The idea is that the approximation is similar to the original graph in some way.

Here, we are interested in preserving cuts. Our aim is to find a (sparse) subgraph $G'$ of $G$ that satisfies

$$(1-\varepsilon)|\nabla_G(S, V \setminus S)| \leq |\nabla_{G'}(S, V \setminus S)| \leq (1+\varepsilon)|\nabla_G(S, V \setminus S)| \tag{7.3}$$

for all $S \subseteq V$. If we can do this, than we will be able to compute min-cuts in shorter time (as the subgraph contains much fewer edges).

Note that in general we cannot expect (7.3) to hold. Suppose that we have an algorithm that keeps a $p$-fraction of all edges, than we expect the size of a cut to decrease roughly by a factor $p$ as well. We will therefore consider weighted graphs, i.e. graphs equipped with a function $w : E \to \mathbb{R}$.

For a weighted graph, when we say size of a cut, we refer to the total weight of the edges crossing the cut, so

$$|\nabla_G(S, V \setminus S)| = \sum_{e \in \nabla_G(S,V\setminus S)} w(e). \tag{7.4}$$

Note that the unweighted situtation can be retrieved by setting all edge weights equal to 1.

Roughly, in what follows, we will assign to each edge in the subgraph $G'$ a weight $1/p$, so that (7.3) is satisfied.

Throughout, we will refer to the size of the minimum cut in $G$ as the **edge-connectivity** of $G$, and we will denote it $\kappa(G)$ (or simply $\kappa$).

**2.1. Random sampling.** Let $G$ be an (unweighted) graph, and let $\mathbf{G}'$ be the graph obtained by randomly pruning this graph in the following way: each edge is selected (and kept) with probability $p$, independently of all other edges. Each selected edge gets weight $1/p$.

Let $C$ be any cut (in fact, any edge set) in $G$, and let $\mathbf{C}'$ be the corresponding cut in $\mathbf{G}'$. Then

$$\mathbb{E}\left[|\mathbf{C}'|\right] = \sum_{e \in C} \mathbb{P}\left(e \in C'\right) w(e) = \sum_{e \in C} p \frac{1}{p} = |C|, \tag{7.5}$$

so that (7.3) is satisified at least in expectation.

Now suppose that $G$ has edge-connectivity $\kappa$. Let $C$ be any cut in $G$ (obviously, we have $|C| \geq \kappa$). We will show that if $p$ is sufficiently large, the size of $\mathbf{C}'$ is tightly concentrated around its mean. For this, we will need the following variant of Chernoff's bound, which follows easily from Theorem 4.3.

LEMMA 7.3. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be independent binary random variables. Define $\mathbf{X} = \sum_{i=1}^n \mathbf{X}_i$ and let $\mu = \mathbb{E}[\mathbf{X}]$. For all $0 < \varepsilon < 1$,*

$$\mathbb{P}\left((1-\varepsilon)\mu \leq \mathbf{X} \leq (1+\varepsilon)\mu\right) \geq 1 - 2\exp\left(-\frac{1}{3}\varepsilon^2 \mu\right). \tag{7.6}$$

The following result shows that for a fixed cut, (7.3) holds with high probability.

THEOREM 7.4. *Let $0 < \varepsilon < 1$ and let $c > 0$. Suppose that $\kappa$ is so large that $p := \frac{3c \ln n}{\varepsilon^2 \kappa} < 1$. Let $C$ be any cut, then (7.3) holds with probability at least $1 - 2n^{-c}$.*

PROOF. Let $\mathbf{X} = p|\mathbf{C}'|$ be the number of edges in the cut. Note that $\mathbf{X} = \sum_{e \in C} \mathbb{1}_{\{e \text{ kept}\}}$ is a sum of independent indicator random variables, and $\mathbb{E}[\mathbf{X}] = p|C|$. Hence it follows from Lemma 7.3 that

$$\begin{aligned}\mathbb{P}\left((1-\varepsilon)|C| \leq |\mathbf{C}'| \leq (1+\varepsilon)|C|\right) &= \mathbb{P}\left((1-\varepsilon)p|C| \leq \mathbf{X} \leq (1+\varepsilon)p|C|\right) \\ &\geq 1 - 2\exp\left(-\frac{1}{3}\varepsilon^2 p|C|\right). \end{aligned} \tag{7.7}$$

Noting that $|C| \geq \kappa$, we find that $\frac{1}{3}\varepsilon^2 p|C| \geq c \ln n$, from which the claim follows. □

By Corollary 7.1, the number of min-cuts in $G$ is at most $\binom{n}{2}$, so by an application of the union bound, (7.3) holds for all min-cuts in $G$ simultaneously with probability at least $1 - \binom{n}{2} 2n^{-c}$, which tends to 1 if only $c > 2$.

