1 Preliminaries Proof

Claim 1.1. *Let $S, T \subseteq V$ be two disjoint sets. It is possible to find $w(E(S,T))$ using $O(1)$ non-adaptive cut queries.*

Proof. Observe that $\mathrm{cut}_G(S \cup T) = \mathrm{cut}_G(S) + \mathrm{cut}_G(T) - 2w(E(S,T))$. Hence, we can find $w(E(S,T))$ using three queries, one for each of $\mathrm{cut}_G(S)$, $\mathrm{cut}_G(T)$, and $\mathrm{cut}_G(S \cup T)$. □

2 Uniform Sampling Proof

Corollary 2.1. *There exists a matrix $M \in \{\pm 1\}^{q \times a}$ with $q = O(\log^2 a / \log \log a)$, such that every $(200 \log a)$-sparse vector $v \in \mathbb{R}^a$ can be recovered from the measurement vector Mv.*

Proof of sampling result. Let $w \in \mathbb{R}^{|V|}$ be the weights vector adjacent to v. Assume we know that $\deg(v) \leq O(\log n)$. Each index of the measurement vector b can be written as $\langle w, A_i \rangle$, where A_i is the i-th row of A. Notice that we can write A_i as $A_i^+ - A_i^-$ where each vector corresponds to indices with ± 1 respectively. Then, let $B^+ \subseteq V$ be the vertices corresponding to the non-zero entries of A_i^+. Notice that, $\langle w, A_i^+ \rangle = |E(v, A_i^+)|$, which we can query using Claim 1.1. Therefore, we can write $\langle w, A_i \rangle = \langle w, A_i^+ - A_i^- \rangle = |E(v, A_i^+)| - |E(v, A_i^-)|$ and we can recover the entire vector b using $O(\log^2 n)$ queries.

A simple two round algorithm Round 1 - check degree of d. Denote $\deg(v) = d$, then subsample the set V at a rate of $p = 10 \log n / \deg(v)$ to obtain T. Notice that for every $(v, u) = e \in E(v, V \setminus \{v\})$ we have $e \in E(v, T)$ iff $u \in T$. For each $e \in E(v, V \setminus \{v\})$ let X_e be the event that e is in $E(v, T)$ and set $X = \sum_{e \in E(v, V \setminus \{v\})} X_e$. Therefore, $\mathbb{E}[X] = 10 \log n$ Using Chernoff-Hoeffding,

$$\Pr\left[|X - 10 \log n| \notin [5 \log n, 20 \log n]\right] \leq \exp\left(-\frac{-1/2 \cdot \mathbb{E}[X]}{3}\right) \leq \exp\left(-\frac{-5 \log n}{3}\right) \leq n^{-5/3}.$$

A single round Subsample at $\log n$ levels $V \setminus \{v\} = T_0 \supseteq T_1 \ldots \supseteq T_{\log n}$, where each set T_i is obtained by keeping each vertex in T_{i-1} independently with probability $1/2$. Apply sparse recovery on all levels at the same time. By the argument for above some level we will have $1 \leq |E(v, T_i)| \leq 10 \log n$. Query complexity: To check $|E(v, T_i)|$ we need $O(1)$ queries per level using Claim 1.1. For the second process, note that the algorithm performs $O(\log n)$ instances of sparse recovery, one on each level T_i. Since each requires $O(\log^2 n)$ queries we are done. □

3 Star Contraction

Cut Contraction Intuition Fixing a minimum cut $\mathcal{C} \subseteq E$ of G, denote by $c(v)$ the number of edges of $\mathcal{C}$ incident to v. Observe that the probability that when a contraction of a vertex $v \in H \setminus R$ hits an edge in $\mathcal{C}$ is equal to $c_R(v)/d_R(v)$, where $c_R(v) = |\mathcal{C} \cap E(v, R)|$ and recalling $H = \{v \in V \mid d(v) \geq \tau\}$. It can be shown that this is about $c(v)/\deg(v)$, so we will bound this quantity. The proof of this follows using linearity of expectation and Chernoff's bound. Therefore,

$$\Pr\left[\mathcal{C} \text{ is not contracted}\right] = 1 - \prod_{v \in H \setminus R} \left(1 - \frac{c_R(v)}{d_R(v)}\right) \approx 1 - \prod_{v \in H \setminus R} \left(1 - \frac{c(v)}{d(v)}\right).$$

Proposition 3.1. *Let n be a positive integer, $a \in [0,1)$, and $b \geq 1$. Define*

$$\begin{aligned} F(a,b) = \min_{x \in \mathbb{R}^n} \quad & \prod_{i \in [n]} (1 - x_i) \\ \textit{subject to} \quad & \sum_{i \in [n]} x_i = b, \\ & \max_{i \in [n]} 0 \leq x_i \leq a. \end{aligned}$$

Then $F(a,b) \geq (1-a)^{\lceil b/a \rceil}$.

