Faster Minimum k-cut of a Simple Graph

Jason Li
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA, USA
jmli@cs.cmu.edu

Abstract—We consider the (exact, minimum) k-CUT problem: given a graph and an integer k, delete a minimum-weight set of edges so that the remaining graph has at least k connected components. This problem is a natural generalization of the global minimum cut problem, where the goal is to break the graph into $k = 2$ pieces.

Our main result is a (combinatorial) k-CUT algorithm on simple graphs that runs in $n^{(1+o(1))k}$ time for any constant k, improving upon the previously best $n^{(2\omega/3+o(1))k}$ time algorithm of Gupta et al. [FOCS'18] and the previously best $n^{(1.981+o(1))k}$ time combinatorial algorithm of Gupta et al. [STOC'19]. For combinatorial algorithms, this algorithm is optimal up to $o(1)$ factors assuming recent hardness conjectures: we show by a straightforward reduction that k-CUT on even a simple graph is as hard as $(k-1)$-clique, establishing a lower bound of $n^{(1-o(1))k}$ for k-CUT. This settles, up to lower-order factors, the complexity of k-CUT on a simple graph for combinatorial algorithms.

Keywords-minimum k-cut; graph algorithms; randomized algorithms

I. Introduction

We consider the (exact, minimum) k-CUT problem: given a graph and an integer k, delete a minimum-weight set of edges so that the remaining graph has at least k connected components. This problem is a natural generalization of the global minimum cut problem, where the goal is to break the graph into $k = 2$ pieces. This problem has been actively studied in theory of both exact and approximation algorithms, where each result brought new insights and tools on graph cut algorithms.

Goldschmidt and Hochbaum gave the first polynomial-time algorithm for fixed k, with $O(n^{(1/2-o(1))k^2})$ runtime [1]. Since then, the exact exponent in terms of k has been actively studied. The textbook minimum cut algorithm of Karger and Stein [2], based on random edge contractions, can be adapted to solve k-CUT in $\tilde{O}(n^{2(k-1)})$ (randomized) time. This bound was improved recently for the first time by an algorithm of Gupta et al. [3], which runs in $n^{(1.981+o(1))k}$ (randomized) time. The deterministic algorithms side has seen a series of improvements since then [4], [5], [6]. The fastest algorithm for general edge weights is due to Chekuri et al. [6]. It runs in $O(mn^{2k-3})$ time and is based on a deterministic tree packing result of Thorup [5]. Lastly, if the edge weights of the input graph are integers bounded by

$n^{O(1)}$ (in particular, exponent independent of k), then k-CUT can be solved in $n^{(2\omega/3+o(1))k}$ deterministic time [7], where $\omega < 2.373$ is the matrix multiplication constant [8], [9].

Lower bounds for the k-CUT problem have also been studied actively in the past decade. k-CUT on real-weighted graphs is at least as hard as minimum weighted $(k-1)$-clique [7], the latter of which is conjectured to require $n^{(1-o(1))k}$ time for any constant k [10]. For k-CUT on unweighted graphs, the lower bound is weakened to $n^{(\omega/3-o(1))k}$, again from a reduction to $(k-1)$-clique, where $\omega < 2.3727$ is the matrix multiplication constant [7]. However, for "combinatorial" algorithms as described in [11], [10], this lower bound is again $n^{(1-o(1))k}$ even for unweighted graphs, under the stronger hardness conjecture of k-clique for combinatorial algorithms [11], [10].

In this paper, we consider the k-CUT on *simple* graphs: graphs that are unweighted and have no parallel edges. Our main result is a (combinatorial) k-CUT algorithm on simple graphs that runs in $n^{(1+o(1))k}$ time for any constant k, improving upon the previously best $n^{(2\omega/3+o(1))k}$ time algorithm [7] and the previously best $n^{(1.981+o(1))k}$ time combinatorial algorithm [3]. For combinatorial algorithms, this algorithm is optimal up to $o(1)$ factors assuming recent hardness conjectures: we show by a straightforward reduction that k-CUT on even a simple graph is as hard as $(k-1)$-clique, establishing a lower bound of $n^{(1-o(1))k}$ for k-CUT. This settles, up to lower-order factors, the complexity of k-CUT on a simple graph for combinatorial algorithms. We remark that this is the first setting for k-CUT, except the restricted $k = 2$ case, where the running time has been determined up to $o(1)$ factors.

Theorem I.1 (Main Result). *For any parameter k, there is a (combinatorial, randomized) algorithm that computes the k-CUT of a simple graph in $k^{O(k)}n^{(1+o(1))k}$ time.*

Theorem I.2 (Lower Bound). *Suppose we assume the conjecture that every combinatorial algorithm for k-clique requires $n^{(1-o(1))k}$ time for any constant k. Then, for any constant k, every combinatorial algorithm for k-CUT of a simple graph also requires $n^{(1-o(1))k}$ time.*

IEEE
computer
society

A. Our Techniques

Our k-CUT algorithm incorporates algorithmic techniques from a wide array of areas, from graph sparsification to fixed-parameter tractability to tree algorithms.

Graph Sparsification: Our first algorithmic ingredient is the Kawarabayashi-Thorup (KT) sparsification algorithm, which originated from the breakthrough paper of Kawarabayashi and Thorup on the deterministic minimum cut problem [12]. At a high level, given any simple graph G with minimum cut λ, the algorithm contracts G into a multigraph of $\tilde{O}(m/\lambda)$ [1] edges so that any minimum cut of G that has at least two vertices on each side gets "preserved" in the contraction. That is, we never contract an edge in any such minimum cut. Kawarabayashi and Thorup used their contraction procedure to provide the first $\tilde{O}(m)$-time deterministic algorithm for minimum cut of a simple graph. They first applied the contraction to G, obtaining a multigraph $\overline{G}$ on $\overline{m} = \tilde{O}(m/\lambda)$ edges, and then ran the $\tilde{O}(\overline{m}\lambda)$-time minimum cut algorithm of Gabow on $\overline{G}$, which works for multi-graphs. This covers the case when the minimum cut of G has at least two vertices on each side; the other case, where the minimum cut consists of a singleton vertex on one side, is trivial. In the KT sparsification, we can also ensure that $\overline{G}$ has $\tilde{O}(n/\lambda)$ vertices, which is the property that we focus on.

One of our technical contributions is extending the KT sparsification algorithm to work for k-CUT. Briefly, we show that modulo a regularity condition, we can contract the input graph G into a multi-graph $\overline{G}$ on $\overline{n} = \tilde{O}(n/\lambda_k)$ vertices, where λ_k is the minimum k-cut of G, to preserve all minimum k-cuts where each side has at least two vertices. (For the case when one component of the minimum k-cut is a singleton vertex, we handle it separately with a simple branching procedure: try each vertex as a singleton component and recursively solve $(k-1)$-cut.) In the spirit of Kawarabayashi and Thorup, we then solve k-CUT on the contracted graph in $\overline{n}^{(1+o(1))k}\lambda_k^k$ time. Since $\overline{n} = \tilde{O}(n/\lambda_k)$, this running time becomes $n^{(1+o(1))k}$, as needed.

Tree Packing: To solve k-CUT in $n^{(1+o(1))k}\lambda_k^k$ time on multi-graphs,[2] we begin with a tree packing result of Thorup [5], which says that we can compute a small collection of trees so that for one tree T, at most $2k-2$ edges of T have endpoints in different components of the minimum k-cut. we apply a reduction in [7] which, at a multiplicative cost of $O(n^k)$ in the running time, produces a tree T such that exactly $k-1$ edges of T have endpoints in different components of the minimum k-cut. (Note that $k-1$ here is the smallest possible.) In other words, if we remove these edges from T, then the connected components

---

[1] Throughout the paper, we use the standard $\tilde{O}(\cdot)$ notation to hide polylogarithmic factors in the running time.

[2] Here, we assume the contracted graph is our input graph now, hence n instead of $\overline{n}$.

in the remaining forest are exactly the components of the minimum k-cut.

Color Coding: Following the last paragraph, our problem thus reduces to this: given a graph $G = (V, E)$ and a tree T on the vertices V, remove some $(k-1)$ edges F of T to minimize the k-cut (in G) formed by the k connected components in $T - F$. Our main technical contribution is providing such an algorithm that runs in $n^{o(k)}\lambda^k$ time.

One major ingredient in our algorithm is the technique of *color coding* due to Alon et al. [13], who first used it for the k-PATH problem in the fixed-parameter setting. To illustrate our approach, let us assume (*with* loss of generality, for sake of exposition) that the tree T is a "spider": it can be decomposed into an edge-disjoint set of paths sharing a common endpoint r; see the black edges in Figure 1 for an example. Call each of the edge-disjoint paths from r be a *branch*. Let us further assume (again, *with* loss of generality) that the optimal set F^* of $(k-1)$ edges consists of one edge from each of $(k-1)$ distinct branches $B_1^*, \ldots, B_{k-1}^*$. Let $S_1^*, \ldots, S_k^* \subseteq V$ be the components of $T - F^*$ with $r \in S_k^*$, and let $OPT \subseteq E$ be the minimum k-cut (in G) with components $S_1^*, \ldots, S_k^*$.

Our first observation is that if there were no edges between S_i^* and S_j^* for every $1 \le i < j \le k-1$, then the problem becomes easy, because the following (polynomial-time) algorithm works:

1) For each branch B, pick the edge e in the branch to remove so that the two connected components $C_1, C_2 \subseteq V$ of $T - e$ minimize $|E[C_1, C_2]|$ (the number of edges in G between C_1 and C_2). Let $f(B)$ be the minimum value $|E[C_1, C_2]|$ for branch B.

2) Select the $(k-1)$ branches $B_1, \ldots, B_{k-1}$ with the smallest values of $f(B_i)$. The $(k-1)$ edges to remove are the edges selected in each of these branches, and the total k-cut value is $\sum_{i=1}^{k-1} f(B_i)$.

In other words, the algorithm processes each branch independently and selects the best $(k-1)$ branches.

In general, if there is an edge in the minimum k-cut between S_1 and S_2 (say), then this edge may contribute to both $f(B_1)$ and $f(B_2)$, in which case it is double-counted. So we always have $\sum_{i=1}^{k-1} f(B_i) \ge |OPT|$, and strict inequality is possible. But if every edge in OPT connects $f(B_k)$ to some $f(B_i)$ ($i \le k-1$), then every edge is counted at most once, so $\sum_{i=1}^{k-1} f(B_i) = OPT$.

This algorithm works in this special setting because no "double-counting" occurs: every edge in OPT is counted exactly once in $\sum_{i=1}^{k-1} f(B_i)$. What if we consider the opposite case scenario, where *many* edges are double-counted? In particular, suppose that there is at least one edge between every two branches B_i^*, B_j^* with $i, j \le k-1$. Let $E_2 \subseteq OPT$ be the edges in OPT in between two such branches.

These edges may fool the simple algorithm described above, but they serve a different useful purpose. Namely,

the edges E_2 *connect* the branches $B_1^*, \ldots, B_k^*$ together, and that is a property we will exploit as follows: Let us randomly color each edge of E either red or green, hoping for the following two properties:

1) For each $i < j \leq k-1$, there is at least one edge in E_2 colored green.
2) All edges in $E \setminus E_2$ incident to vertices in $B_1^*, \ldots, B_k^*$ are colored red.

The properties we require are a bit more specific, but the two conditions above suffice for illustration. If both properties as satisfied, then if we consider the connected components of branches, where two branches are pairwise connected if they share a green edge, then $B_1^*, \ldots, B_k^*$ exactly mark out a single component. And if the algorithm iterates over all such components of branches (and processes each one independently, say), then it will come across exactly $\{B_1^*, \ldots, B_k^*\}$ at some point. Thus, in some sense, we may *assume* that the algorithm knows $B_1^*, \ldots, B_k^*$.

Even with this knowledge, the issue of double-counting still remains. To handle it, we pinpoint down a small set of edges $E' \subseteq E_2$ such that if all edges in E' are green, then the algorithm learns enough information about the double-counted edges to provide the correct answer. The size of E' needs to be small, because we color each edge green with small probability, and yet require that all of E' is colored green with sufficiently large probability.

Tree Algorithms: The algorithm above only handles the case when T is a "spider". What about the general case? Our first idea is to apply *heavy-light decomposition* on the (rooted) tree, breaking it into a disjoint union of branches, with the property that every path from leaf to root intersects the edges of $O(\log n)$ branches. This $O(\log n)$ factor will be paid in the running time as $O(\log n)^k$, which (perhaps surprisingly) can be bounded by $\max\{k^{O(k)}, n\}$; morally, this means that the $O(\log n)$ factor is negligible. The benefit of the HLD is that we once again have a disjoint union of branches. Although these branches may not have a common endpoint, they *almost* do, in the sense that every path from leaf to root now intersects $O(\log n)$ branches instead of 1, which is still small.

