On Mimicking Networks Representing Minimum Terminal Cuts

Arindam Khan[a,*], Prasad Raghavendra[b,**]

[a]*School of Computer Science, Georgia Institute of Technology, Atlanta, USA*
[b]*Electrical Engineering and Computer Sciences, University of California, Berkeley, USA*

Abstract

Given a capacitated undirected graph $G = (V, E)$ with a set of terminals $K \subset V$, a *mimicking network* is a smaller graph $H = (V_H, E_H)$ which contains the set of terminals K and for every bipartition $[U, K - U]$ of the terminals, the cost of the minimum cut separating U from $K - U$ in G is exactly equal to the cost of the minimum cut separating U from $K - U$ in H.

In this work, we improve both the previous known upper bound of 2^{2^k} [1] and lower bound of $(k+1)$ [2] for mimicking networks, reducing the doubly-exponential gap between them to a single-exponential gap as follows:

- Given a graph G, we exhibit a construction of mimicking network with at most k'th Hosten-Morris number ($\approx 2^{\binom{(k-1)}{\lfloor (k-1)/2 \rfloor}}$) of vertices (independent of the size of V).
- There exist graphs with k terminals that have no mimicking network with less than $2^{\frac{k-1}{2}}$ number of vertices.

Keywords: Algorithms, analysis of algorithms, approximation algorithms, combinatorial problems, graph algorithms

---

*Principal corresponding author
**Corresponding author
Email addresses: `akhan67@gatech.edu` (Arindam Khan), `prasad@cs.berkeley.edu` (Prasad Raghavendra)

1. Introduction

Suppose that there are small number of terminals or clients that are part of a huge network such as the internet. Often, it is useful to construct a smaller graph which preserves the properties of the huge network that are relevant to the terminals. For example, if the terminals or the clients are interested in routing flows through the large network, one would want to construct a small graph preserving the routing properties of the original network. The notion of *mimicking networks* introduced by Hagerup et al. [1] is an effort in this direction.

Let G be an undirected graph with edge capacities c_e for all $e \in E$, and a set of k terminals $K(\subset V) := \{v_1, v_2, \ldots v_k\}$. A *mimicking network* for G is an undirected capacitated graph $H = (V_H, E_H)$ such that $K \subseteq V_H$ and for each subset $U \subset K$ of terminals, the cost of the minimum cut separating U from $K-U$ in H is exactly equal to the cost of the minimum cut separating U and $K-U$ in the graph G. Let us assume G to be connected; otherwise we can consider each component separately. Here, we will use edge costs and edge capacities interchangeably. As a corollary, the set of realizable external flows (possible total flows at terminals) in G is preserved in a mimicking network. The vertices of the mimicking network that are not terminals, namely $(V_H - K)$ will be referred to as *Steiner* vertices.

The work of Hagerup et al. [1] exhibited a construction of mimicking networks with at most 2^{2^k} vertices for every graph with k terminals. Subsequently, Chaudhuri et al. [2] proved that there exist graphs that require at least $(k+1)$ vertices in its mimicking network. The same work also obtained improved constructions of mimicking networks for special classes of graphs, namely, bounded treewidth and outerplanar graphs.

Mimicking networks constituted the main building block in the development of an $O(n)$ time algorithm for computing the maximum $s-t$ flow in a bounded treewidth network [1] and for obtaining an optimal solution for the all-pairs minimum-cut problem in the same class of networks [3].

Closely tied to mimicking networks, is the more general notion of *vertex sparsifiers* [4] which only approximately preserve the cut values. While there has been progress [5; 6] in efficient constructions of vertex sparsifiers without Steiner nodes, the power of vertex sparsifiers with Steiner nodes is poorly understood [7]. The following question originally posed by Moitra [4] remains open: *Do there exist cut sparsifiers with $k^{O(1)}$ additional Steiner nodes that yield a better than $O(\log k/\log\log k)$ approximation?* In fact, Moitra [4] points out that there could exist exact cut sparsifiers with only k additional Steiner nodes.

1.1. Our results:

In this paper, we show improved upper and lower bounds for mimicking networks a.k.a. vertex cut sparsifiers with quality 1.