In fact, similar results hold for all polynomial-sized sets of cuts. However, the number of cuts in a graph is $\Theta(2^n)$, so Theorem 7.4 combined with a union bound does not suffice to show that (7.3) holds for all cuts simultaneously. However, a more careful analysis will give the desired result.

THEOREM 7.5. *Suppose that $\kappa \geq \omega(\log n)$. Let $0 < \varepsilon < 1$ and $c > 4$. If $p = \frac{3c \ln n}{\varepsilon^2 \kappa}$, then with probability at least $1 - n^{-(c-4)}$ (7.3) holds for all cuts simultaneously.*

PROOF. Let $C$ be a cut of size $\alpha\kappa$, then (7.3) fails with probability at most

$$2\exp\left(-\frac{1}{3}\varepsilon^2 p|C|\right) = 2n^{-\alpha c}. \tag{7.8}$$

As $\alpha \geq 1$, this improves the previous bound of $2n^{-c}$. By Corollary 7.2, the number of cuts of size $\alpha\kappa$ is at most $\binom{n}{2\alpha}2^{2\alpha}$. By an application of the union bound, it follows that the probability that (7.3) fails for some cut $C$ is at most

$$\sum_{\alpha}\binom{n}{2\alpha}2^{2\alpha}\times 2n^{-\alpha c} \leq \sum_{\alpha} n^{2\alpha}2^{2\alpha}\times 2n^{-\alpha c} = \sum_{\alpha} 2^{2\alpha+2\alpha\log n-\alpha c\log n+1} \leq \sum_{\alpha} 2^{-(c-2-o(1))\log n}. \tag{7.9}$$

The number of different $\alpha$ is at most $\binom{n}{2}$, so the right-hand side is bounded by $\binom{n}{2}n^{-(c-2-o(1))} = o(1)$, which concludes the proof. □

## 3. Random selection, revisited

In the statement of Theorem 7.5 we assume that the edge-connectivity is sufficiently large. The following example will indicate that indeed graphs with small edge-connectivity pose a problem to our method.

The **dumbbell graph** (or **barbell graph**) $D_n$ is obtained by connecting two disjoint copies of $K_n$ by a single edge. The resulting graph is a graph on $2n$ vertices, with edge-connectivity $\kappa(D_n) = 1$, see Figure 1.

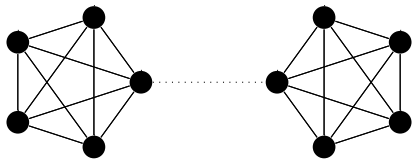

FIGURE 1. The dumbbell graph $D_n$ (here, $n = 5$). The (unique, for $n \geq 3$) minimum cut is indicated by a dotted line.

If all edges are sampled with the same probability $p$, there is a reasonable probability (namely, $1-p$) that the minimum cut in $D_n$ is destroyed, in which case (7.3) fails.

In the case of a dumbbell graph, the natural way to overcome this problem is to slightly adapt the selection procedure, so that the minimum cut is always kept, and the other edges are sparsified in the usual way. This simple idea suggests a more general approach, in which the edge probabilities are chosen in a non-uniform way.

The selection procedure now works as follows: each edge $e$ is selected with probability $p_e$; each edge that is sampled is then given the weight $1/p_e$. As before, the weighting ensure that (7.3) is satisfied in expectation.

We still have to specify the $p_e$ in a useful way. Roughly, the idea is to have $p_e$ large for edges that cross small cuts, and $p_e$ small for edges that do not cross small cuts. We will consider the idea of Benczúr and Karger [BK96; BK02], for which we will need the following definition.

