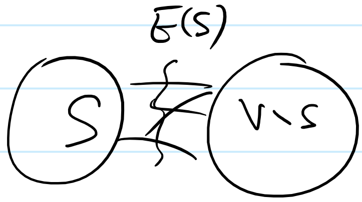
Graph Sparsification II: Cut Sparsifiers

Freitag, 5. Juli 2019 09:40

Graph $G=(V,E)$ undirected

For $S \subseteq V : 1 \leq |S| < n$: $(S, V\setminus S)$ is called a cut

Let us define $E(S)$ as the set of edges across the cut



Size of a cut $(S, V\setminus S)$: $e(S) = |E(S)|$

(if G is weighted $e(S)$ is the sum of the weights in $E(S)$)

Minimum cut size: $\lambda(G)$ (a cut of size $\lambda(G)$ is called a minimum cut)

Cut Sparsifier: Goal

Given an undir. and unweighted graph $G=(V,E)$, find a weighted graph $H=(V,E')$ with $E' \subseteq E$ and weights $w: E' \to \mathbb{R}^+$ s.t. for all cuts $(S, V\setminus S)$:

$$(1-\varepsilon)\cdot e_G(S) \leq \underbrace{e_H(S)}_{\text{weighted}} \leq (1+\varepsilon)\cdot e_G(S)$$

Warm-up example: $G = K_n$

Sample each edge with prob. $p = \dfrac{c \cdot \ln n}{\varepsilon^2 n}$ (for c suff. large)

weight of sampled edges: $1/p$

(assume $|S| \leq |V\setminus S|$)

$\#\text{edges} = |S|\cdot|V\setminus S| \geq \frac{1}{2} n|S| \Rightarrow X(S)$: # sampled edges

$$\mathbb{E}[X(S)] = p\cdot|S|\cdot|V\setminus S| \geq \frac{1}{2} pn|S|$$

$$\mathbb{P}\left(|X(S) - \mathbb{E}[X(S)]| > \varepsilon\cdot\mathbb{E}[X(S)]\right) \leq 2\cdot e^{-\varepsilon^2 \mathbb{E}[X(S)]/3}$$

$$\leq 2\cdot e^{-\frac{\varepsilon^2}{3}\cdot\frac{1}{2}\frac{c\ln n}{\varepsilon^2}|S|}$$

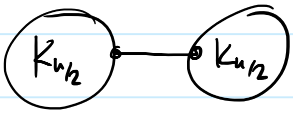
#cuts where the smaller side is $|S|$ is $\leq n^{|S|}$

$$= 2\cdot\frac{1}{n^{\frac{c}{6}\cdot|S|}}$$

General graphs?

Freitag, 5. Juli 2019 09:40

Uniform edge sampling does not work



→ We will see that non-uniform independent edge sampling works

Edge sampling in general graphs

Sample each edge with prob p_e, how do the cuts behave?

Cut Counting

Lemma: In every weighted graph, the number of min. cuts $\leq \binom{n}{2} \leq \frac{n^2}{2}$ and the number of cuts of size $\leq \alpha \cdot \lambda(G)$ is at most $n^{2\alpha}$

Proof! Follows from the properties of the random contraction alg. (see algorithm theory lec.)

Define $q_{\varepsilon,d} := \frac{3(d+4)}{\varepsilon^2} \cdot \ln n = O\left(\frac{\log n}{\varepsilon^2}\right)$ (for $d = O(1)$)

Lemma: Let $G=(V,E)$ be an unweighted graph and let p_e be an assignment of edge prob. and let $H=(V,E')$ to be the graph where each edge $e \in E$ is cont. in E' indep. with prop. p_e

If for all cuts $(S, V \setminus S)$: $p(S) := \sum_{e \in E(S)} p_e \geq q_{\varepsilon,d}$, then

the prob. that $|e_H(S) - p(S)| \leq \varepsilon p(S)$ for all cuts $(S, V \setminus S)$ is at least $1 - \frac{1}{n^d}$.