Lastly, what about the case when multiple edges are cut from the same branch? To handle this situation, we apply dynamic programming on a tree in a manner similar to [7]. The key observation is that if we remove an edge e from T, then the two components of T_e become independent subproblems, thus making the situation amenable to dynamic programming.

B. Related Work

The k-CUT problem has been studied extensively in the approximate and fixed-parameter settings as well.

Approximation algorithms.: The first approximation algorithm k-CUT was a $2(1-1/k)$-approximation of Saran and Vazirani [14]. Later, Naor and Rabani [15], and also

Ravi and Sinha [16] gave 2-approximation algorithms using tree packing and network strength respectively. Xiao et al. [17] extended Kapoor [18] and Zhao et al. [19] and generalized Saran and Vazirani to give an $(2-h/k)$-approximation in time $n^{O(h)}$. On the hardness front, Manurangsi [20] showed that for any $\epsilon > 0$, it is NP-hard to achieve a $(2-\epsilon)$-approximation algorithm in time $\mathrm{poly}(n,k)$ assuming the Small Set Expansion Hypothesis.

Recently [7], Gupta et al. gave a 1.81-approximation for k-CUT in FPT time $f(k)\mathrm{poly}(n)$ and a $(1+\epsilon)$-approximation in $f(k)n^{(1+o(1))k}$ time. These algorithms do not contradict Manurangsi's work, since k is polynomial in n for his hard instances.

FPT algorithms.: The k-CUT problem was shown to be $W[1]$-hard when parameterized by k by Downey et al. [21]. Kawarabayashi and Thorup give the first $f(\mathsf{Opt}) \cdot n^2$-time algorithm [22] for unweighted graphs. Chitnis et al. [23] used a randomized color-coding idea to give a better runtime, and to extend the algorithm to weighted graphs. Here, the FPT algorithm is parameterized by the cardinality of edges in the optimal k-CUT, not by the number of parts k. For more details on FPT algorithms and approximations, see the book [24], and the survey [25].

II. PRELIMINARIES

All graphs in this paper will be unweighted, undirected multigraphs without self-loops. We denote by n and m the number of vertices and edges in the input graph, respectively. A graph is *simple* if for every two vertices in the graph, there is at most one edge between them. For a graph G, let $V(G)$ and $E(G)$ be its vertex set and edge set, respectively. For vertex-disjoint sets $S_1, \ldots, S_\ell$, denote by $E[S_1, \ldots, S_\ell]$ the set of edges whose endpoints lie in distinct sets S_i, S_j ($i \neq j$). When there are multiple graphs in our context, we use $E_G[S_1, \ldots, S_\ell]$ instead to indicate that the graph in question is G; we use similar notation for other graph functions. For a vertex set $S \subseteq V(G)$ of a graph G, denote $\partial S = E[S, V \setminus S]$ as the set of edges with exactly one endpoint in S. For disjoint vertex sets $S_1, \ldots, S_\ell \subseteq V(G)$, denote $\partial[S_1, \ldots, S_\ell] = E[S_1, \ldots, S_\ell, V \setminus \bigcup_{i=1}^{\ell} S_i]$ as the set of edges with at least one endpoint in some S_i, but not both endpoints in the same S_i; note that $\partial S = \partial[S]$ for any subset $S \subseteq V(G)$. The *degree* of a vertex $v \in V(G)$ is the number of edges incident to it, which equals $|\partial(\{v\})|$.

For a graph $G = (V, E)$, the minimum k-cut will either be denoted as the subsets $S_1^*, \ldots, S_k^* \subseteq V$ that comprise the components of the k-cut, or as the edge set $OPT = E[S_1^*, \ldots, S_k^*]$, the set of edges in the k-cut. We now define the concept of a *nontrivial* minimum k-cut:

Definition II.1 (Nontrivial minimum k-cut)**.** *A minimum k-cut $\{S_1^*, \ldots, S_k^*\}$ is nontrivial if none of the sets $|S_i^*|$ ($i \in [k]$) have size 1.*

Since we work with multigraphs throughout the paper,

every edge has a unique identifier. Whenever we declare a variable e as an edge, we mean its identifier. In particular, if two edges e, e' both have endpoints u, v, then it is possible that $e \neq e'$. We identify each edge by its identifier, rather than its two endpoints (u, v). That is, every variable e designated to an edge is set to the edge's identifier, rather than the tuple (u, v) or the set $\{u, v\}$. We may still say "an edge $(u, v) \in E(G)$", by which we mean an arbitrary edge with endpoints u, v in G (and we do not care about its identifier). Likewise, "for every edge $(u, v) \in E(G)$" means every edge with endpoints u, v in G.

Whenever we *contract* two vertices $u, v \in V(G)$ in a graph G, all edges that survive keep their identifiers. More formally, contraction produces the following graph G':

1) We have $V(G') = V(G) \setminus \{u, v\} \cup \{x\}$ for some new vertex x.
2) For every edge e with endpoints u', v' distinct from u and v, add edge e with the same endpoints u', v' in G'.
3) For every edge e with endpoints u, v' for $v' \neq v$, add edge e with endpoints x, v' in G'.
4) For every edge e with endpoints u', v for $u' \neq u$, add edge e with endpoints u', x in G'.
5) For every edge e with endpoints u, v in G, do not add it to G'.

In particular, the edge identifiers in G and G' match. One benefit to this formulation is that if $OPT \subseteq E$ is a minimum k-cut with components $S_1^*, \ldots, S_k^*$ and we contract two vertices $u, v \in S_i^*$ for some i, then the same set OPT is still a minimum k-cut in the contracted graph.

For a positive integer ℓ, we denote by $[\ell]$ the set of integers from 1 to ℓ (inclusive), $\{1, 2, \ldots, \ell\}$.

Tree Terminology (Section IV): The terminology in this paragraph are specific to Section IV.

For a rooted tree T, let $T(v) \subseteq T$ denote the subtree of T rooted at $v \in V(T)$. For any set $S \subseteq V(T)$, define $T(S) := \bigcup_{v \in S} T(v)$, the union of all (vertices and edges of) trees $T(v)$ over all $v \in S$.

For an edge $e = (u, v) \in V(T)$ where v is the child of u, we say that v is the *child vertex* of edge e and e is the *parent edge* of v.

Given a rooted tree T, the *depth* depth(v) of vertex $v \in V(T)$ is the (unweighted) distance of v to the root of T. Every time we use depth(v), the tree T will be implicit.

A *branch* of T is a path in T that travels "downwards" the rooted tree. More formally, it is a path whose vertices have distinct depths. The vertex with the minimum depth is the *root* of the branch. A *maximal* branch is a branch from the root to a leaf in the tree. When we say a *branch from u to v*, we mean the path from u and v.

For a rooted tree T and $u, v \in V(T)$, we say that u *precedes* v if $v \in T(u)$. We say that two vertices $u, v \in V(T)$ are *incomparable* if u does not precede v and v does

not precede u. Note that u and v are incomparable iff they do not lie on a common branch.

Given a set $S \subseteq V(T)$, denote by $S_\downarrow$ the *minimal elements of S*, the minimal set $S' \subseteq S$ such that every vertex in S is preceded by some vertex in S' (that is, $S \subseteq T(S')$).

Tree Packing.: Our algorithm will use the concept of *tree packing* which, at a high level, reduces the k-CUT to a problem of finding the best way to remove edges in a tree. Tree packings for the k-CUT problem were first introduced by Thorup [5], who used them to obtain a deterministic k-CUT algorithm in time $O(mn^{2k-2})$.

Definition II.2 (T-tree, Definition 2.1 of [7]). *A tree T of G is a ℓ-T-tree if it crosses some minimum k-cut at most ℓ times; i.e., $E_T(S_1^*, \ldots, S_k^*) \leq \ell$. If $\ell = k - 1$, the minimum value possible, then we call it a tight T-tree.*

Theorem II.3 (Thorup [5], rephrased in Corollary 2.3 of [7]). *We can find a collection $\mathcal{T}$ of $\tilde{O}(k^3 m)$ trees such that there exists a $(2k - 2)$-T-tree in $\mathcal{T}$.*

Theorem II.4 (Lemma 2.4 of [7]). *There is an algorithm that takes as input a tree T such that $|E_T(S_1^*, \ldots, S_k^*)| \leq 2k - 2$, and produces a collection of $k^{O(k)} n^{k-1} \log n$ trees, such that one of the new trees T' satisfies $|E_{T'}(S_1^*, \ldots, S_k^*)| = k - 1$ w.h.p. The algorithm runs in time $k^{O(k)} n^{k-1} m \log n$.*

Combining Theorem II.3 and Theorem II.4, we obtain the following:

Corollary II.5. *We can find a collection $\mathcal{T}$ of $k^{O(k)} n^{k+O(1)}$ trees such that there exists a tight T-tree in $\mathcal{T}$.*

For the rest of the paper, only Definition II.2 and Corollary II.5 will be used.

III. ALGORITHM OUTLINE

In this section, we outline our main algorithm, assuming our two main technical results below. The former is proved in Section IV and the latter in Section V.

Theorem III.1. *Let G be an unweighted multigraph, let T be a tight T-tree of G, and let s be a parameter. There is an algorithm* TreeCut(G, T, λ) *with the following guarantee: if the minimum k-cut in G has size $\leq \lambda$, then* TreeCut *outputs a minimum k-cut of G. The running time of* TreeCut *is $k^{O(k)} \lambda^k n^{o(k)}$.*

Theorem III.2. *Let G be a simple graph with minimum degree $\delta > \omega(\max\{\alpha \log n, \alpha k\})$, and let $\alpha \geq 1$ be a parameter. Then, we can contract G into a (multi-)graph $\overline{G}$ such that:*

1) *Suppose the minimum k-cut has size $\leq \alpha\delta$ in G. Then, every nontrivial minimum k-cut is preserved in $\overline{G}$. That is, no edge of such a cut is contracted in $\overline{G}$.*
2) *$\overline{G}$ has $\tilde{O}(\alpha m/\delta)$ edges and $\tilde{O}(\alpha m/\delta^2)$ vertices.*

We will also use the following sparsification routine due to Nagamochi and Ibaraki below, for which we provide a quick proof for self-containment:

Theorem III.3 (Nagamochi-Ibaraki [26]). *Given a simple graph G and parameter λ, there is a polynomial-time algorithm* NISparsify(G, λ) *that computes a subgraph H with at most λn edges such that all k-cuts of size $\leq \lambda$ are preserved. More formally, all sets S with $|\partial_G S| \leq \lambda$ satisfy $|\partial_G S| = |\partial_H S|$.*

Proof: For $i = 1, 2, \ldots, \lambda$, let F_i be a maximal forest in $G \setminus \bigcup_{j<i} F_j$. Set $H := \bigcup_i F_i$, which can easily be computed in polynomial time. (We note that Nagamochi and Ibaraki [26] present a way to compute H in *linear* time, although we do not need this.) For any edge (u, v) in $G - H = G \setminus \bigcup_i F_i$, there must be a (u, v) path in each F_i, otherwise we would have added edge (u, v) to F_i. These λ paths, along with edge (u, v), imply that every cut that separates u and v has size $\geq \lambda + 1$. Therefore, u and v must lie in the same component of any k-cut of size $\leq \lambda$, so removing edge (u, v) cannot affect any such k-cut. ∎

Let us now describe our algorithm in pseudocode:

Algorithm 1 MinKCut$(G = (V, E), k)$

1: **if** $k = 1$ **then**   ▷ Base case $k = 1$
2:      **return** $\{V\}$
3: $\mathcal{S}^* \leftarrow \{S_1^0, \ldots, S_k^0\}$, an arbitrary initial k-cut   ▷ k-cuts will be represented as partitions of V of size k
4: **for** each $v \in V$ **do**
5:      $\mathcal{S}^1 \leftarrow \{v\} \cup$ MinKCut$(G - v, k - 1)$▷ Recursively call minimum $(k-1)$-cut
6:      If $\mathcal{S}^1$ is a better k-cut than $\mathcal{S}^*$, then set $\mathcal{S}^* \leftarrow \mathcal{S}^1$
7: **if** $\delta > \omega(\max\{k^2 \log n, k^3\})$ **then**   ▷ Assumption of Theorem III.2. δ is the minimum degree of G
8:      $H \leftarrow$ NISparsify$(G, k^2\delta)$ ▷ Nagamochi-Ibaraki sparsification: see Theorem III.3
9:      $G \leftarrow$ KT-Sparsification(H, k^2) ▷ Replace G with the KT-sparsification of G (Theorem III.2)
10: $\mathcal{T} \leftarrow$ TreePacking(G, k) ▷ Corollary II.5
11: **for** each tree $T \in \mathcal{T}$ **do**
12:      $\mathcal{S}^T \leftarrow$ TreeCut$(G, T, k^2\delta)$ ▷ Theorem III.1
13:      If $\mathcal{S}^T$ is a better k-cut than $\mathcal{S}^*$, then set $\mathcal{S}^* \leftarrow \mathcal{S}^T$
14: **return** $\mathcal{S}^*$

A. Correctness

We first state an easy claim from [7]. We then use it to bound the size of a nontrivial minimum k-cut by the minimum degree δ of the graph.