Theorem 1. *There exist graphs with k terminals, for which every mimicking network has at least* $2^{k-1} - 1$ *edges and* $2^{(k-1)/2}$ *vertices.*

Theorem 2. *For every graph G, there exists a mimicking network that has at most k'th Hosten-Morris number* ($\approx 2^{\binom{(k-1)}{\lfloor (k-1)/2 \rfloor}}$) *of vertices.*

Related Work:. In a concurrent and independent work, Krauthgamer et al. [8] showed a slightly weaker lower bound of $\binom{k}{k/2}$ ($< 2^{(k-1)}$) for the number of *edges* of mimicking networks. Chambers et al. had mentioned an upper bound of Dedekind number of vertices for mimicking networks, without an elaborate proof in [9]. Dedekind number is the number of antichains in the partial order $\subseteq$ induced on the subsets of a $(k-1)$-element set by containment. Hosten-Morris number is the number of *intersecting* antichains in this partial order and thus gives a slight improvement over this bound.

2. Preliminaries

In this section, we set up the notation and present formal definitions of the terms related to mimicking networks. Let $c : E \to \mathbb{R}_0^+$ be the capacity function of the graph. Let $h_G : 2^V \to \mathbb{R}_0^+$ denote the cut function of G:

$$h_G(A) = \sum_{e \in \delta(A)} c(e)$$

where $\delta(A)$ denotes the set of edges crossing the cut $[A, V \setminus A]$. Now we define the terminal cut function $h_K^G : 2^K \to \mathbb{R}_0^+$ on K as

$$h_K^G(U) = min_{A \subset V, A \cap K = U} h_G(A)$$

In words, $h_K^G(U)$ is the cost of the minimum cut separating U from $K \setminus U$ in G. Let $S(U)$ be the smallest subset of V such that $h_G(S(U)) = h_K^G(U), S(U) \cap K = U$ i.e., $S(U)$ is the partition containing U in the minimum terminal cut separating U from $K - U$. For any fixed $U \subset K$, the minimum cut $h_K^G(U)$ can be computed efficiently. We will sometimes abuse this notation and use $h_K^G(U)$ to denote both the size of the minimum terminal cut and the set of edges belonging to the minimum terminal cut.

Contraction of edges will be our main tool to construct mimicking networks. Note that given a graph G and an edge e whose endpoints are not both terminals, contracting the edge e in the graph G will not decrease the value of any minimum terminal cut.

Definition 1. *A graph $H = (V_H, E_H)$ is a contraction-based mimicking network of graph $G = (V, E)$ with terminal set K if there exists a function $f : V \to V_H$ such that the edge cost function of H is defined as follows: $c_H(y, z) = \sum_{u,v|f(u)=y,f(v)=z} c(u, v)$ where $(y, z) \in E(H)$ and $(u, v) \in E(G)$.*

3. Exponential Lower bound

In this section we will exhibit the lower bound on the size of mimicking networks using a subtle rank argument. For a set of k terminals K, there are $2^{k-1}-1$ minimum terminal cuts. Let us enumerate these cuts by $[U_i, K \setminus U_i]$ for $i \in \{1, 2, \cdots p(= 2^{k-1} - 1)\}$. Fix $p = 2^{k-1} - 1$ for the remainder of the section. Let $h^G_K(U_i)$ be the minimum terminal cut separating U_i from the rest of the terminals for $i \in \{1, 2, \cdots p(= 2^{k-1} - 1)\}$.

Definition 2. *A minimum terminal cut vector (MTCV) $m^{G,K}$ for graph G with terminal set K, is a p-dimensional vector where i'th coordinate $m^{G,K}_i = h^G_K(U_i)$.*

Let M_k be the set of all possible minimum terminal cut vectors with k terminals. Not all vectors $v \in \mathbb{R}^{2^{k-1}-1}$ can be minimum terminal cut vectors. The submodularity of the cut function introduces constraints on the coordinates of the minimum terminal cut vector. For example there are 3 possible terminal cuts for graphs with terminal set size 3. However [**0.1, 0.1, 0.8**] is not a valid MTCV. First we prove that these minimum terminal cut vectors form a convex set.

Lemma 1. *M_k is a convex cone in $\mathbb{R}^{2^{k-1}-1}$.*

Proof. Note that by scaling the edges of a graph G, the corresponding minimum terminal cut vector also scales. Therefore, it is sufficient to show the convexity of the set M_k.