Proof:

Consider a cut $(S, V \setminus S)$, define $\alpha_S := \frac{p(S)}{\mu}$, where $\mu := \min_S p(S) \geq q_{\varepsilon,d}$

$$\begin{aligned}
\mathbb{P}(|e_H(S) - p(S)| > \varepsilon \cdot p(S)) &< 2 \cdot e^{-\varepsilon^2 \cdot p(S)/3} \quad \text{(using Chernoff)}\\
&= 2e^{-\alpha_S \cdot \varepsilon^2 \cdot \mu/3}\\
&\leq 2e^{-\alpha_S \cdot \frac{\varepsilon^2}{3} \cdot \frac{3(d+4)}{\varepsilon^2} \ln n}\\
&= 2 \cdot n^{-\alpha_S (d+4)}\\
&\leq n^{-\alpha_S (d+3)}
\end{aligned}$$

Union bound over all cuts

$$\mathbb{P}\big(\exists S : |e_H(S) - p(S)| > \varepsilon p(S)\big) \le \sum_S n^{-\alpha_S(d+3)}$$

$$= \sum_{\alpha_S} \underbrace{\sum_{S:\, p(S) = \alpha_S \cdot \mu}}_{\text{cut counting: } \le n^{2\alpha_S} \text{ such cuts}} n^{-\alpha_S(d+3)}$$

$$= \sum_{\alpha_S} n^{-\alpha_S(d+1)}$$

$$\le n^{-d} \qquad \underline{\underline{\#\text{possible } \alpha_S \text{ is polyn. in } n}}$$

(D)

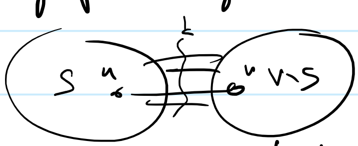
Strong Connectivity

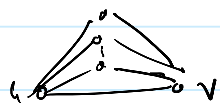
Def: G is called k-connected if every cut is of size $\ge k$

Def: A k-strong component of a graph G is a maximal k-connected vertex-induced subgraph of G

Def: The strong connectivity k_e of an edge e is the maximum k s.t. e is in some k-strong component

Remark: The (standard) connectivity of an edge $\{u,v\}$ is the size of smallest cut separating u & v.





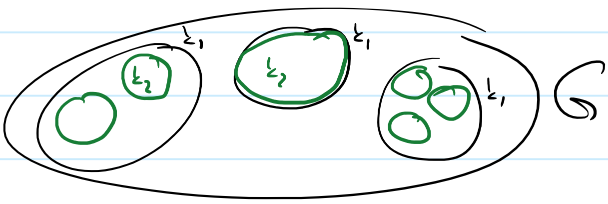
std. connectivity $\ge$ strong connectivity

Lemma: The following holds for k-strong components

(1) k_e is uniquely defined for each edge e

(2) For any k, the k-strong components are disjoint

(3) For $k_1 < k_2$, the k_2-strong comp. are a refinement of the k_1-strong comp.



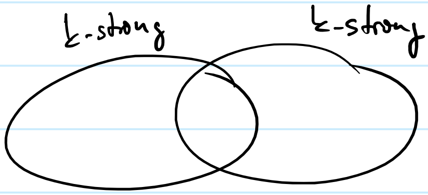
(4) $\sum_{e \in E} \frac{1}{k_e} \le n-1$

Freitag, 5. Juli 2019 09:40

Proof:

(1) by def. ✓

(2) suppose not:



union is k-connected
$\Rightarrow$ contr. to maximality

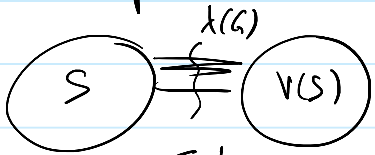
(3) refinement ($k_1 < k_2$)

comp. k_2-conn. $\rightarrow$ k_1-conn $\rightarrow$ subset of a maximal k_1-conn. component

(4) $\sum_{e\in E} \frac{1}{k_e} \le n-1$

min cut of G

$k_e \ge \lambda(G)$ for all $e \in E$



contr. to sum: $\sum_{e\in E(S)} \frac{1}{k_e} \le 1$

remove edges $E(S)$ $\rightarrow$ recurse over comp.