Claim III.4 (Claim 2.8 of [7]). *Given a set of $k + 1$ components $S_1, \ldots, S_{k+1}$ that partition V, we have*

$$|OPT| \leq \left(1 - \binom{k+1}{2}^{-1}\right)|E[S_1, \ldots, S_{k+1}]|.$$

Lemma III.5. *Suppose there exists a nontrivial minimum k-cut in the graph. Then, we have $|OPT| \leq k^2\delta$.*

Proof: Fix a nontrivial minimum k-cut $S_1^*, \ldots, S_k^*$. Applying Claim III.4 on $S_i := S_1^* \setminus \{v\}$ for $i \in [k]$ and $S_{k+1} = \{v\}$, we get

$$|OPT| < \left(1 - \frac{1}{k^2}\right)|E[S_1, \ldots, S_{k+1}]|$$
$$\leq \left(1 - \frac{1}{k^2}\right)(|E[S_1^*, \ldots, S_k^*]| + \delta)$$
$$= \left(1 - \frac{1}{k^2}\right)(|OPT| + \delta),$$

so $|OPT| \leq k^2(1 - k^2)\delta \leq k^2\delta$, as needed. ∎

Lemma III.6 (Correctness). MinKCut(G, k) *outputs a minimum k-cut of G.*

Proof: If there is a nontrivial minimum k-cut v, then consider the iteration of line 5 on the vertex v. By induction on k, we may assume that MinKCut$(\cdot, k - 1)$ outputs a minimum k-cut. Therefore, $\mathcal{S}^1$ will be an optimal k-cut on this iteration of line 5.

Otherwise, suppose there is no nontrivial minimum k-cut, so by Lemma III.5, $|OPT| \leq k^2\delta$. First, suppose that line 7 holds. Then, by Theorem III.3 with $s := k^2\delta$, the graph H in line 8 has the same minimum k-cuts as G and has at most $k^2\delta n$ edges. By Theorem III.2 with $\alpha := k^2$, every (nontrivial) minimum k-cut in H also exists in KT-Sparsification(H, k^2), which we set as our new G. By the correctness of TreePacking(G, k) (Corollary II.5) and TreeCut(G, T) (Theorem III.1), the algorithm computes an optimal k-cut. Otherwise, if line 7 does not hold, then our situation is even easier, since G does not change. ∎

B. Running time

Fact III.7. $(\log n)^{O(k)} \leq \max\{k^{O(k)}, n\}$.

Proof: If $k < \frac{\log n}{\log \log n}$, then $(\log n)^k \leq (\log n)^{\frac{\log n}{\log \log n}} = n$. Else $\log n \leq O(k \log k)$, and hence $(\log n)^k \leq O(k \log k)^k \leq k^{O(k)}$. ∎

Lemma III.8. MinKCut(G, k) *runs in $k^{O(k)}n^{(1+o(1))k}$ time.*

Proof: First, we bound the running time outside the recursive calls in line 5. Suppose first that line 7 does not hold. Then, $\delta \leq O(k^2 \log n + k^3)$, so

$$k^2\delta \leq O(k^4 \log n + k^5). \tag{1}$$

By Corollary II.5, the collection $\mathcal{T}$ has size $k^{O(k)}n^{(1+o(1))k}$. For each $T \in \mathcal{T}$, TreeCut$(G, T, k^2\delta)$ is executed in line 12,

which runs in $k^{O(k)}(k^2\delta)^k n^{o(k)}$ time by Theorem III.1. In total, this is $k^{O(k)}n^{(1+o(1))k}\cdot k^{O(k)}(k^2\delta)^k n^{o(k)}$ time, which is at most $k^{O(k)}(\log n)^k n^{(1+o(1))k}$ by (1). The $(\log n)^k$ factor is negligible by Fact III.7.

Otherwise, line 7 holds. By Theorem III.3 with $s := k^2\delta$, the graph H has at most $k^2\delta n$ edges, so by Theorem III.2, the graph G in line 7 has $\overline{n} = \tilde{O}(k^2(k^2\delta)/\delta^2) = \tilde{O}(k^{O(1)}n/\delta)$ vertices. By the same arguments as in the previous paragraph, the calls to TreeCut$(G,T,k^2\delta)$ take $k^{O(k)}(k^2\delta)^k \overline{n}^{(1+o(1))k}$ time, which is bounded by $k^{O(k)}(k^2\delta)^k(\tilde{O}(k^{O(1)}n/\delta))^{(1+o(1))k} = k^{O(k)}n^{(1+o(1))k}$.

Finally, we handle the recursive component of the algorithm. Fix an arbitrarily small $\epsilon > 0$, and fix constants c_1, c_2 such that the nonrecursive part takes time $k^{c_1 k}n^{(1+\epsilon)k+c_2}$ for all k. We prove by induction on k that the total algorithm takes time $(2k)^{c_1 k}n^{(1+\epsilon)k+c_2}$, with the trivial base case $k = 1$. For $k > 1$, the n recursive calls to $(k-1)$-cut take $n \cdot (2k)^{c_1(k-1)}n^{(1+\epsilon)(k-1)+c_2} = (2k)^{-c_1}(2k)^{c_1 k}n^{(1+\epsilon)k+c_2}$ time total. The nonrecursive part takes time $k^{c_1 k}n^{(1+\epsilon)k+c_2} = 2^{-c_1 k}(2k)^{c_1 k}n^{(1+\epsilon)k+c_2}$ by assumption. As long as $(2k)^{-c_1} + 2^{-c_1 k} \le 1$, which holds for any constant $c_1 \ge 1$, the sum of the two running times is at most $(2k)^{c_1 k}n^{(1+\epsilon)k+c_2}$, preserving the induction. Hence, the running time is $k^{O(k)}n^{(1+o(1))k}$. ∎

IV. ALGORITHM ON TIGHT T-TREES

In this section, we prove the running time guarantee of TreeCut(G,T,s) in line 12.

Theorem III.1. *Let G be an unweighted multigraph, let T be a tight* T-tree *of G, and let s be a parameter. There is an algorithm* TreeCut(G,T,λ) *with the following guarantee: if the minimum k-cut in G has size $\le \lambda$, then* TreeCut *outputs a minimum k-cut of G. The running time of* TreeCut *is $k^{O(k)}\lambda^k n^{o(k)}$.*

Since the minimum k-cut can be obtained by deleting $k-1$ edges of T and taking the connected components as the k-cut, our algorithm will pursue this route: it will look for the best $k-1$ edges of T to delete. Let $E_T^* := E_T[S_1^*,\ldots,S_k^*]$ be the optimal set of $(k-1)$ edges to delete.

First, observe that we can assume that for every edge (u,v) in T, the minimum (2-)cut that separates u and v has size $\le \lambda$. This is because if an edge (u,v) in T does not satisfy this property, then no minimum k-cut of size $\le \lambda$ can separate u and v, so we can contract u and v in T. Moreover, since s–t minimum cut is polynomial time solvable, the algorithm can detect which edges to contract.

Assumption IV.1. *For every edge (u,v) in T, the minimum (2-)cut that separates u and v has size $\le \lambda$.*

A. Restricted Case: Union of Branches

We first begin with an algorithm when the tree T is "spider-like", as discussed in Section I-A.

Theorem IV.2. *Let G be an unweighted multigraph, let T be a tight* T-tree *of G, and let λ be a parameter. Suppose in addition that:*

1) *We can root T at a vertex $r \in V(T)$ so that T is a disjoint union of maximal branches.*
2) *There is an optimal minimum k-cut $S_1^*,\ldots,S_k^*$ such that $E_T[S_1^*,\ldots,S_k^*]$ contains at most one edge from each maximal branch.*

Then, there is an algorithm TreeCut(G,T,s) *with the following guarantee: if the minimum k-cut in G has size $\le s$, then* TreeCut *outputs a minimum k-cut of G. The running time of* TreeCut *is $k^{O(k)}s^k n^{o(k)}$.*

In this section, we develop an algorithm to solve this restricted case. Throughout, we assume that the minimum k-cut of G is indeed at most λ, since otherwise, the algorithm can output anything.

Since the edges of E_T^* lie in distinct branches rooted at the same vertex, the child vertices of the edges in E_T^* are incomparable, and each subtree rooted at a child vertex is a component in $\mathcal{S}$. Without loss of generality, let S_k^* be the component containing the root of T. Let $v_1^*,\ldots,v_{k-1}^*$ be the child vertices of E_T^* such that component S_i^* is exactly $T(v_i^*)$. For each $i \in [k-1]$, consider the maximal branch containing v_i^*, and let u_i^* be the child of the root r that lies on this branch. Lastly, define $E' \subseteq E$ to be the edges whose endpoints are incomparable (i.e., they do not lie on a common branch).

Definition IV.3. *An edge $(u,v) \in E(T)$ is* partially preceded *by a vertex $x \in V(T)$ if either x precedes u or x precedes v (or both).*

Define the multigraph H as the graph obtained from starting with $G[\bigcup_{i\in[k-1]} V(T(v_i^*))]$ and contracting each vertex set $V(T(v_i^*))$ ($i \in [k-1]$) into a single vertex v_i^*, with self-loops removed. More precisely, H has vertex set $\{v_1^*,\ldots,v_{k-1}^*\}$, and its edge set is as follows: for each edge $e \in E'$ with endpoints in $T(v_i^*)$ and $T(v_j^*)$ ($i \ne j$), add that same edge e between v_i^* and v_j^*. Note that the two graphs share common vertices and edges; we make it this way to facilitate transitioning between the two graphs. Observe that $|E(H)| \le \lambda$, since every edge in $E(H)$ corresponds to an edge in $E_G(S_1^*,\ldots,S_{k-1}^*)$. Moreover, for an edge $e \in E(H)$ with endpoints v_i^*, v_j^*, edge e connects $T(v_i^*)$ and $T(v_j^*)$ in G; let $e|v_i^*$ and $e|v_j^*$ denote the endpoint of e in $T(v_i^*)$ and $T(v_j^*)$, respectively.

Lemma IV.4. *For each connected component C in H, there exists a spanning tree T_C of C satisfying the following property: Let U be the set of endpoints of edges in T_C (more formally, $U := \bigcup_{(u,v)\in E(T_C)}\{u,v\}$). Then, every edge e in C is partially preceded by some vertex in U (in the tree T).*

Proof: Fix a connected component C of H, and construct a weighted digraph H' as follows: for each edge

$e \in E(C)$ with endpoints v_i^*, v_j^*, add an arc (v_i^*, v_j^*) with weight equal to the depth of vertex $e|v_j^*$ in $T(v_j^*)$, and an arc (v_j^*, v_i^*) with weight equal to the depth of vertex $e|v_i^*$ in $T(v_i^*)$. Since C is connected, H' is strongly connected. Let $r \in V(H')$ be arbitrary, and let $A \subseteq E(H')$ be a minimum cost (out-)arborescence of H' rooted at r (so that r is the only vertex with no in-arcs in A). We claim that the tree $\overline{A}$ formed by un-directing every arc in A is our desired spanning tree T_C.

To prove this claim, let $e \in E(C)$ be arbitrary, with endpoints v_i^*, v_j^*, and assume without loss of generality that v_i^* does not precede v_j^* in the tree $\overline{A}$ rooted at r (otherwise we can swap v_i^* and v_j^*). Let $a \in E(H')$ be the arc originating from e in the direction (v_j^*, v_i^*), and let $a' \in H[C]$ be the in-arc of v_i^* in A (note that v_i^* cannot be the root since it does not precede v_j^*, and that it is possible that $a = a'$). Observe that $A \setminus a' \cup a$ is also an arborescence. Since A is the minimum cost arborescence, the weight of a' is at most the weight of a. Let e' be the edge originating from a'; we have $\mathrm{depth}(e'|v_i^*) \le \mathrm{depth}(e|v_i^*)$. It follows that the endpoint $e|v_i^*$ is preceded by $e'|v_i^* \in U$. ∎

1) Restricted Case: Algorithm: We first present the main steps our algorithm. Suppose for simplicity that H is connected; that is, there is only one connected component H. In fact, **we encourage the reader to assume that H is connected on their first reading**, since it simplifies the presentation while still preserving all the key insights. Our first insight is *color-coding* to mark out the spanning tree T_H guaranteed by Lemma IV.4. In particular, all the edges in the spanning tree should be colored one color (say, green), while all other edges in $\partial_G(T(u_i^*))$ should be colored a different color (say, red); see Figure 1. This color-coding process will succeed with probability roughly λ^{-k}, so we need to repeat it roughly λ^k many times. This is where we pay the λ^k multiplicative factor in the running time. Then, construct a graph with the maximal branches as vertices, where two vertices are connected by an edge if their corresponding branches have a green edge between them. Assuming that H is connected, one of these connected components corresponds exactly to the branches $T(u_1^*), \ldots, T(u_{k-1}^*)$. Finally, we iterate over the connected components C of size $k-1$ (one of which captures $T(u_1^*), \ldots, T(u_{k-1}^*)$) and, with the information of the green edges, compute an overestimate of the minimum possible k-cut formed by cutting one edge from each of the corresponding $k-1$ branches of C. The catch is that for the component containing $T(u_1^*), \ldots, T(u_{k-1}^*)$, this estimate will actually be *exact*. This ensures that the minimum k-cut is indeed returned.