Let G_1 and G_2 be graphs with terminal set K of size k. Let N_1 and N_2 be their set of non-terminals respectively i.e., $N_i \cup K = V(G_i)$ for $i = 1, 2$. Note that these graphs might have different edge weights or different number of vertices. So depending on the edge values minimum terminal cuts will have different values. Let us assume that t_1 and t_2 be the minimum terminal cut vectors for graphs G_1 and G_2 with the same terminal set K and non negative edge cost functions $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively. We claim that for any nonnegative λ_1, λ_2 such that $\lambda_1 + \lambda_2 = 1$, there exists a graph H with the same terminal set K and edge cost function $\mathcal{C}'$ such that its minimum terminal cut vector $t' = \lambda_1 t_1 + \lambda_2 t_2$. Let H be a complete graph with $V(H) = K \cup N_1 \cup N_2$. We start with all edge costs in H to be 0. Then for $i = 1$ and 2, for all edges $(u, v) \in E(G_i)$, we increase the cost of edge (u, v) in H by $\lambda_i \mathcal{C}_i(u, v)$. The final graph has a MTCV of value $\sum_{i=1}^{2} \lambda_i t_i$. We call H to be a convex combination of G_1 and G_2 with respect to K. □

Now we show the central lemma regarding the range of the minimum terminal cut vectors.

Lemma 2. *The set M_k has nonzero volume.*

Proof. The **0** vector is an MTCV for a completely disconnected graph. For each $i \in \{1, \ldots, p\}$, we will show that a line segment in the i^{th} direction belongs to M_k. By the convexity of the set M_k (Lemma 1) this will imply that the set M_k has nonzero volume, i.e., it is full dimensional.

To demonstrate a line segment along direction $i \in \{1, \ldots, p\}$, we will show that there exist two MTCVs that differ only in the i'th coordinate and are the same in all other $p - 1$ coordinates. Fix a subset U_i of terminals. To construct MTCVs that differ only on the i^{th} coordinate, construct a graph H_i for terminal sets U_i as shown in Figure 1. Connect all terminals in $K - U_i$

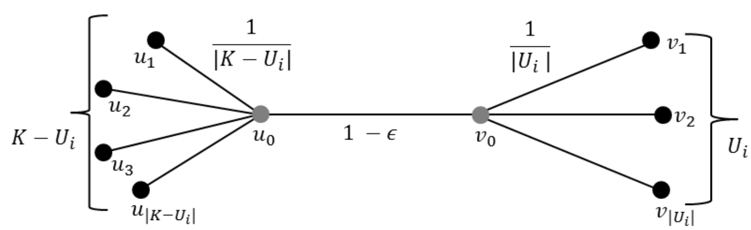

Figure 1: Graph corresponding to terminal cut $[U_i, K \setminus U_i]$

to a non-terminal u_0 with edge costs $1/|K - U_i|$. Connect all terminals in U_i to another non-terminal v_0 with edge costs $1/|U_i|$. Put an edge between u_0 and v_0 with edge cost $(1-\epsilon)$ where $0 < \epsilon < min\{1/|U_i|, 1/|K - U_i|\}$. So, the value of the minimum terminal cut separating U_i from $K - U_i$ is $(1 - \epsilon)$ and it contains only the edge (u_0, v_0). All other minimum terminal cuts have value ≤ 1 and do not contain the edge (u_0, v_0) as any cut containing (u_0, v_0) and some other edge has value at least $(1-\epsilon) + min\{1/|U_i|, 1/|K - U_i|\} > 1$. So, we can change the value of ϵ between 0 and $min\{1/|U_i|, 1/|K - U_i|\}$ to obtain a line segment contained in M_k along the direction i. □

Definition 3. *For a given graph G with terminal set K, the cut matrix S_G is a $p \times |E(G)|$ matrix where $S_{ij} = 1$ if edge $e_j \in h_K^G(U_i)$ and 0 otherwise.*

Theorem 3. *(Restatement of Theorem 1) There exist graphs with k terminals, for which every mimicking network has at least $2^{k-1} - 1$ edges and $2^{(k-1)/2}$ number of vertices.*

Proof. Suppose every graph G with k terminals has a mimicking network with t vertices.

Consider a mimicking network H with t vertices for a graph G with k terminals. There are $2^{t-1} - 1$ possible cuts in the graph H. Therefore, there are at most $(2^{t-1} - 1)^p$ different cut matrices S_H of H. The specific cut matrix S_H depends on the weights of the edges in H.

Let us refer to these matrices as $S_1, S_2, \ldots, S_{(2^{t-1}-1)^p}$. Each matrix S_i can be thought of as a linear map $S_i : \mathbb{R}^{\binom{t}{2}} \to \mathbb{R}^{2^{k-1}-1}$. For every graph G, there exists a choice of weights w_{ij} for the edges of H, and a choice of the cut matrix S_ℓ (determined by the weights), such that $S_\ell w$ is equal to the minimum terminal cut vector h_K^G of the graph G. Therefore, the set M_k of all the MTCVs is in the union of the ranges of the linear maps $\{S_i\}_{i=1}^{(2^{t-1}-1)^p}$.