DEFINITION 7.6. *The strength $\kappa_e$ of an edge $e$ is $\kappa_e := \max_{e\subseteq U\subseteq V}\kappa(G[U])$. i.e. the maximum edge-connectivity of subgraphs induced by some set $U \subseteq e$.*

In words, the strength of an edge is the maximum edge-connectivity of an induced subgraph containing the edge $e$.

EXAMPLE. *In the dumbbell graph $D_n$, every edge has strength $n-1$, except the edge connecting the two cliques, which has strength $n-1$.*

Intuitively, the smaller the strength, the more important the edge will be for finding the min-cut, so we will choose $p_e$ inversely proportional to $\kappa_e$.

Later, we will need the following technical result.

LEMMA 7.7. *If $G$ has $n$ vertices and at least $\frac{n \log n}{\varepsilon^2}$ edges, then*

*(i)* $\sum_{e \in E} \frac{1}{\kappa_e} \leq n - 1$;

*(ii)* $\sum_{e \in E} \frac{1}{\kappa_e} \geq \frac{n}{2}$.

PROOF.

(i) Consider a component of $G$ with edge-connectivity $\kappa$. This component contains a cut $C$ of size $\kappa$, and for each edge in the cut $\kappa_e \geq \kappa$. Thus $\sum_{e \in C} \frac{1}{\kappa_e} \leq \frac{|C|}{\kappa} = 1$.

As $C$ is a cut, removing the edges from $C$ increases the number of components in $G$, hence after at most $n - 1$ such steps, we are left with a graph without any edges left.

(ii) Note that the edge-connectivity of any graph is at most the minimum degree in that graph, which is at most the average degree $\frac{2|E(G)|}{n}$. It follows that $\sum_{e \in E(G)} \frac{1}{\kappa_e} \geq \frac{|E(G)|}{2|E(G)|/n}$. □

We can now state and prove the main result of this lecture.

THEOREM 7.8 ([BK02, Theorem 2.6]). *Let $G$ be a connected graph on $n$ vertices, and let $0 < \varepsilon < 1$. Then there is a (weighted) subgraph $G'$ of $G$ with the following properties:*

*(i) $G'$ has $O\left(\frac{n \log n}{\varepsilon^2}\right)$ edges;*

*(ii) for every cut $C$ in $G'$, its value is within a $(1 \pm \varepsilon)$-fraction of the size of $C$ in $G$.*

The proof that we provide here is slightly weaker than the proof in [BK02].

PROOF. We will show that in a suitably constructed random subgraph $\mathbf{G}'$, both properties hold with probability tending to 1, which implies the theorem. Let $c > 7$ be a constant. For each edge $e$ in $G$, let $p_e = \frac{3c \ln n}{\varepsilon^2 \kappa_e}$; for convenience assuming that each $p_e \leq 1$. The random graph $\mathbf{G}'$ is obtained from $G$ by first weighting each edge with weight $1/p_e$, and then retaining it with probability probability $p_e$, independently of all other edges. We will show that both properties hold with probability tending to 1.

Property (i): Note that if $G$ already has at most $\frac{n \log n}{\varepsilon^2}$ edges, the same holds for $\mathbf{G}'$, so the first property holds automatically. We will show that if $G$ has strictly more than $\frac{n \log n}{\varepsilon^2}$ edges, then the expected number of edges in $\mathbf{G}'$ is $O\left(\frac{n \log n}{\varepsilon^2}\right)$ and with high probability the number of edges is not much larger than its expectation.

For convenience, write $\mathbf{X} = |E(\mathbf{G}')|$ for the number of edges in $\mathbf{G}'$. First, let us compute the expected value of $\mathbf{X}$,

$$\mathbb{E}[\mathbf{X}] = \sum_{e \in E(G)} p_e = \frac{3c \ln n}{\varepsilon^2} \sum_{e \in E(G)} \frac{1}{\kappa_e}, \tag{7.10}$$

so by Lemma 7.7

$$\frac{3cn \ln n}{2\varepsilon^2} \leq \mathbb{E}[\mathbf{X}] \leq \frac{3c(n-1) \ln n}{\varepsilon^2}. \tag{7.11}$$

By Chernoff's bound we conclude

$$\mathbb{P}(\mathbf{X} \geq 1.5\mathbb{E}[X]) \leq \exp\left(-\frac{1}{12}\mathbb{E}[\mathbf{X}]\right) \leq \exp\left(-\frac{cn \ln n}{8\varepsilon^2}\right) = o(1), \tag{7.12}$$

so that with probability tending to 1

$$\mathbf{X} \leq \frac{4.5c(n-1) \ln n}{\varepsilon^2}. \tag{7.13}$$

Property (ii): For the second part, we cannot directly apply Chernoff's bound to find strong concentration of cut sizes in $\mathbf{G}'$. The reason is that different edges get different weights depending on their strength. We will solve this problem by decomposing the graph in such a

way that we obtain the required concentration by applying Chernoff's bound to each of the summands.