The key insight in the algorithm is coloring the edges of this tree T_H, which serves two purposes. First, it allows the algorithm to figure out which $k-1$ branches contain T_H by computing the connected components as described above. Second, the edges of T_H partially precede
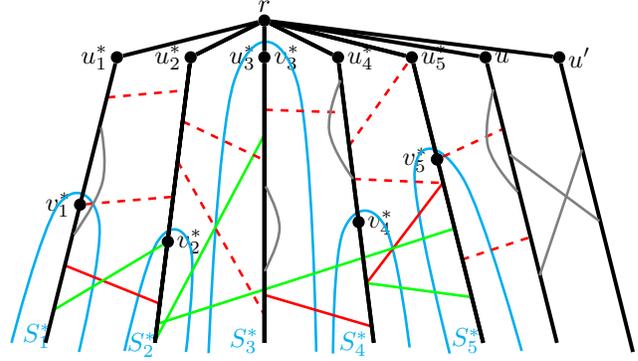

Figure 1. In this example, the graph H is connected, so there is only one component C_1^*. The green edges form T_1^* (see Condition IV.5); the (solid and dashed) red edges form $\bigcup_i \partial_G(T(u_i^*)) \setminus \bigcup_i T_i^*$ (see Condition IV.6); the solid red and green edges form $E[S_1^*, \ldots, S_{k-1}^*]$; the gray edges can be either red or green without affecting Conditions IV.5 or IV.6. (Note that $E[S_1^*, \ldots, S_5^*]$ is not *actually* the minimum 6-cut in the graph, but that is not the focus of this example.)

all edges in H, including all edges that appear *twice* in $\partial_G T(v_1^*), \ldots, \partial_G T(v_{k-1}^*)$. It turns out that these edges are the hardest to deal with, since they are the ones double-counted when merely summing up the boundaries $\partial_G T(v_1^*), \ldots, \partial_G T(v_{k-1}^*)$.[3] However, with the knowledge of T_H, any edge partially preceded by an endpoint in T_H is cut for sure, and this includes all double-counted edges! And once the double-counted edges are dealt with, we can simply treat the $k-1$ branches independently.[4]

It turns out that both of these properties—finding the $k-1$ branches and dealing with the double-counted edges—can each be done separately with a λ^k multiplicative factor, each with standard color-coding techniques.[5] Focusing on this special tree T_H is what enables us to achieve both with just one λ^k factor.

We now proceed to the algorithm. Let the connected components of H be $C_1^*, \ldots, C_z^*$ for $z \le V(H) = r$, ordered in an arbitrary order, and let T_i^* be the spanning tree for C_i^* promised by Lemma IV.4. The algorithm now colors the edges of E' red and green such that the following two conditions hold:

Condition IV.5. *For each C_i^*, all edges in T_i^* are colored* green.

Condition IV.6. *All edges in $\bigcup_i \partial_G(T(u_i^*)) \setminus \bigcup_i T_i^*$ are colored* red.

---

[3]Indeed, a reader familiar with the k-Partial Vertex Cover problem may be familiar with the difficulty of double-counting, an complication that alone justifies why the problem is $W[1]$-hard.

[4]An illustrative analogy for the minimum/maximum k-Partial Vertex Cover problem is that if we somehow knew that there were *no* edges between the k optimal vertices to select, then the problem becomes easy: simply output the k vertices of minimum/maximum degree.

[5]Indeed, the $(1 + \epsilon)$-approximate k-cut algorithm time of [7] can be adapted this way to solve exact minimum k-cut.

We do this with the following color-coding procedure: color each edge red with probability $1 - 1/\lambda$ and green with probability $1/\lambda$. For each i, $\partial_G(T(u_i^*))$ is the (2-)cut in G formed by removing the parent edge of u_i^*, so by Assumption IV.1, we have $|\partial_G(T(u_i^*))| \leq \lambda$. Therefore, the success probability is at least

$$
\left(1 - \frac{1}{\lambda}\right)^{|\bigcup_i \partial_G(T(u_i^*))|} \cdot \prod_{i \in [z]} \left(\frac{1}{\lambda}\right)^{|E(T_i^*)|}
$$
$$
\geq \left(1 - \frac{1}{\lambda}\right)^{k\lambda} \cdot \left(\frac{1}{\lambda}\right)^{k}
$$
$$
= 2^{-O(k)} \lambda^{-k}. \tag{2}
$$

We describe this algorithm and its guarantees succinctly as follows:

Algorithm IV.7. *Color each edge in E' red with probability $1 - 1/\lambda$ and green with probability $1/\lambda$. With probability $2^{-O(k)}\lambda^{-k}$, Conditions IV.5 and IV.6 hold.*

Next, build a graph whose vertices are the children of r, and for every two children u, u' of r, connect them by an edge if there is a green edge between $T(u)$ and $T(u')$. Consider all (maximal) connected components in this graph; for each one, add its set U of vertices into a collection $\mathcal{U}$. Observe that if Conditions IV.5 and IV.6 hold, then for each connected component C_i^*, there exists a set $U_i^* \in \mathcal{U}$ of size $|V(C_i^*)|$ such that each vertex v_j^* in C_i^* belongs on a (different, unique) branch $T(u)$ ($u \in U_i^*$). Therefore, with a success probability of $2^{-O(k)}\lambda^{-k}$, we can assume the following:

Assumption IV.8. *After running Algorithm IV.7, Conditions IV.5 and IV.6 hold, and for each connected component C_i^*, there exists a set $U_i^* \in \mathcal{U}$ of size $|V(C_i^*)|$ such that each vertex v_j^* in C_i^* belongs on a (different, unique) branch $T(u)$ ($u \in U$). (This assumption holds with probability at least $2^{-O(k)}\lambda^{-k}$.)*

We will not assume Assumption IV.8 unconditionally throughout the remainder of this section; rather, we will explicitly state where Assumption IV.8 is assumed. Hopefully, this provides more intuition as to how Assumption IV.8 is used.

2) Restricted case: Processing the sets $U \in \mathcal{U}$: In this section, we process the sets $U \in \mathcal{U}$, solving a certain cut problem on a specific graph for each U (in polynomial time). Recall that if H is connected, then this is the step where we should compute the exact value $|OPT|$ for the corresponding vertex set $U_1^* = V(H)$. In general, for each set U_i^* corresponding to a component C_i^* of H with vertices $v_{i_1}^*, \ldots, v_{i_\ell}^* \in V_T^*$, we want to compute the exact value $|E[S_{i_1}^*, \ldots, S_{i_\ell}^*]|$, whose sum over all C_i^* will turn out to equal $|OPT|$. Moreover, for every other set $A \in \mathcal{A}$, we want to make sure we compute some sort of *overestimate*,

so that these extraneous sets do not mislead the algorithm.

For each $U \in \mathcal{U}$, define $\mathsf{MinElts}(U) := \left(\bigcup_{u \in U} \bigcup_{(v,v') \in \partial T(u_i)} \mathsf{green}(\{v, v'\} \cap T(u))\right)\downarrow$ as the minimal elements on $\bigcup_u T(u)$ of the set of endpoints of green edges in $\bigcup_i \partial T(u_i)$, which is clearly pairwise incomparable. We now define the *minimum ancestor cut* problem:

Definition IV.9 (Minimum ancestor cut). *Fix some set $U = \{u_1, \ldots, u_\ell\} \in \mathcal{U}$, and let $\mathsf{MinElts}(U) = \{s_1, \ldots, s_\ell\}$ where $s_i \in V(T(u_i))$. The* minimum ancestor cut *is the following problem: For each $i \in [\ell]$, select one edge in the branch from s_i to r, and consider the $(\ell + 1)$-cut in G formed by removing these selected edges in T. We want to compute the $(\ell + 1)$-cut of minimum size, denoted $\mathsf{MinAncCut}(U)$.*

Why are minimum ancestor cuts relevant? We first show that for components C_i^* of H, $\mathsf{MinAncCut}(U_i^*)$ has a close connection with $S_1^*, \ldots, S_k^*$.

Claim IV.10. *Assuming Assumption IV.8, for each $U_i^* = \{u_{i_1}^*, \ldots, u_{i_\ell}^*\}$, we have $|\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]| = |\mathsf{MinAncCut}(U_i^*)|$. Moreover, $\sum_{U_i^*} |\mathsf{MinAncCut}(U_i^*)| = OPT$.*

Proof: Recall that the corresponding component C_i^* has vertices $v_{i_1}^*, \ldots, v_{i_\ell}^*$. First, we claim that an ancestor cut of size $|\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]|$ is achievable: simply cut the parent edges of $v_{i_1}^*, \ldots, v_{i_\ell}^*$. We now show that this cut is indeed an ancestor cut. By Conditions IV.5 and IV.6, the only green edges in any $\partial T(u_{i_j}^*)$ lie in T_i^*. Since $T_i^* \subseteq H$, and since any edge in H between $T(u_{i_j}^*)$ and $T(u_{i_{j'}}^*)$ must lie between $T(v_{i_j}^*)$ and $T(v_{i_{j'}}^*)$, we have that $v_{i_j}^*$ precedes any endpoint of a green edge on $T(u_{i_j}^*)$. Therefore, the parent edges of $v_{i_j}^*$ lie on the branches from endpoints in $\mathsf{MinElts}(U_i^*)$ to r, so our cut is a valid ancestor cut.

We now show that no better ancestor cut is possible; suppose otherwise. Then, let $S_{i_1}, \ldots, S_{i_\ell}$ be the components not containing r in $\mathsf{MinAncCut}(U_i^*)$. First, merge the components $S_{i_1}^*, \ldots, S_{i_\ell}^*$ together with S_k^*. We claim that the number of cut edges drops by exactly $|\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]|$: since C_i^* is a component of H, all edges in $\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]$ have endpoints in $S_{i_1}^*, \ldots, S_{i_\ell}^*$ or S_k^* (instead of outside these components). Next, split $S_{i_1}^* \cup \ldots \cup S_{i_\ell}^* \cup S_k^*$ into $S_{i_1}, \ldots, S_{i_\ell}$ and the remaining component; this increases the number of cut edges by at most $|\mathsf{MinAncCut}(U_i^*)|$. We arrive at a k-cut of smaller size, contradicting the choice of $S_1^*, \ldots, S_k^*$.

Finally, to prove that $\sum_{U_i^*} |\mathsf{MinAncCut}(U_i^*)| = |\partial_G[S_1^*, \ldots, S_k^*]|$, observe that over different components C_j^* of H with vertices $v_{j_1}^*, \ldots, v_{j_\ell}^* \in V_T^*$, the edges $\partial_G[S_{j_1}^*, \ldots, S_{j_\ell}^*]$ are disjoint over distinct C_j^*. (This is because if two distinct $C_j^*, C_{j'}^*$ shared an edge in their respective edge sets $\partial_G[\cdot, \ldots, \cdot]$, then the components $C_j^*, C_{j'}^*$ should have become a single connected component.) Therefore, $|\partial_G[S_1^*, \ldots, S_k^*]|$ equals the sum of the $|\partial_G[S_{j_1}^*, \ldots, S_{j_\ell}^*]|$ values, which equals the sum of the
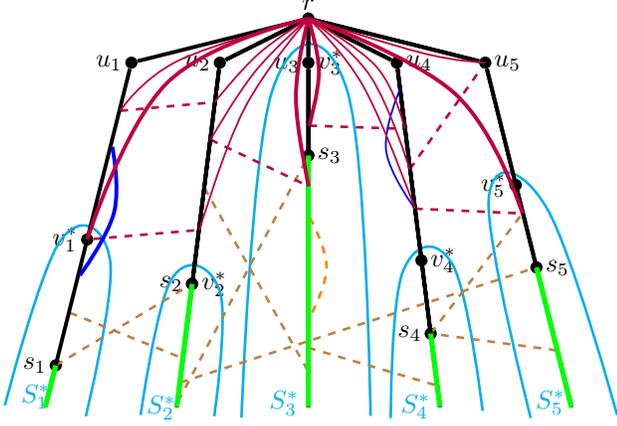
Figure 2. The graph G' for U_1^* from Figure 1. The green branches are the vertices in $T(\mathsf{MinElts}(U_1^*))$. The dotted orange, brown, blue, and purple edges are the edges in E considered in step (a), (b), (c), and (d), respectively. The solid blue and purple edges are the edges in G'. The bold blue and purple edges are the ones cut in $\mathsf{MinAncCut}(U_1^*)$, which is also the cut which produces $S_1^*, \ldots, S_5^*$.