However, since M_k has non-zero volume (is of full dimension), at least one of the linear maps S_i must have rank at least $2^{k-1} - 1$. Therefore $\binom{t}{2} \geq 2^{k-1} - 1$, which implies that $t \geq 2^{(k-1)/2}$.

□

4. Improved Upper Bounds on Size of Mimicking Networks

Theorem 4. *(Restatement of Theorem 2) For every graph G, there exists a mimicking network that has at most k'th Hosten-Morris number ($\approx 2^{\binom{(k-1)}{\lfloor (k-1)/2 \rfloor}}$) of vertices.*

Proof. We present a more accurate analysis of the algorithm in [1] (restated in Algorithm 1) that constructs the mimicking network from the graph.

While the algorithm creates $2^{2^{k-1}-1}$ clusters, we show that by the properties of the terminal cut structure many of the clusters are empty. We upper bound the number of vertices in H by k'th Hosten-Morris number to complete the proof.

For a terminal cut $[U, K - U]$ where $v_k \notin U$, let $\{S(U), V_G - S(U)\}$ denote the partition induced by the minimum cut separating $[U, K - U]$. If there are multiple minimum terminal cuts, we take any one with smallest cardinality $|S(U)|$. Now let us prove two structural properties of these minimum terminal cuts.

1. Find all minimum terminal cuts using the max-flow algorithm;
2. Partition the graph into $2^{2^{k-1}-1}$ clusters $\mathcal{C}_1, \mathcal{C}_2 \cdots \mathcal{C}_{2^{2^{k-1}-1}}$ such that two vertices u, v belong to the same cluster if they appear on the same side of all the minimum terminal cuts ;
3. Contract each non-empty cluster into a single node ;
4. *Return* the contracted graph H ;

Algorithm 1: ALGORITHM FOR EXACT-CUT-SPARSIFIER

Lemma 3. *If $X \subseteq Y \subseteq K$ then $S(X) \subseteq S(Y)$.*

Proof. From the submodularity property of cuts, we get

$$\begin{aligned}(h_G(S(X)) + h_G(S(Y))) &\geq (h_G(S(X) \cup S(Y)) + h_G(S(X) \cap S(Y))) \\ &\geq (h_G(S(X \cup Y) + h_G(S(X \cap Y))) = (h_G(S(Y)) + h_G(S(X))).\end{aligned}$$

Here the second inequality follows from the fact that $h_G(S(X) \cup S(Y)) \geq h_G(S(X \cup Y))$ and $h_G(S(X) \cap S(Y)) \geq h_G(S(X \cap Y))$. Now as all the inequalities are tight, we get $h_G(S(X) \cup S(Y)) = h_G(S(X \cup Y)) = h_G(S(Y))$ and $h_G(S(X) \cap S(Y)) = h_G(S(X \cap Y)) = h_G(S(X))$. We have $h_G(S(X) \cap S(Y)) = h_G(S(X))$, but recall that among all the minimum cuts separating $(X, K - X)$, $S(X)$ has the smallest cardinality. This implies $S(X) \subseteq S(Y)$. □

Lemma 4. *If $X \cap Y = \phi$ then $S(X) \cap S(Y) = \phi$.*

Proof. The lemma follows from Lemma 3 and from $X \subseteq K - Y$. □

Note that each cluster created by Algorithm 1, is basically the intersection of partitions containing $S(X)$ for some minimum terminal cuts $(X, K - X)$ and the complement of $S(X)$ for the remaining minimum terminal cuts. Let $B \subseteq \{U \subset K, v_k \notin U\}$ i.e., B is a collection of subsets of K that do not contain v_k. Let us define $A(B) = (\cap_{Z \in B} S(Z)) \cap (\cap_{W \subseteq \{K - v_k\}, W \notin B} \overline{S(W)})$. Each $A(B)$ corresponds to a cluster produced by the algorithm. We will show that $A(B)$ is empty for many choices of B.

Lemma 5. *If $A(B) \neq \phi$ then B is an upward closed set i.e., $(\forall P \in B, P \subseteq Q \subseteq \{K - v_k\} \Rightarrow Q \in B)$.*

Proof. Suppose that there exists a $Q \notin B$ such that for some $P \in B$ and $Q \supseteq P$. From Lemma 3, $S(P) \subseteq S(Q)$. By definition, $A(B) \subseteq S(P) \cap \overline{S(Q)}$. Hence, we get $A(B) \subseteq S(P) \cap \overline{S(Q)} = \phi$, a contradiction. □