To apply this proposition we need to bound $\sum_{v \in H \setminus R} c(v)/d(v)$ and $\max_{v \in H \setminus R} c(v)/d(v)$. To bound the maximum, notice that $c(v) \leq d(v)/2$ since otherwise one can move the vertex v to the other side of the cut and reduce its size. To bound the sum notice that,

$$\sum_{v \in V} \frac{c(v)}{d(v)} \leq \sum_{v \in V} \frac{c(v)}{\delta(G)} \leq \frac{2|C|}{\delta(G)} \leq 2, \tag{1}$$

where the last inequality is since $|C| \leq \delta(G) \leq \tau$.

Bound on Number of Edges To bound the number of edges, we will first show that there are only two types of edges. Those between centers and those incident to low degree vertices.

1. Edges incident to low degree vertices remains since they don't contract
2. Edges between centers must remain
3. Need to show no vertices in H remain uncontracted. This happens whp since they have degree at least τ (hitting set argument).

If this holds then the number of edges is bounded by $\tilde{O}(|R|^2 + n \cdot \tau) = \tilde{O}((n/\tau)^2 + n\tau)$. Balancing these two terms we set $\tau = n^{1/3}$.

4 Weighted Edge Sampling

4.1 Heavy Hitters

The proof is based on the count-min sketch data structure and standard analysis. We begin by defining the count-min sketch data structure.

Definition 4.1 (Count-Min Sketch)**.** *A Count-Min Sketch is a streaming data structure that maintains a 2 dimensional array of counters $C[1,\ldots,d][1,\ldots,w]$, where all the entries are initially zero. Given parameters (ϵ,δ) we set $w = \lceil e/\epsilon \rceil$ and $d = \lceil \log(1/\delta) \rceil$. We also choose d pairwise independent hash functions $h_1,\ldots,h_d : [n] \to [w]$.*

Whenever we insert an element into the data structure, the data structure increments the counter at the corresponding position by setting $C[i, h_i(x)] \leftarrow C[i, h_i(x)] + 1$ for all $i \in [d]$. To estimate the frequency of an element x we return the minimum over the counters $\min_{i \in d} C[i, h_i(x)]$. The following lemma states the guarantees of the count-min sketch.

Lemma 4.2 (Theorem 1 in [**?**])**.** *Let x be an element with frequency f_x and estimated frequency $\tilde{f}_x$. The count-min sketch gives an estimate $\tilde{f}_x$ such that $f_x \le \tilde{f}_x \le f_x + \epsilon \cdot \sum_{y \in [n]} f_y$ with probability at least $1 - \delta$.*

We restate the lemma for ease of reference.

Lemma 4.3. *Given a graph G on n vertices with integer edge weights bounded by W, parameter $\alpha > 0$, a source vertex $s \in V$, and a target vertex set $T \subseteq V$, one can find an approximate weight vector $\tilde{w} \in \mathbb{R}^{|T|}$ such that with probability $1 - n^{-10}$,*

$$\forall e \in E(s,T), \qquad w(e) \le \tilde{w}(e) \le w(e) + w(E(s,T))/\alpha.$$

The algorithm uses $O(\alpha \log n)$ non-adaptive cut queries.

Proof. The algorithm is based on a count-min data structure with parameters $\epsilon = 1/\alpha$ and $\delta = 1/n^{10}$ to estimate the weights of the edges, where we treat the edge weight vector as the frequency vector we wish to estimate. Notice that there is an explicit mapping from the edges to the bins in the count-min sketch. Therefore, each counter in the count-min sketch is associated with the weight of some subset $F \subseteq E(s,T)$.

Hence, we can recover all the weights with the requisite error and success probability using Lemma 4.2 if we recover the weight of $dw = O(\alpha \log n)$ such subsets. To conclude the proof, we show that each subset can be recovered using $O(1)$ queries. Notice that each subset $F \subseteq E(s,T)$ corresponds to a subset $T' \subseteq T$ such that $F = E(s,T')$. Therefore, we can recover the weight of F by querying the weight of $E(s,T')$ using $O(1)$ queries by Claim 1.1. $\square$