$|\mathsf{MinAncCut}(U_i^*)|$ values. ∎

Therefore, for each U_i^*, we may define $\mathsf{MinAncCut}(U_i^*)$ as not just any arbitrary minimum ancestor cut, but the specific one formed by cutting the edges $E_T[S_{i_1}^*, \ldots, S_{i_\ell}^*]$. (By Claim IV.10, this is a minimum ancestor cut.)

We now compute a function $f(U)$ for each set $U \in \mathcal{U}$. For each U_i^*, we want $f(U_i^*) = |\mathsf{MinAncCut}(U_i^*)|$ so that it exactly captures the contribution of U_i^* to the minimum k-cut. (We will need Assumption IV.8 to achieve this.) For any other U, we want $f(U) \geq |\mathsf{MinAncCut}(U)|$. Lastly, we want $f(U)$ to be computable in polynomial time.

Fix a set $U = \{u_1, \ldots, u_\ell\}$ and let $\mathsf{MinElts}(U) = \{s_1, \ldots, s_\ell\}$ where $s_i \in V(T(u_i))$. To compute $f(U)$, we construct the following multigraph G' on the vertices $(\bigcup V(T(u_i))) \cup r$ (see Figure 2):

(a) For each edge in $E \setminus E'$ [6] with both endpoints in $T(\mathsf{MinElts}(U))$, do nothing: these edges are not cut in any ancestor cut.

(b) For each edge in E' with at least one endpoint in $T(\mathsf{MinElts}(U))$, do nothing: these edges are *always* cut in an ancestor cut, and we will account for these edges separately.

(c) For each edge in $E \setminus E'$ with at most one endpoint in $T(\mathsf{MinElts}(U))$, add it to G'.

(d) For each edge $(u, v) \in E'$ with both endpoints not in $T(\mathsf{MinElts}(U))$, add edges (u, r) and (v, r) to G'.

Note that all edges in G' have both their endpoints in the same branch. Now, for each $i \in [\ell]$, compute the vertex $t_i \in T(u_i)$ that minimizes $|\partial_{G'} T(t_i)|$. Take the sum of the costs of these ℓ cuts, and finally, add the number of edges considered in step (b) to this sum. The final value is $f(U)$.

---
[6] Recall that $E' \subseteq E$ is the edges with incomparable endpoints in T.

Algorithm IV.11. *For each set $U \in \mathcal{U}$, construct the graph G' as above, and compute the vertex $t_i \in T(u_i)$ that minimizes $|\partial_{G'} T(t_i)|$. Take the sum of the costs of these cuts, and add the number of edges considered in step (b) to this sum. Let $f(U)$ be the final value.*

Let us now explain the intuition of the construction of G', relating it to $\mathsf{MinAncCut}(U)$. First, every edge considered in (a) has both endpoints on the same maximal branch, both of which are below s_i on the appropriate branch, so it is never cut in an ancestor cut and can therefore be ignored. Every edge considered in (b) is always cut in an ancestor cut: if the edge is (u, v) with $u \in T(s_i)$, then u and s_i will always belong in the same component in the ancestor cut, but never v and s_i because v is on a different branch. Every edge in (c) can either be cut or not cut depending on the specific ancestor cut, and it is easy to see that it is included in $\bigcup \partial_{G'} T(t_i)$ iff it is cut in the ancestor cut that cuts the parent edges of each t_i. Finally, every edge in (d) splits into two edges, possibly adding two edges to a cut in G'. Indeed, if neither endpoint of edge (u, v) in (d) is in r's side of the cut, then both corresponding edges in G' are cut. This is where the overestimate $f(U) \geq |\mathsf{MinAncCut}(U)|$ will come from. However, assuming Assumption IV.8, there cannot be any overestimate for each U_i^*: by definition of T_i^*, every edge in $E(S_1^*, \ldots, S_{k-1}^*)$ with an endpoint (equivalently, both endpoints) in the maximal branches containing vertices in $\mathsf{MinElts}(U_i^*)$ is partially preceded by $\mathsf{MinElts}(U_i^*)$. Therefore, we have equality for each U_i^*: $f(U_i^*) = |\mathsf{MinAncCut}(U_i^*)|$.

We now formalize our intuition. Define the following natural correspondence between ancestor cuts and the "G'-cuts" $\bigcup_i \partial_{G'} T(t_i)$ over the choices of $\{t_i : i \in [\ell]\}$: two correspond to each other if the edges in the ancestor cut are the same as the parent edges of t_i.

Lemma IV.12. *For two corresponding cuts, the size of the ancestor cut is at most the size of the G'-cut plus the number of edges in step (b).*

Proof: The proof essentially following the intuition paragraph above. An edge in (a) contributes 0 to the sizes of both cuts, and an edge in (b) contributes 1 to the ancestor cut and 1 to the number of edges in step (b). An edge in (c) contributes an equal amount to both cuts. Finally, if an edge in (d) contributes 0 to the ancestor cut, then both of its endpoints belong to the component containing r, so in the G'-cut, neither of its endpoints is in their respective $T(t_i)$; hence, it also contributes 0 to the G'-cut. If the edge contributes 1 to the ancestor cut, then it contributes either 1 or 2 to the G'-cut depending on whether exactly one endpoint belongs to r's component in the ancestor cut (1 to G'-cut), or no endpoints belong to it (2 to G'-cut). ∎

Lemma IV.13. *Assuming Assumption IV.8, for each U_i^*, $\mathsf{MinAncCut}(U_i^*)$ has size exactly $f(U_i^*)$.*

Proof: In the proof of Lemma IV.12, the only potential source of inequality is in (d): an edge with no endpoints in r's component of $\mathsf{MinAncCut}(U_i^*)$ contributes 1 to $|\mathsf{MinAncCut}(U_i^*)|$ and 2 to the corresponding G'-cut. If such an edge e existed, then it must be in $\partial_G[S_{i_1}^*, \dots, S_{i_\ell}^*]$ where $V(C_i^*) := \{v_{i_1}^*, \dots, v_{i_\ell}^*\}$, which means e is in component C_i^*. Also, neither of its endpoints is preceded by a vertex in $\mathsf{MinElts}(U_i^*)$, which means neither of its endpoints is preceded by any endpoint in G of any edge in C_i^*, and therefore any endpoint in G of any edge in T_i^* as well. We thus have an edge e in C_i^* not partially preceded by any endpoint in T_i^*, contradicting the definition of T_i^* (see beginning of Section IV-A1). Therefore, no such edge exists, and we have equality. ∎

We run Algorithm IV.11, computing the value $f(U)$ for each set $U \in \mathcal{U}$ in polynomial time. Finally, the algorithm seeks to minimize

$$\min_{\substack{U_1, \dots, U_\ell \\ \sum_i |U_i| = k-1}} \sum_{i=1}^{\ell} f(U_i). \tag{3}$$

The expression (3) can be formulated as a knapsack problem with small, integral costs, which can easily be solved in polynomial time.

Since the branches in any two distinct U, U' are disjoint, the sum $\sum_i f(U_i)$ for any $U_1, \dots, U_\ell$ is a $(\sum_i |U_i|)$-cut. This fact, along with Lemma IV.12, proves that (3) is at least OPT. Furthermore, assuming Assumption IV.8, OPT can be achieved by Lemma IV.13. Thus, as long as Assumption IV.8 is true, (3) is exactly OPT.

Algorithm IV.14. *Compute (3) in polynomial time by formulating it as a knapsack problem. Assuming Assumption IV.8, the result is exactly OPT.*

Lemma IV.15. *(3) is always at least OPT.*

Proof: Let $U_1, \dots, U_\ell$ be the sets achieving the minimum in (3). For each U_i, consider the best G'-cut in the graph G' constructed for U_i. By Lemma IV.12, the corresponding ancestor cut has size at most $f(U_i)$. Also, the ancestor cut has one edge sharing a maximal branch with each vertex in $\mathsf{MinElts}(U_i)$. Since $\mathsf{MinElts}(U_i)$ and $\mathsf{MinElts}(U_j)$ lie on different branches for $i \neq j$, if we take the union of the ancestor cuts over all U_i, then every maximal branch is cut at most once, so we cut one edge from each of exactly $k-1$ maximal branches. This union is a k-cut of cost at most $\sum_i f(U_i)$ (it could be smaller if an edge appears twice in the union, once from each side), which means that $\sum_i f(U_i) \geq OPT$. ∎

Lemma IV.16. *Assuming Assumption IV.8, (3) equals OPT.*

Proof: By Lemma IV.15, it suffices to show that (3) is at most OPT (assuming Assumption IV.8). By Lemma IV.13, $|\mathsf{MinAncCut}(U_i^*)| = f(U_i^*)$ for each i. Finally, by Claim IV.10, $\sum_{U_i^*} |\mathsf{MinAncCut}(U_i^*)| = OPT$.

Thus, $\sum_i f(U_i^*) = OPT$, and the minimum in (3) can only be smaller. ∎

Thus, by Lemma IV.16, the algorithm below outputs an optimum k-cut w.h.p., proving THM-MAIN.

Algorithm IV.17. *For $O(2^{O(k)} \lambda^k \log n)$ repetitions, run Algorithms IV.7, IV.11 and IV.14, and output the minimum value of (3) ever computed.*

B. General Case

In this section, we present our general algorithm, proving Theorem III.1, restated below.

Theorem III.1. *Let G be an unweighted multigraph, let T be a tight T-tree of G, and let s be a parameter. There is an algorithm $\mathsf{TreeCut}(G, T, \lambda)$ with the following guarantee: if the minimum k-cut in G has size $\leq \lambda$, then $\mathsf{TreeCut}$ outputs a minimum k-cut of G. The running time of $\mathsf{TreeCut}$ is $k^{O(k)} \lambda^k n^{o(k)}$.*

Before we begin, let us briefly describe the differences of the general setting and state the techniques we will use to overcome the new difficulties.

1) The first difference is that in general, the $k-1$ edges in $E_T[S_1^*, \dots, S_k^*]$ may not be incomparable. For example, in the extreme case, they can all lie on a single maximal branch. We resolve this issue with *dynamic programming* on the tree, in a similar fashion to Section 2.3.2 of [7]. At a high level, we only focus on the *minimal* vertices, which are incomparable, and capture the remaining vertices through dynamic programming.

2) The second difference is that the tree T is no longer a union of disjoint maximal branches. However, we still want to define a suitable ordering on the endpoints of the relevant edges, so that we can define a similar notion of partial precedence and use Lemma IV.4. Intuitively, we want a set of disjoint branches, one containing each minimal vertex, such that a variant of Lemma IV.4 still holds, so that we can set up a similar minimum s–t cut problem. We handle this issue with *heavy-light decomposition*, a well-known routine that breaks up a tree into long chains, combined with color-coding as before, since we do not know beforehand which chains are useful for us.

1) General Case: Algorithm: We will perform dynamic programming on the tree T, rooted at an arbitrary vertex r_0.[7] We define the dynamic programming states as follows:

Definition IV.18 (DP State). *For vertex $x \in T$ and integer $k' \in [0, k-1]$, define $\mathsf{State}(x, k')$ as the minimum number of edges cut in $G[V(T(x))]$ over all partitions of $V(T(x))$*

---

[7]In this section, we will free up variable r to be used as an integer, to be more consistent with the variable choice in Section 2.3.2 of [7].

obtained by cutting $k' - 1$ edges from $T(x)$. More formally,

$$\mathsf{State}(x, k') := \min_{S_1, \ldots, S_{k'}} |E_{G[V(T(x))]}[S_1, \ldots, S_{k'}], \quad (4)$$

where the minimum is over all partitions $S_1, \ldots, S_{k'}$ of $V(T(x))$ satisfying $|E_{T(x)}[S_1, \ldots, S_{k'}]| = k' - 1$. If there is no valid partition $S_1, \ldots, S_{k'}$, then $\mathsf{State}(x, k') = \infty$.

Observation IV.19. $|OPT| = \mathsf{State}(r, k - 1)$.

For the rest of this section, we will only be concerned with computing the actual *value* $\mathsf{State}(r, k - 1) = |OPT|$. The k-cut that achieves this value can be recovered from the dynamic program using standard backtracking procedures (at no asymptotic increase in running time).

The base cases are:

1) $\mathsf{State}(x, 0) = 0$ for all vertices $x \in V(T)$, and
2) $\mathsf{State}(x, k') = \infty$ for all leaves x and integer $k' \in [1, k - 1]$.