From Lemma 5, if $A(B) \neq \phi$ then B is an upward closed set. Now the minimal elements of the upper sets form an antichain. Moreover from Lemma 4, two completely disjoint elements in B lead to an empty region. So the number of non-empty clusters is upper bounded by the number of antichains of subsets of $(k-1)$-element sets where any two members of the antichain have non-empty intersection. This number is k'th Hosten-Morris number. $\square$

The above techniques can be extended to get improved upper bounds for mimicking networks for several special graph families. For example, for planar graphs with all terminals on the outer face, the number of vertices (non-empty clusters) produced by Algorithm 1 is $O(k^4)$. The key observation is that instead of all 2^k terminal cuts $[U, K-U]$, considering only $O(k^2)$ minimum terminal cuts where both U and $K-U$ are contiguous sets of terminals on the outer face, suffice to get all non-empty regions.

Using this improved upper bound for general graphs and a more careful analysis of the algorithm for bounded treewidth graphs in [2], we obtain an improved upper bound of $k \cdot 2^{\binom{(2t+1)}{(2t+1)/2}}$ for graphs with treewidth t. Similarly, a tighter analysis on the number of vertices in the dual graph of the planar triangulation of an outerplanar graph gives an improved bound of $5k-9$ for the number of vertices in the mimicking network.

Another interesting observation is that we can use $Y-\Delta$ reduction, well-studied in graph theory literature (See [10]), to construct better mimicking networks for trees. $Y-\Delta$ reduction preserves minimum terminal cut values by applying appropriate edge capacities in each reduction. For example, one can apply $Y-\Delta$ reductions on a tree T with k terminals to obtain a tree T' with k terminals such that each non-terminal vertex in T' has degree at least 3 and all leaves are terminals. So at most there are $(2k-2)$ vertices and $(2k-3)$ edges. We can add appropriate $d(>> \Sigma_{e \in E} c(e))$ capacity edges (if needed) in T' to make the tree 3-regular and the set of terminals as the set of leaves. Adding these edges does not affect any minimum terminal cuts. Notice that then we can apply $Y-\Delta$ transformation on all nonterminals on either odd level or even level of this tree. The maximum among the number of nonterminals in odd and even level is at least $\lceil \frac{k-2}{2} \rceil$ and we delete all such vertices when we apply $Y-\Delta$ transformation. Thus $|V_H| \leq k + \lfloor \frac{k-2}{2} \rfloor = \lfloor (\frac{3k}{2}) - 1 \rfloor$. This gives a mimicking network for trees with at most $\lfloor (\frac{3k}{2}) - 1 \rfloor$ vertices.

5. Conclusion

The results of this work reduce the gap between the upper and lower bounds for the size of the mimicking networks on general graphs from doubly-exponential to exponential. Our techniques also yield improved bounds for special classes of graphs. The main question that remains open is whether there exist exponential sized mimicking networks for general graphs.

Acknowledgements

We thank Jugal Garg, Ruta Mehta, Prasad Tetali and László A. Végh for helpful discussions and the reviewers for their constructive comments.

References

[1] T. Hagerup, J. Katajainen, N. Nishimura, P. Ragde, Characterizations of k-terminal flow networks and computing network flows in partial k-trees, in: SODA, 1995, pp. 641–649.

[2] S. Chaudhuri, K. V. Subrahmanyam, F. Wagner, C. D. Zaroliagis, Computing mimicking networks, Algorithmica 26 (1) (2000) 31–49.

[3] S. R. Arikati, S. Chaudhuri, C. D. Zaroliagis, All-pairs min-cut in sparse networks, J. Algorithms 29 (1) (1998) 82–110.

[4] A. Moitra, Approximation algorithms for multicommodity-type problems with guarantees independent of the graph size, in: FOCS, 2009, pp. 3–12.

[5] T. Leighton, A. Moitra, Extensions and limits to vertex sparsification, in: STOC, 2010, pp. 47–56.

[6] K. Makarychev, Y. Makarychev, Metric extension operators, vertex sparsifiers and lipschitz extendability, in: FOCS, 2010, pp. 255–264.

[7] J. Chuzhoy, On vertex sparsifiers with steiner nodes, in: STOC, 2012, pp. 673–688.

[8] R. Krauthgamer, I. Rika, Mimicking networks and succinct representations of terminal cuts, in: SODA, 2013, pp. 1789–1799.

[9] E. W. Chambers, D. Eppstein, Flows in one-crossing-minor-free graphs, in: ISAAC (1), 2010, pp. 241–252.

[10] T. Feo, J. Provan, Delta-wye tranformations and the effective reduction of two-terminal planar graphs, Operations Research 41 (1993) 572–582.