Suppose that the graph $G$ has $r$ distinct strengths, say $\kappa_{(1)} < \kappa_{(2)} < \ldots < \kappa_{(r)}$. For convenience, define $\kappa_{(0)} = 0$. Let $G^w$ be a weighted version of $G$, in which each edge $e$ in $G$ is weighted by $1/p_e$. We will decompose $G^w$ according to the edge weights: for $i = 1, 2, \ldots, r$, let $F_i$ the graph $G$ restricted to the set of edges with edge weight at least$\kappa_{(i)}$, so that[1]

$$G^w = \frac{\varepsilon^2}{3c \ln n} \sum_{i=1}^{r} (\kappa_{(i)} - \kappa_{(i-1)}) F_i. \tag{7.14}$$

Indeed, if the edge $e$ has weight $\frac{\varepsilon^2}{3c \ln n} \kappa_{(I)}$, then it is an edge in each of $F_1, F_2, \ldots, F_I$. Instead of selecting a random subset of edges in $G$ and then weighting each selected edge, we will consider a random subset of edges in $G^w$. It should be clear that this yields the same random graph $\mathbf{G}'$.

The key observation now is that each $F_i$ is a graph without edge weights. Sampling edges in $G$ means sampling from each of the $F_i$. Note that edges in different $F_i$ are not sampled independently, but that edges within a single $F_i$ are. This means that we can apply Chernoff's bound to each of the graphs $F_i$.

First, consider $F_1$, which is just the original graph $G$ (as all edges have strength at least $\kappa_{(1)} \geq 1$). As the edge-connectivity of $F_1$ is $\kappa_{(1)} \geq \frac{3c \ln n}{\varepsilon^2}$, we can apply Theorem 7.5 to show that with probability at least $1 - n^{-(c-4)}$ all cuts in $F_1$ satisfy (7.3) simultaneously.

Any other $F_i$ is not necessarily connected. However, each component is a maximal $\kappa_{(i)}$-edge-connected induced subgraph of $G$. Fix such a component $H$, and some edge $e$ in $H$. Then $\kappa_e \geq \kappa_{(i)}$, and the maximal $\kappa_e$-edge-conneced sugraph containing $e$ is fully contained in $H$. It follows that strength is conserved when moving to the subgraph induced by $H$. Therefore, by Theorem 7.5, with probability at least $1 - n^{-(c-4)}$ all cuts in $F_i[H]$ satisfy (7.3) simultaneously. As every cut in $F_i$ breaks into a number of cuts in different components, and the number of components is at most $n$, by the union bound, it follows that with probability at least $1 - n^{-(c-5)}$ (7.3) holds for all cuts in $F_i$ simultaneously.

Now suppose that $C$ is some cut in $G$. If $|\mathbf{C}'|$ is not within a $(1 \pm \varepsilon)$-factor of $|C|$, then the corresponding random cut in one of the $F_i$ must be too large or too small. As the number of different $F_i$ is at most $\binom{n}{2}$, it follows that simultaneously all cuts in $G$ satisfy 7.3 with probability at least

$$1 - \binom{n}{2} n^{-(c-5)} > 1 - n^{-(c-7)}, \tag{7.15}$$

which tends to 1 by the assumption that $c > 7$. □

## Bibliography


[BK02] András A. Benczúr and David R. Karger. "Randomized approximation schemes for cuts and flows in capacitated graphs". Available on arXiv [cs/0207078v1], `http://arxiv.org/abs/cs/0207078v1`. 2002.

[BK96] András A. Benczúr and David R. Karger. "Approximating $s - t$ minimum cuts in $\tilde{O}(n^2)$ time". In: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing ACM. 1996, pp. 47–55.


[1] Although we have not defined this notation formally, it should be intuitively clear what it means.