Fix a non-leaf vertex x, and assume that the values $\mathsf{State}(v, s)$ have already been computed for all $v \in T(x) \setminus \{x\}$ [8] and integer $k' \in [0, k - 1]$. We seek to compute the states $\mathsf{State}(x, k')$ for $k' \in [0, k - 1]$.

We can easily detect whether or not $\mathsf{State}(x, k') = \infty$: it is ∞ iff there are less than $k' - 1$ edges in $T(x)$. Therefore, let us also assume that $\mathsf{State}(x, k') < \infty$. Consider the components that achieve the minimum in (4), as well as the $k' - 1$ edges in $T(x)$ cut by those components. Let the children of these $k' - 1$ edges be $v_1^*, \ldots, v_{k'-1}^*$. Furthermore, suppose that they are ordered so that $\{v_1^*, \ldots, v_{k'-1}^*\}\!\downarrow = \{v_1^*, \ldots, v_r^*\}$ for some $r \in [k' - 1]$. For each $i \in [r]$, let $k_i^* := |\{v_1^*, \ldots, v_{k'-1}^*\} \cap V(T(v_i^*))| \in [k' - 1]$ be the number of these vertices preceded by v_i^*, so that $\sum_{i=1}^{r} k_i^* = k' - 1$.

Definition IV.20 (HLD). *A heavy-light decomposition (HLD) of a tree T is a partition $\mathcal{B}$ of the edges of T into disjoint branches, such that for each vertex $v \in V(T)$, the branch from v to the root x of T shares edges with at most $O(\log n)$ branches in $\mathcal{B}$.*

Fact IV.21. *For any tree, a HLD of the tree exists and can be computed in linear time.*

Fix a HLD of $T(x)$ with branches $\mathcal{B}$. For a vertex $v \in V(T(x)) \setminus \{x\}$, define $B(v)$ to be the branch in $\mathcal{B}$ containing the parent edge of v. For convenience, define $B_i^* := B(v_i^*)$ for $i \in [r]$; here, r is now an integer, not the root of $T(x)$, since x is now that root. Note that since $v_1^*, \ldots, v_r^*$ are pairwise incomparable, the branches $B_1^*, \ldots, B_r^*$ are always distinct. For each branch $B \in \mathcal{B}$, define $\mathsf{subroot}(B)$ to be the child of the root of B that lies on B (hence *sub*root). We now focus on the "minimal" branches B_i^*, formalized as follows: Let $q := |\{\mathsf{subroot}(B_i^*) : i \in [r]\}\!\downarrow|$,

[8] For the rest of this section, the reader may assume that $T = T(x)$ for convenience. Every time we refer to $T(v)$, we always have $v \in T(x)$, so $T(v) = (T(x))(v)$, but the latter is more cumbersome to write.

and order $v_1^*, \ldots, v_r^*$ so that $\{\mathsf{subroot}(B_i^*) : i \in [q]\} = \{\mathsf{subroot}(B_i^*) : i \in [r]\}\!\downarrow$. In particular, the vertices $\mathsf{subroot}(B_i^*)$ for $i \in [q]$ are incomparable. Finally, for each $i \in [r]$, define $u_i^* := \mathsf{subroot}(B_i^*)$.

This time, define $E' \subseteq E$ as the edges whose endpoints u, v satisfy the property that $\mathsf{subroot}(B(u))$ and $\mathsf{subroot}(B(v))$ are incomparable. Define H to be the graph on $\{v_1^*, \ldots, v_r^*\}$ with edge set as follows: for each edge $e \in E'$ with endpoints in $T(u_i^*)$ and $T(u_j^*)$ ($i \neq j$, $i, j \in [r]$), add that same edge e between v_i^* and v_j^*. The main difference of H compared to Section IV-A is that this H does not include edges between v_i^* and v_j^* in a common minimal branch B_h^*, $h \in [q]$.

For each $i \in [r]$ and vertex $v \in T(v_i^*)$, let $b(v)$ be the first vertex in B_i^* on the branch from v to r. Clearly, such a vertex always exists and is inside $T(v_i^*)$. For a vertex set $U \subseteq \bigcup_i T(v_i^*)$, define $b(U) := \bigcup_{u \in U} b(u)$. We now have a lemma resembling Lemma IV.4 from Section IV-A:

Lemma IV.22. *For each connected component C in H, there exists a spanning tree T_C of C satisfying the following property: Let U be the set of endpoints of edges in T_C. Then, every edge in C is partially preceded by some vertex in $b(U)$.*

Proof: The proof essentially follows from applying Lemma IV.4 on an appropriately chosen graph. Construct the graph G_b as follows: the vertex set is $\bigcup_i B_i^*$, and for each edge in C with endpoints u, v in G, add that edge with endpoints $b(u), b(v)$ in G_b. Note that if we contract each vertex set $V(B_i^*)$ in G_b, we get exactly H. Define a tree T_b by taking the branches $B_1^*, \ldots, B_r^*$ and adding a root r_b of T_b connected to the root of each B_i^*. Apply Lemma IV.4 on the graph $G_b \cup T_b$ (we take union with T_b only to include the root r_b), which gives us a spanning tree T_C. Let U be the endpoints of T_C in G, which means that the set of endpoints in G_b is exactly $b(U)$. By the guarantee of Lemma IV.4 and the definition of partially precede Definition IV.3, every edge e in C with endpoints u, v in G has one of its endpoints $b(u), b(v)$ preceded by a vertex in $b(U)$ in T_b, and hence also in $T(x)$. Since $b(u)$ always precedes u in $T(x)$ and the same for $b(v)$ and v, one of u, v is preceded by a vertex in $b(U)$ in $T(x)$. ∎

As in Section IV-A, let the connected components of H be $C_1^*, \ldots, C_z^*$ for $z \leq V(H) = r$, ordered in an arbitrary order, and let T_i^* be the spanning tree for C_i^* promised by Lemma IV.4. We first color the edges of E' identically to Section IV-A1:

Condition IV.23 (Same as Condition IV.5). *For each C_i^*, all edges in T_i^* are colored* green.

Condition IV.24 (Same as Condition IV.6). *All edges in $\bigcup_{i \in [q]} \partial_G(T(u_i^*)) \setminus \bigcup_i T_i^*$ are colored* red.

Algorithm IV.25 (Same as Algorithm IV.7). *Color each edge in E'* red *with probability $1 - 1/\lambda$ and* green
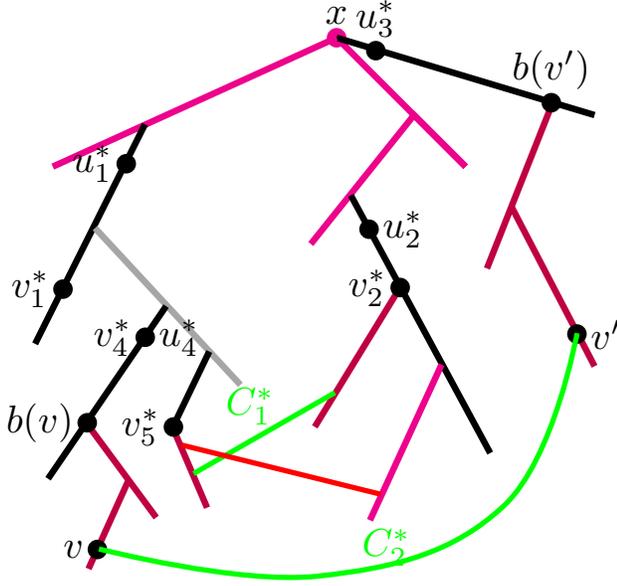
Figure 3. An example with $r = 6$ and $q = 3$. There are two components C_1^*, C_2^*, and one component C_1^+. For Condition IV.26 to hold, none of the black branches can be contracted. For Condition IV.27 to hold, the purple branches must be contracted. For Condition IV.28 to hold, the magenta branches must be contracted. Whether the gray branch is contracted or not does not affect the three conditions.

with probability $1/\lambda$. With probability $2^{-O(k)}\lambda^{-k}$, Conditions IV.23 and IV.24 hold.

Unlike Section IV-A1, we do not build the graph on the children of the root yet. This time, the algorithm now contracts every branch of the HLD independently with probability $1/\log n$. We would like the following three conditions to hold (see Figure 3):

Condition IV.26. *For each $i \in [r]$, B_i^* is not contracted.*

Condition IV.27. *For each C_i^* and each edge e in T_i^* with endpoints u, v in G, all branches in $\mathcal{B}$ intersecting the branch from u to $b(u)$ are contracted, and the same holds for the branch from v to $b(v)$.*

Condition IV.28. *For all $i \in [q]$, all branches in $\mathcal{B} \setminus \{B_i^*\}$ intersecting the branch from v_i^* to x are contracted.*

Clearly, Condition IV.26 holds with probability exactly $1/(\log n)^r$, which is negligible (see Fact III.7). We now claim the following for Condition IV.27:

Claim IV.29. *Conditioning on the event that Condition IV.26 holds, Condition IV.27 is true with probability $2^{-O(r)}$.*

Proof: Every T_i^* has $O(|V(C)|)$ endpoints in G, and for each such endpoint v, at most $O(\log n)$ branches in $\mathcal{B}$ intersect the branch from v to $b(v)$ by the guarantee of HLD. Over all T_i^*, this is $\sum_C O(|V(C)|) \cdot O(\log n) = O(r \log n)$

many branches. We now show that none of these branches in $\mathcal{B}$ can be B_j^* for some $j \in [r]$.

Fix $i \in [r]$, and consider an endpoint $v \in T_i^*$. First, suppose for contradiction that the branch from v to $b(v)$ intersects the branch B_i^*. Let (u', v') be an edge in B_i^* on the branch from v to $b(v)$, with v' as the child of u'. Since v' lies on the branch from v to $b(v)$, we have that $b(v)$ precedes v'. But then v' is a vertex on B_i^* encountered before $b(v)$ on the branch from v to x, contradicting the choice of $b(v)$.

Next, suppose for contradiction that the branch from v to $b(v)$ intersects the branch B_j^* for $i \neq j$. Extend the branch from v to $b(v)$ into the branch from v to v_i^*, which by assumption contains a vertex in B_j^*. First, if v_i^* is on B_j^*, then both v_i^* and v_j^* lie on a common branch B_j^*, contradicting the fact that they are incomparable. Otherwise, the path from v to v_i^* must travel beyond B_j^*, so every vertex in B_j^* is preceded by v_i^*. In particular, $v_j^* \in V(B_j^*)$ is preceded by v_i^*, again contradicting the fact that they are incomparable.

Therefore, even if we condition on Condition IV.26, none of the $O(r \log n)$ relevant branches are automatically contracted (so that the probability of success is not automatically 0). Thus, the probability that none of the $O(r \log n)$ branches are contracted is $(1 - 1/\log n)^{O(r \log n)} = 2^{-O(r)}$. ∎

Finally, conditioned on Conditions IV.26 and IV.27, Condition IV.28 holds with probability at least $(1 - 1/\log n)^{O(r \log n)} = 2^{-O(r)}$, since by the HLD property, there are $O(\log n)$ many branches that still need to be contracted for each v_i^* in Condition IV.28. Thus, we have the following:

Algorithm IV.30. *Compute a HLD of $T(x)$ into branches $\mathcal{B}$, and contract each branch with probability $1/\log n$. With probability at least $1/(\log n)^k \cdot 2^{-O(k)}$, Conditions IV.26 to IV.28 hold. Let T' be the resulting tree; note that x is still the root of T'.*

Define T' as in Algorithm IV.30; we have the observation below:

Observation IV.31. *Assuming Condition IV.28, all minimal branches $B_1^*, \ldots B_q^*$ now have x as their root.*

Moreover, the following observation follows immediately from Lemma IV.22:

Observation IV.32. *Assume Condition IV.27. Fix any component C_i^*, and let U be the set of endpoints in T' of edges in T_i^*. Then, every edge in C_i^* is partially preceded by some vertex in U (in the tree T').*

We now define the set $\mathcal{U}$ similarly to Section IV-A2, except this time using the tree T'. Build a graph whose vertices are the children of x in T', and for every two children u, v of x, connect them by an edge if there is a green edge between $T'(u)$ and $T'(v)$. For each (maximal) connected component in this graph, add its set U of vertices

into a collection $\mathcal{U}$.

Next, define the following graph H'. Its vertex set is $\{v_1^*, \ldots, v_q^*\}$, and it is obtained from H as follows: for each $i \in [q]$, contract into v_i^* all vertices v_j^* ($j \in [r]$) such that u_j^* is preceded by u_i^*. For each connected component C of H', let C^+ be the set of vertices in H contracted to a vertex in C. Every C^+ is a union of some connected components C_i^* of H; let $C_1^+, \ldots, C_y^+$ be all such vertex sets C^+.