Appendix A. Comparison of Different Bounds

Hosten-Morris numbers are a rapidly-growing integer sequence defined as follows: Consider the partial order $\subseteq$ induced on the subsets of an $(n-1)$-element set by containment. The n^{th} Hosten-Morris number $Z(n)$ counts the number of intersecting antichains in this partial order. Equivalently, Hosten-Morris numbers counts the number of self-dual monotone Boolean functions of n variables or the number of simplicial complexes on the set $\{1, ..., n-1\}$ such that no pair of faces covers all of $\{1, ..., n-1\}$. *Dedekind numbers* $M(n-1)$ are related numbers which count the number of antichains in this partial order. Kleitman and Markowsky showed that:

$$\binom{n}{\lfloor n/2 \rfloor} \leq \log M(n) \leq \binom{n}{\lfloor n/2 \rfloor}(1 + O(\log n/n)) \qquad \text{(A.1)}$$

Note that $M(k-2) \leq Z(k) \leq M(k-1)$. Asymptotically $Z(k) \approx 2^{\binom{(k-1)}{\lfloor (k-1)/2 \rfloor}}$.

Table A.1: Different bounds related to N(k)

k	Lower bound	Upper bound	k'th Hosten-Morris Number $Z(k)$	$(k-1)$th Dedekind No. $M(k-1)$	$2^{2^{k-1}} - 1$
2	2	2	2	2	3
3	3	3	4	5	15
4	5	5	12	19	255
5	6	6	81	167	65535
6	9	2646	2646	7580	4.29×10^9

Appendix B. Improved Constructions for Special Classes of Graphs

Appendix B.1. Bounded Treewidth Graphs

Theorem 5. *Let G be a n-vertex network of treewidth t and K is the set of terminals such that $|K| = k$, then we can construct a mimicking network for G that has size at most $k \cdot 2^{\binom{(2t+1)}{(2t+1)/2}}$.*

Proof. Let us prove this using induction on k. For $k \leq 2(t+1)$, the induction hypothesis holds from Theorem 2. Now we assume that the theorem holds for all $k' < k$ and we will show that it holds for the graph with k terminals too.

Let $T = (I, F)$ be an augmented binary tree decomposition of G. Let G_v^- and G_v^+ are subgraphs of G spanned by T_v^-(the subtree of T rooted at $v \in I$) and $T_v^+ = T \setminus T_v^-$ (the rest of the tree). Define $K_v^+ = K \cap V(G_v^+)$, $K_v^- = K \cap V(G_v^-)$ and $U = (X(parent(v)) \cap X(v))$. Note that from the continuity property of tree decomposition, U is a vertex separator. Now if $|K_v^+ \cup U| < k$ and $|K_v^- \cup U| \leq 2(t+1)$ then we can construct a mimicking network of size $(k-1) \cdot 2^{\binom{2(t+1)}{(t+1)}}$ for G_v^+ with $K_v^+ \cup U$ as terminal set and a mimicking network of size $2^{\binom{2(t+1)}{(t+1)}}$ for G_v^- with $K_v^- \cup U$ as terminal set. *Glueing* these two networks at U will give a mimicking network of the required size for G with terminals K.

Now let us show how to find such a v satisfying $|K_v^+ \cup U| < k$ and $|K_v^- \cup U| \leq 2(t+1)$. Let $w \in I$ such that $|K_w^-| > 2(t+1)$. Such a vertex exists as $k > 2(t+1)$. Let l and r be the children of w in T. Let $X(l) \cap X(w) = U, U \cap K = Z_1, U \setminus K = N$ i.e., Z_1 and N are the set of terminals and nonterminals in U. Also assume $(V(G_l^-) \setminus X(w)) \cap K = A, (V(G_r^-) \setminus X(w)) \cap K = B$ i.e., A (or B) are the terminals present only in G_l^-(or G_r^-) but not in $X(w)$. Let $Z_2 = (X(w) \setminus (A \cup B)) \cap K$ i.e., the set of terminals present in $X(w)$ but not in G_l^- or G_r^-. Let $|A| = a, |B| = b, |Z_1| = z_1, |Z_2| = z_2$. Then the total number of terminals in $Q_w^- = a + b + z_1 + z_2 > 2(t+1)$. W.l.o.g., assume $a > b$. Then we get $a > (t+1) - \frac{z_1+z_2}{2}$. Also we get, $z_1 + z_2 + N \leq |X(w)| \leq (t+1)$. On the other hand, $|K_l^+ \cup U| \leq (k - a + N) \leq k - (t+1) + \frac{z_1+z_2+N}{2} + \frac{N}{2} \leq k - (t+1) + \frac{(t+1)}{2} + \frac{N}{2} \leq k + \frac{N-(t+1)}{2} < k$. Hence if $|K_l^- \cup U| \leq 2(t+1)$ then $v = l$. Otherwise repeat this with l and its children to get the desired vertex v. □

Appendix B.2. Trees

We will use $Y - \Delta$ reduction to construct better mimicking network for tree. $Y - \Delta$ reducibility is well-studied in graph theory literature (See [10]).