Cut Sparsifier Algorithm

Set $p_e := \frac{1}{k_e} \cdot q_{\varepsilon,d+2} = O\left(\frac{\log n}{\varepsilon^2}\right) \cdot \frac{1}{k_e}$ $\quad (d = O(1))$

Output $H = (V, E')$, where E' cont. every edge $e \in E$ indep. with prob. p_e and weight $1/p_e$

Theorem: W.h.p., $|E'| = O\left(\frac{n \log n}{\varepsilon^2}\right)$ and for every cut $(S, V\setminus S)$

$$(1-\varepsilon)e_G(S) \le e_H(S) \le (1+\varepsilon)e_G(S)$$

Proof: $E[|E'|] = \sum_{e\in E} p_e = q_{\varepsilon,d+2} \cdot \sum_{e\in E} \frac{1}{k_e} \le (n-1)\cdot q_{\varepsilon,d+2} = O\left(\frac{n \log n}{\varepsilon^2}\right)$

Sampling prob. $p_e = \frac{q_{\varepsilon,d+2}}{k_e}$ $\quad (q := q_{\varepsilon,d+2})$

Let $k_1 < k_2 < \dots < k_s$ all the different edge strengths

Consider G as a weighted graph with edge weights

$$w_e = \frac{1}{p_e} = \frac{k_e}{q}$$

Define unweighted $F_1, \dots, F_s$, $F_i = (V, E_i)$, $E_i := \{e \in E : k_e \geq k_i\}$

Observations:

- for all $i \in \{1, \dots, s\}$ and all $e \in E_i$, strength of e in F_i is k_e
- $F_1 = G$, F_{i+1} is a subgraph of F_i
- If we define $k_0 = 0$, we can write G as

$$G = \sum_{i=1}^{s} \frac{k_i - k_{i-1}}{q} \cdot F_i$$

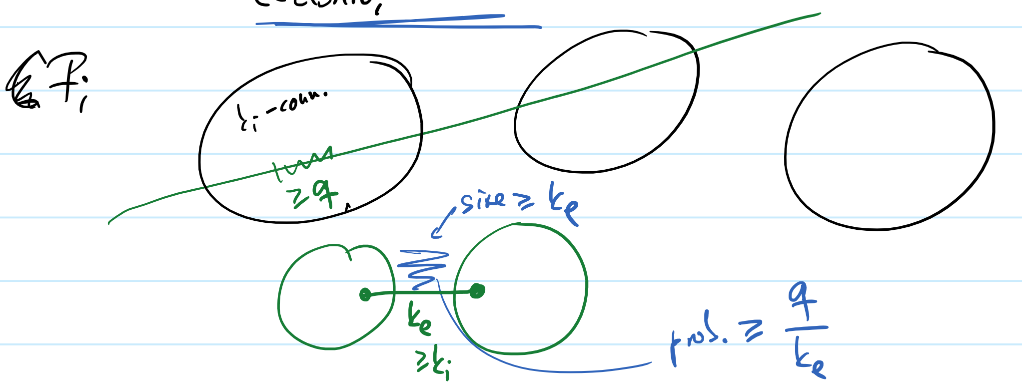
need to show that weight of e in G is $\frac{k_e}{q} = \frac{\overbrace{k_1 - k_0 + k_2 - k_1 + k_3 - k_2 + \dots k_i - k_{i-1}}^{k_i = k_e}}{q}$

where $k_i = k_e$

$\Rightarrow$ if all cuts are close to their exp. size in all F_i

Need to show that in all F_i, for all cuts $(S, V \setminus S)$

$$\sum_{e \in E(S) \cap E_i} p_e \geq q_{\varepsilon,d+2}$$



in each F_i, all cuts are within $(1 \pm \varepsilon)$-factor of exp. with prob. $1 - \frac{1}{n^{d+2}}$

$\Longrightarrow$ union bound over F_i $\Rightarrow$ all cuts in all F_i are good w. pr. $\geq 1 - \frac{1}{n^d}$

$\square$