Lemma IV.33. *If Conditions IV.23, IV.24 and IV.28 hold, then for each C_i^+, there exists a set $U_i^+ \in \mathcal{U}$ whose subtrees contain precisely all vertices v_i^* in C_i^+ (and no more).*

Proof: Consider a vertex set C_i^+ whose vertices get contracted into $v_{i_1}^*, \ldots, v_{i_\ell}^*$ in H'. For each vertex v_h^* in C_i^+, there is a vertex $v_{i_j}^*$ ($i \in [\ell]$) such that u_h^* is preceded by $u_{i_j}^*$, which means that v_h^* is also preceded by $u_{i_j}^*$. Therefore, all vertices v_h^* in C_i^+ must be inside $T'(u_{i_j}^*)$ for some vertex $v_{i_j}^*$.

By Condition IV.28, all the vertices $u_{i_1}^*, \ldots, u_{i_\ell}^*$ are children of the root x of T'. By Condition IV.23, there must be green edges connecting each connected component $C_h^* \subseteq C^+$. These components C_h^* are connected together through vertices in different C_h^* getting contracted into the same $v_{i_j}^*$, which means they share vertices in a common $T'(u_{i_j}^*)$. It follows that all vertices in C_i^+ belong to subtrees in a single $U \in \mathcal{U}$. By Condition IV.24, there cannot be any green edges in $\bigcup_{i_j} \partial_G(T(u_{i_j}^*)) \setminus \bigcup_i T_i^*$, so there cannot be any other vertices v_h^* in U. ∎

Assumption IV.34. *After running Algorithms IV.25 and IV.30, Conditions IV.23, IV.24, IV.26, IV.27, and IV.28 hold, and for each C_i^+, there exists a set $U_i^+ \in \mathcal{U}$ whose subtrees contain precisely all vertices v_j^* in C_i^+ (and no vertices v_j^* outside C_i^+). (Note that $\bigcup_i U_i^+ = \{u_1^*, \ldots, u_q^*\}$. This assumption holds with probability at least $2^{-O(k)}(\log n)^{-k}\lambda^{-k}$.)*

Before we move on to the next section, for each $U \in \mathcal{U}$, let us define

$$\mathsf{MinElts}(U) := \Big(\bigcup_{u \in U} \{v : v \in T'(u) \text{ and }$$
$$\exists (v, v') \in \partial T'(u) \text{ green}\}\Big)\downarrow.$$

Note that if we assume Conditions IV.23, IV.24 and IV.27, then for each $U_i^+ = \{u_{i_1}^*, \ldots, u_{i_\ell}^*\}$, we also have $\mathsf{MinElts}(U_i^+) = \big(\bigcup_j \{b(v) : v \in T(u_{i_j}) \text{ and } \exists (v, v') \in \partial T(u_{i_j}) \text{ green}\}\big)\downarrow \subseteq \bigcup_j B_{i_j}^*$. After constructing $\mathsf{MinElts}(U)$ for each $U \in \mathcal{U}$, the algorithm may forget the construction of T'; indeed, T' was only needed to construct the sets U and $\mathsf{MinElts}(U)$. However, we will still need T' for our analysis.

2) General Case: Processing the Sets $U \in \mathcal{U}$: We begin similarly to Section IV-A2: define a similar *minimum ancestor p-cut* problem for $p \in [k' - 1]$:

Definition IV.35 (Minimum ancestor p-cut). *Fix some set $U = \{u_1, \ldots, u_\ell\} \in \mathcal{U}$, and let $\mathsf{MinElts}(U) = \{s_1, \ldots, s_m\}$. The minimum ancestor p-cut is the following problem: For each $i \in [\ell]$, remove at least one edge in $T(u_i)$ (note: not $T'(u_i)$) such that after removal, no $s_j \in V(T(u_i))$ is in the same component as u_i, and such that exactly $p - 1$ edges are removed in total (over all i). Consider the p-cut in G from the connected components of the remaining forest. We want to compute the p-cut of minimum size, denoted $\mathsf{MinAncCut}(U, p)$.*

Claim IV.36. *Assuming Assumption IV.34, for each $U_i^+ = \{u_{i_1}^*, \ldots, u_{i_\ell}^*\}$ with $p_i^+ := |\bigcup_j V(T(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\}|$, we have $|\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]| = |\mathsf{MinAncCut}(U_i^+, p_i^+)|$. Moreover, $\sum_{U_i^+} |\mathsf{MinAncCut}(U_i^+, p_i^+)| = OPT$.*

Proof: The proof resembles the proof of Claim IV.10, except with the additional HLD and Assumption IV.41 added in.

Let the set C_i^+ have vertices $v_{i_1}^*, \ldots, v_{i_{\ell'}}^*$ for $\ell' \geq \ell$. Assume that the components $C_1^*, \ldots, C_y^*$ are reordered so that $C_1^*, \ldots, C_z^* \subseteq C_i^+$ are the connected components of H whose vertices form C_i^+. For simpler notation, define $p := p_i^+$, and let $v_{i_{\ell+1}}^*, \ldots, v_{i_p}^*$ be the other vertices in $\bigcup_j V(T'(u_{i_j}^*))$, that is, those in $\bigcup_j V(T'(u_{i_j}^*)) \cap \{v_{r+1}^*, \ldots, v_{k'-1}^*\}$. First, we claim that a p-ancestor cut of size $|\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]|$ is achievable: simply cut the parent edges of $v_{i_1}^*, \ldots, v_{i_p}^*$. We now show that this cut is indeed an ancestor cut. By Conditions IV.23 and IV.24, the only green edges in any $\partial_G(T(u_{i_j}^*))$ lie in some T_i^*, $i \in [z]$. By Condition IV.27, all endpoints of $T_1^*, \ldots, T_z^*$ in T' are on the branches $B_{i_1}^*, \ldots, B_{i_{\ell'}}^*$. Inside each branch $B_{i_j}^*$, vertex $v_{i_j}^*$ precedes any endpoint in $T(x)$ of a green edge on $B_{i_j}^*$, and these branches cover all green edges. Therefore, every vertex in $\mathsf{MinElts}(U_i^+)$ is preceded by some vertex $v_{i_j}^*$ in T'. Since T' is a contraction of $T(x)$, precedence is unchanged, so every vertex in $\mathsf{MinElts}(U_i^+)$ is also preceded by some vertex $v_{i_j}^*$ in $T(x)$. In other words, the parent edges of $v_{i_j}^*$ lie on the branches in $T(x)$ from vertices in $\mathsf{MinElts}(U_i^+)$ to x, so our cut is a valid ancestor cut.

The proof that no better ancestor cut is possible is identical (see the proof of Claim IV.10), so we omit it. ∎

We now construct our graph G', similar to the one in Section IV-A2. Fix a set $U = \{u_1, \ldots, u_\ell\}$ and let $\mathsf{MinElts}(U) = \{s_1, \ldots, s_\ell\}$ where $s_i \in V(T(u_i))$. For each $p \in [k' - 1]$, we will compute a value $f(U, p)$. First, we construct the following multigraph G' on the vertices $(\bigcup_i V(T(u_i))) \cup \{x\}$:

(a) For each edge in $E \setminus E'$ with both endpoints in $T(\mathsf{MinElts}(U))$, do nothing: these edges are not cut in any ancestor cut.

(b) For each edge in E' with at least one endpoint in $T(\mathsf{MinElts}(U))$, do nothing: these edges are *always*

cut in an ancestor cut, and we will account for these edges separately.

   (c) For each edge in $E \setminus E'$ with at most one endpoint in $T(\mathsf{MinElts}(U))$, add it to G'.
   (d) For each edge $(u, v) \in E'$ with both endpoints not in $T(\mathsf{MinElts}(U))$, add edges (u, r) and (v, r) to G'.

Note that all edges in G' have both their endpoints in $V(T(u_i)) \cup \{x\}$ for some i. Next, for each sequence $p_1, \ldots, p_\ell$ of positive integers summing to p, do the following: For each $i \in [\ell]$, consider all ways to select p_i vertices $v_1, \ldots, v_{p_i} \in T(u_i)$ so that if we remove their parent edges, then no vertex in $\mathsf{MinElts}(U) \cap V(T(u_i))$ is in the same component as u_i; find the way that minimizes $|\bigcup_j \partial_{G'}(T(v_j))|$. Then, sum over the costs of the $(p_i + 1)$-cuts for each $i \in [\ell]$. Finally, compute the minimum sum over all sequences $p_1, \ldots, p_\ell$, and add the number of edges considered in step (b) to this minimum. The final value is $f(U, p)$.

It is not clear how to compute $f(U, p)$ quickly, and this is where we will use the previously computed $\mathsf{State}(\cdot, \cdot)$'s; we defer this to Section IV-B3. The intuition for the construction of G' is the same as in Section IV-A2.

Again, define the following natural correspondence between ancestor cuts and the "G'-cuts": two correspond to each other if their cut edges are identical.

Lemma IV.37. *For two corresponding cuts, the size of the ancestor cut is at most the size of the G'-cut plus the number of edges in step (b).*

   Proof: The proof is identical to the one in Lemma IV.12, so we omit it. ∎

Lemma IV.38. *Assuming Assumption IV.8, for each U_i^+ and p_i^+, $\mathsf{MinAncCut}(U_i^+, p_i^+)$ has size exactly $f(U_i^+, p_i^+)$.*

   Proof: In the proof of Lemma IV.12 (adapted to suit Lemma IV.37), the only potential source of inequality is in (d): an edge with no endpoints in x's component of $\mathsf{MinAncCut}(U_i^+, p_i^+)$ contributes 1 to $|\mathsf{MinAncCut}(U_i^+, p_i^+)|$ and 2 to the corresponding G'-cut. If such an edge e existed, then it must be in $\partial_G[S_{i_1}^*, \ldots, S_{i_\ell}^*]$ where $V(C_i^+) := \{v_{i_1}^*, \ldots, v_{i_\ell}^*\}$, which means e is in component C_i^+. Also, neither of its endpoints is preceded by a vertex in $\mathsf{MinElts}(U_i^+)$, which means that for any edge in any $C_j^* \subseteq C_i^+$ with endpoints u and v in G, neither of e's endpoints is preceded by $b(u)$ or $b(v)$. In particular, this is true for any edge in any T_j^* as well. We thus have an edge e in some $C_j^* \subseteq C_i^+$ not partially preceded by any endpoint in T_j^*, contradicting the definition of T_j^*. Therefore, no such edge exists, and we have equality. ∎

Let us first assume that for each $U \in \mathcal{U}$ and $p \in [k' - 1]$, we can compute $f(U, p)$ in time $n^{o(k)}$. (Unlike the one in Section IV-A2, this assumption is nontrivial and is covered in Sections IV-B3 and IV-B4.) Then, as in Section IV-A2,

we can compute the minimum similar to (3) as a knapsack problem:

Algorithm IV.39. *Compute*

$$\min_{\substack{U_1, \ldots, U_\ell \\ \sum_i |U_i| = r - 1 \\ p_1, \ldots, p_\ell \\ \sum_i p_i = k' - 1}} \sum_{i=1}^{\ell} f(U_i, p_i). \tag{5}$$

in polynomial time by formulating it as a knapsack problem. Assuming Assumption IV.34, the result is exactly OPT.

We have the final algorithm below, which proves Theorem III.1:

Algorithm IV.40. *For $O(2^{O(k)} \lambda^k \log n)$ repetitions, run Algorithms IV.25 and IV.39, and output the minimum value of (5) ever computed.*

   3) Computing $f(U, p)$: In this section, we describe how to compute $f(U, p)$ in a rather ad-hoc way, which is where we pick up the $n^{o(k)}$ multiplicative factor. Note that it can be easily computed in $O(k^{O(k)} n^{p + O(1)})$ time by brute force, which is fine if $p = O(1)$ or even $o(k)$. However, p could be as large as $\Theta(k)$.

First, let us make an additional assumption for now. We later show how to remove this assumption, at a cost of a multiplicative $n^{o(k)}$ factor.