A graph G is said to be $Y - \Delta$ reducible to a simple graph H, if G can be reduced to H by repeated applications of following four reductions and two transformations. We claim that $Y - \Delta$ reduction preserves minimum terminal cut values by applying appropriate edge capacities in each reduction as follows:

R0: *Loop reduction:* Just delete the loop.

R1: *Degree-one reduction:* Delete a degree one non-terminal and its incident edge. Note that the incident edge on the non-terminal does not take part in any minimum terminal cut

R2: *Series reduction:* Delete a degree two non-terminal y and its two incident edges xy and yz and add a new edge xz with capacity $min(c(x,y), c(y,z))$.
R3: *Parallel reduction:* Replace parallel edges by a single edge with capacity of the new edge as the sum of the capacity of parallel edges .
$Y-\Delta$ *transformation:* Let x be a degree 3 nonterminal with neighbors u, v and w, then we can delete x and add edges $(u,v), (v,w), (w,u)$ with edge capacities $\frac{c(u,x)+c(v,x)-c(w,x)}{2}$, $\frac{c(v,x)+c(w,x)-c(u,x)}{2}$ and $\frac{c(u,x)+c(w,x)-c(v,x)}{2}$ respectively.
$\Delta-Y$ *transformation:* Delete the edges of a Δ xyz, add in a new vertex w and new edges wx, wy, wz with edge capacity $c(x,y)+c(x,z)$, $c(x,y)+c(y,z)$ and $c(x,z)+c(y,z)$ respectively.

Theorem 6. *Given an undirected, capacitated tree $T = (V,E)$ and a set $K \subset V$ of terminals of size k, we can construct a mimicking network $T_H = (V_H, E_H)$ which is a tree and $|V_H| \leq 2k-2$. We can also create a mimicking network of T which is outerplanar and has at most $\lfloor(\frac{3k}{2})-1\rfloor$ vertices and $(2k-3)$ edges.*

Proof. Assume we get the tree T' after applying reductions R0-R3 of $Y-\Delta$ reduction on T. It follows that each non-terminal vertex in T' has degree at least 3 and all leaves are terminals. So at most there are $(2k-2)$ vertices and $(2k-3)$ edges. Let $d >> \Sigma_{e\in E} c(e)$. Now we add appropriate d capacity edges (if needed) in T' to make the tree 3-regular and the set of terminals as the set of leaves. Adding these edges does not affect any minimum terminal cuts. Notice that we can apply $Y-\Delta$ transformation on all nonterminals on either odd level or even level. The maximum among the number of nonterminals in odd and even level is at least $\lceil\frac{k-2}{2}\rceil$ and we delete all such vertices when we apply $Y-\Delta$ transformation. Thus $|V_H| \leq k + \lfloor\frac{k-2}{2}\rfloor = \lfloor(\frac{3k}{2})-1\rfloor$. As an outerplanar graph remains outerplanar after applying $Y-\Delta$ transformation, the resultant graph obtained from T is an outerplanar graph. □

Appendix B.3. Outerplanar Graphs

Theorem 7. *Given an undirected outerplanar graph $G = (V,E)$ with a set $K \subset V$ of terminals of size $k \geq 3$, $5k-9$ vertices are sufficient to construct an exact cut sparsifier.*

Proof. Let us assume the graph to be 2-connected, otherwise appropriate 0-cost edges can be added to make it 2-connected keeping the outerplanarity property intact. Consider the dual graph of the planar triangulation of G. Removing the vertex corresponding to the outer face gives a dual tree T. We will refer to the vertices of the tree T as nodes. Removing the degree 3 nodes divides T into several components that are paths or isolated vertices. We will slightly abuse notation and call isolated vertices as paths of length 0. If some component contains a leaf node of T we call it a *leaf-path* and otherwise call it a *nonleaf-path*. If u_i is a node of T, then we denote $D(u_i)$ to be the dual face corresponding to u_i in G.

We define $\mathcal{D}(P)$ to be the dual complex of a path $P := u_1, u_2 \dots u_m$ in T so that $\mathcal{D}(P) = G[\cup_{u_i \in V(P)} D(u_i)]$ i.e., subgraph induced by vertices in the dual of P. Assume C_0, C_1, C_2, C_3, C_4 are the sets of leaf-paths whose corresponding dual complexes in G contain 0, 1, 2, 3 or > 3 terminals. Similarly assume C_5, C_6, C_7 are the sets of non-leaf paths whose corresponding dual complexes in G contain 0, 1 or > 1 terminals respectively. Let $c_i = |C_i|$ for $i \in [7]$. Note that we can delete paths in C_0 as their dual complexes are not a part of any minimum terminal cuts.

Now we will find the mimicking networks of dual complexes of all leaf-paths and nonleaf-paths and glue them together at the splitting vertices (dual complexes of degree 3 nodes in T). Essentially we are splitting the original graph G into several subgraphs at $D(v_i)$'s where $v_i \in T, deg(v_i) = 3$. After the splitting, apart from the terminals in K, we consider the splitting vertices also as terminals and thus dual complexes of leaf-paths in C_1, C_2, C_3 contain at most 3, 4, 5 terminals respectively. Similarly, dual complexes of non-leafpaths in C_5, C_6 contain at most 4, 5 terminals respectively. So dual complexes of paths in C_1, C_2, C_3, C_5, C_6 can be replaced by mimicking networks of size 3, 5, 6, 5 and 6 respectively. C_4 and C_7 need special care.

For $P \in C_4(|P \cap K| = x_P)$, start with the leaf node and divide the path into disjoint subpaths $P_0^4, P_1^4, \dots, P_s^4$ such that the number of terminals (including splitting vertices as terminals) in the dual complexes of these subpaths are at most 5 and $|\mathcal{D}(P_0^4) \cap K| \geq 4$, and $|P_s^4 \cap K| \geq 1$ and for $P_i (i \neq 0, s), |P_i^4 \cap K| \geq 3$. Find mimicking network of $\mathcal{D}(P_i^4)$ for $i = 0, 1, \dots s$ and then glue them back together to get mimicking network of $\mathcal{D}(P)$. Now each time when we glue back a mimicking network of size m, we increase the number of vertices by $(m-2)$, apart from the first mimicking network. Thus, the mimicking network for P consists of $\leq 2 + (6-2) + \frac{(x_P-4)}{2}(6-2) + (5-2) = (2x_P + 1)$ vertices.

Similarly, for nonleaf-path $P \in C_7$ $(|P \cap K| = w_P)$, we can divide the path into disjoint subpaths $P_0^7, P_1^7, \dots, P_t^7$ such that the number of terminals (including splitting vertices as terminals) in the dual complexes of these subpaths are at most 5. The mimicking network for P consists of $\leq 2 + (6 - 2) + \frac{(w_P-1)}{2}(6-2) = (2w_P + 4)$ vertices.

Counting terminals on all the paths gives the following inequality:

$$(c_1 + 2c_2 + 3c_3 + c_6 + \sum_{P \in C_4} x_P + \sum_{P \in C_7} w_P) \leq k \tag{B.1}$$

On the other hand, the number of non-leaf path is at most $k-3$. Thus we get the following inequality:

$$(c_5 + c_6 + c_7) \leq (k-3) \tag{B.2}$$

Similarly, the number of degree 3 nodes in $T \leq (c_1 + c_2 + c_3 + c_4 - 2)$. Recall that each time when we glue back a mimicking network of size m, we increase the number of vertices by $(m-2)$, apart from the first mimicking network. Then the size of the mimicking network
$\leq 2 + c_1(3-2) + c_2(5-2) + c_3(6-2) + c_5(5-2) + c_6(6-2) + \sum_{P \in C_4}(2x_P + 1 - 2) + \sum_{P \in C_7}(2w_P + 4 - 2) + (c_1 + c_2 + c_3 + c_4 - 2)$
$= 2 + 2c_1 + 4c_2 + 5c_3 + 3c_5 + 4c_6 - 2 + c_4 + \sum_{P \in C_4}(2x_P + 1 - 2) + \sum_{P \in C_7}(2w_P + 4 - 2)$
$= 2c_1 + 4c_2 + 5c_3 + 3c_5 + 4c_6 + 2c_7 + \sum_{P \in C_4}(2x_P) + \sum_{P \in C_7}(2w_P)$
$\leq 2(c_1 + 2c_2 + 3c_3 + c_6 + \sum_{P \in C_4}(x_P) + \sum_{P \in C_7}(w_P)) + 3(c_5 + c_6 + c_7)$
$\leq 2k + (3k - 9) = 5k - 9$.

□