Assumption IV.41. *For each $U_i^+ = \{u_{i_1}^*, \ldots, u_{i_\ell}^*\}$, we have $|V(T'(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_r^*\}| \leq \sqrt{k}$ for all $j \in [\ell]$. (Note that $V(T'(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_r^*\}$ is also $\big(V(T'(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\}\big)\!\downarrow.$)*

In this case, the algorithm computes an estimate $f'(U, p)$ of $f(U, p)$ so that $f'(U, p) \geq f(U, p)$ always, and $f'(U_i^*, p) = f(U_i^*, p)$. First, fix a set $U = \{u_1, \ldots, u_\ell\} \in \mathcal{U}$ and integer $p \in [k' - 1]$. Next, for each sequence $p_1, \ldots, p_\ell$ of positive integers summing to p, for each sequence of positive integers $r_1, \ldots, r_\ell$ with $r_i \leq \min\{p_i, \sqrt{k}\}$ (for all $i \in [\ell]$), do the following: For each $i \in [\ell]$, consider all ways to select r_i incomparable vertices $v_1, \ldots, v_{r_i} \in T'(u_i) \setminus \{u_i\}$ and remove their parent edges so that no vertex in $\mathsf{MinElts}(U) \cap V(T(u_i))$ is in the same component as u_i; find the way that minimizes

$$\min_{\substack{k_1', \ldots, k_{r_i}' \geq 1 \\ \sum_j k_j' = p_i}} \sum_{j=1}^{r_i} \mathsf{State}(v_j, k_j') + \left| \bigcup_{j=1}^{r_i} \partial_{G'}(T(v_j)) \right|. \tag{6}$$

Let the minimum of (6) over all selections of $v_1, \ldots, v_{r_i}$ be M_i. Finally, compute the minimum sum $\sum_{i \in [\ell]} M_i$ over all sequences $p_1, \ldots, p_\ell$, and add the number of edges considered in step (b) to this minimum. The final value is $f'(U, p)$. Clearly, $f'(U, p)$ can be computed in $O(k^{O(k)} n^{\sqrt{k} + O(1)})$ time.

Claim IV.42. *For all $U \in \mathcal{U}$ and $p \in [k'-1]$, $f'(U, p) \geq f(U, p)$.*

Proof: Let $U = \{u_1, \ldots, u_\ell\}$ as before. Given $i \in [\ell]$ and values p_i, r_i and vertices $v_1, \ldots, v_{r_i}$, (6) represents the way to select $(p_i - r_i)$ vertices $v_{r_i+1}, v_{r_i+2}, \ldots, v_{p_i}$ so that $v_1, \ldots, v_{p_i}$ form an ancestor p_i-cut for U and $\{v_1, \ldots, v_{p_i}\}\!\downarrow = \{v_1, \ldots, v_{r_i}\}$. This is because the values $\mathsf{State}(v_j, k'_j)$ capture the way to cut $k'_j - 1$ additional edges inside each $T(v_j)$ that minimizes the number of edges cut in G_j. Note that which edges we cut in each $T(v_j)$ does not affect whether $v_1, \ldots, v_{p_i}$ is a valid ancestor cut, since the component containing u_i in the G'-cut is already determined by $\{v_1, \ldots, v_{p_i}\}\!\downarrow = \{v_1, \ldots, v_{r_i}\}$.

Therefore, $f'(U, p)$ is essentially computing the same as $f(U, p)$, except that in $f'(U, p)$, the values r_i are artificially limited by $\sqrt{k}$. So $f'(U, p)$ can only be larger than $f(U, p)$. ∎

Claim IV.43. *Assuming Assumption IV.41, for each U_i^+, $f'(U_i^+, p) = f(U_i^+, p)$.*

Proof: By the proof of Claim IV.42, the only way for $f'(U_i^+, p) > f(U_i^+, p)$ to happen is if the optimal selection of $r_1, \ldots, r_\ell$ has $r_i > \sqrt{k}$ for some i. However, this is precisely what Assumption IV.41 prevents. ∎

4) Removing Assumption IV.41: In this section, we deal with Assumption IV.41 in another ad-hoc way. Essentially, we show that if Assumption IV.41 does not hold, then we can preprocess the tree $T(x)$ in an efficient way (more precisely, with $n^{o(k)}$ multiplicative overhead) so that Assumption IV.41 does hold.

Definition IV.44 (Rank). *Given a vertex u_i^*, $i \in [q]$ (that is, u_i^* is in some $U_{i'}^+$), define the* rank *of a vertex $v \in T(u_i^*)$ as follows: Consider the branch from v to $T(u_i^*)$, which we call B. If B contains any vertex v_j^*, then the $\mathsf{rank}(v) = -\infty$. Otherwise, let $\mathsf{rank}(v)$ be the number of vertices $v_j^* \in T(u_i^*)$ such that if we travel along the branch from v_j^* to u_i^*, then we encounter a vertex in B before or at the same time as we encounter a vertex in $\{u_1^*, \ldots, u_r^*\} \setminus \{u_j^*\}$.*

Lemma IV.45. *Fix $i \in [q]$, and let a be the maximum rank of a vertex in $T(u_i^*)$. Then, $\left| \left(V(T(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\} \right)\!\downarrow \right| \leq 2^a$.*

Proof: We induct on $a > 1$, with the trivial base case $a = 0$. Now suppose that $a > 0$. Consider the set $S := \left(\{u_1^*, \ldots, u_r^*\} \cap (V(T(u_i^*)) \setminus \{u_i^*\}) \right)\!\downarrow$.

If $S = \emptyset$, then v_i^* is the only vertex in $\{v_1^*, \ldots, v_r^*\} \cap T(u_i^*)$, so $a = 1$ and the bound holds. Otherwise, let $S = \{u_{i_1}^*, \ldots, u_{i_\ell}^*\}$; we claim that for all $j \in [\ell]$, every vertex in $T(u_{i_j}^*)$ has rank at most $a - \ell$. Suppose not: there exists $j \in [\ell]$ and a vertex $v \in T(u_{i_j}^*)$ with rank more than $a - \ell$. Let B' be the branch from v to the parent of $u_{i_j}^*$. Then, there are more than $a - \ell$ vertices $v_{i'}^*$ such that if we travel the branch from $v_{i'}^*$ to the parent of $u_{i_j}^*$, then we encounter a
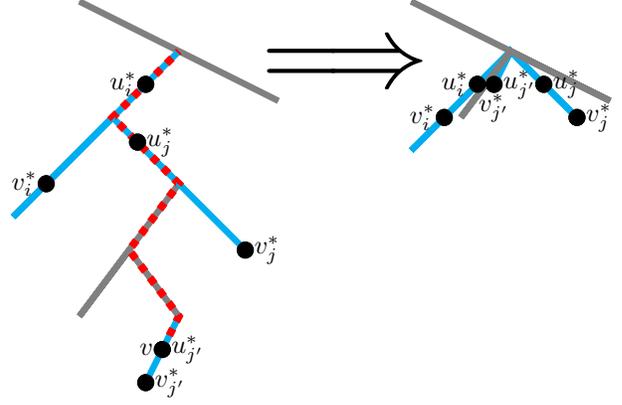

Figure 4. Contracting the branch from v to the parent of u_i^* (dashed red). The rank of v is 3. After the contraction, for the vertices v_j^* and $v_{j'}^*$, their new vertices u_j^* and $u_{j'}^*$ are no longer preceded by any other $u_{i'}^*$.

vertex in B' no later than we encounter some $u_{i''}^*$ for $i'' \neq i'$. Extend the branch B' to B'' so that B'' travels from v to the parent of u_i^*; clearly, the previous statement still holds if we travel the branch from $v_{i'}^*$ to the parent of u_i^* instead, and with B' replaced by B''. Moreover, for each $j' \neq j$, if we travel the branch from $v_{i_{j'}}^*$ to u_i^*, then we also encounter a vertex on B'' no later than we encounter any vertex in $\{u_1^* \ldots, u_r^*\} \setminus \{u_{i_{j'}}^*\}$ (which has to be u_i^*). Thus, these $\ell - 1$ vertices $v_{i_{j'}}^*$ increase the rank of v in $T(u_i^*)$ to more than $(a - \ell) + (\ell - 1) = a - 1$. Finally, v_i^* always increases the count by 1, so v has rank more than a, contradicting the assumption that the maximum rank inside $T(u_i^*)$ is a.

By induction, each $u_{i_j}^* \in S$ satisfies $\left| \left(V(T'(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\} \right)\!\downarrow \right| \leq 2^{a-\ell}$. Therefore,

$$
\begin{aligned}
& \left| \left(V(T'(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\} \right)\!\downarrow \right| \\
&= \left| \{v_i\} \cup \bigcup_{j=1}^{\ell} \left(V(T(u_{i_j}^*)) \cap \{v_1^*, \ldots, v_{k'-1}^*\} \right)\!\downarrow \right| \\
&\leq 1 + \ell \cdot 2^{a-\ell} \\
&< 1 + 2^\ell \cdot 2^{a-\ell} \\
&= 1 + 2^a,
\end{aligned}
$$

as desired. ∎

By Lemma IV.45, either Assumption IV.41 holds, or there exists some $i \in [q]$ such that $a \geq \log_2(\sqrt{n})$. In the former case, we compute $f'(U, p)$ as before. In the latter case, consider this integer i and a vertex v with $\mathsf{rank}(v) \geq \log_2(\sqrt{n})$. Contract the branch from v to the parent of u_i^*, and consider the new contracted graph with the same HLD (except with some edges contracted). Let the branch from v to u_i^* be B. Observe that for each vertex $v_j^* \in V(T(u_i^*))$ satisfying the property in the definition of

rank,[9] the vertex u_j^* is no longer preceded by any other $u_{j'}^*$ (see Figure 4). Previously, this property was only true for v_i^*. Therefore, if we perform the contraction, keep the contracted HLD, and compute the new values $u_1^*, \ldots, u_r^*$, then q increases by (at least) $\mathsf{rank}(v) - 1$. We can repeat this process at most $\left\lceil \frac{k}{\log_2(\sqrt{k})-1} \right\rceil$ times before Assumption IV.41 must hold, since otherwise, q would be greater than k. Let $t^\dagger$ be the number of repetitions, let $v_1^\dagger, \ldots, v\dagger_t$ the selected vertices, and let $T^\dagger$ be the contracted tree.

Our algorithm is as follows. Start with an arbitrary HLD of $T(x)$. Choose a random number $t \in \left[\left\lceil \frac{k}{\log_2(\sqrt{k})-1} \right\rceil\right]$, and choose t random vertices $v_1, \ldots, v_t \in V(T(x))$. Contract all the edges on the branches from each v_i to x. With probability $n^{-o(k)}$, we have $t = t^\dagger$ and $v_i = v_i^\dagger$ for each $i \in [t]$. Note that in this case, our contracted tree is $T^\dagger$ with possibly more edges contracted above any u_i^*. Therefore, the new HLD with branches $\mathcal{B}$ on our new tree satisfies Assumption IV.41. We then proceed with Algorithms IV.25 and IV.30, except we do not recompute the HLD in Algorithm IV.30. The algorithms in Sections IV-B2 and IV-B3 are the same, except we now have Assumption IV.41. Finally, repeat the entire algorithm $n^{o(k)}$ times, starting from guessing v and $v_1, \ldots, v_t$, so that w.h.p., Assumption IV.41 holds in one of the trials.

V. Kawarabayashi-Thorup Sparsification

In this section, we prove the following:

Theorem III.2. *Let G be a simple graph with minimum degree $\delta > \omega(\max\{\alpha \log n, \alpha k\})$, and let $\alpha \geq 1$ be a parameter. Then, we can contract G into a (multi-)graph $\overline{G}$ such that:*

1) *Suppose the minimum k-cut has size $\leq \alpha\delta$ in G. Then, every nontrivial minimum k-cut is preserved in $\overline{G}$. That is, no edge of such a cut is contracted in $\overline{G}$.*
2) *$\overline{G}$ has $\tilde{O}(\alpha m/\delta)$ edges and $\tilde{O}(\alpha m/\delta^2)$ vertices.*

The algorithm iteratively contracts vertices in G, possibly producing a multigraph. It is morally the same as the one in [27], [12], except with worse bounds since we are in the approximate, k-cut setting. The one key difference is that since we are not concerned with a near-linear running time, we replace their inner PageRank/UnitFlow subroutines with a simpler one that iteratively computes low-conductance cuts. Our algorithm is described in pseudocode in Algorithm 3. Below, we introduce the terminology from [12] that we use in our algorithm and analysis.

Terminology: Here, we list the terminology specific to Section V.

For a vertex subset S, define $\mathbf{vol}(S) = \sum_{v \in S} \deg(v)$ as the sum of degrees of vertices in S. A connected subgraph

$H \subseteq G$ is *cut* by edge set $F \subseteq E$ if $H - F$ is disconnected. A set of vertices $C \subseteq V$ is *cut* by F if $G[C] - F$ is disconnected.

The terminology below originate from [12]. In all the definitions, the graph $\overline{G}$ is fixed and has m edges.

Definition V.1 (Regular vertex and supervertex)**.** *A vertex v in $\overline{G}$ is a* supervertex *if more than one vertex in G contracts to v. Otherwise, it is a* regular vertex.

Observation V.2. *Every regular vertex $v \in V(\overline{G})$ satisfies $\deg_{\overline{G}}(v) \geq \delta$.*

Definition V.3 (Passive supervertex)**.** *A vertex $v \in V(\overline{G})$ is a* passive supervertex *if it is a supervertex with $\deg_{\overline{G}}(v) \leq 3\alpha\delta/\gamma$. ($\gamma := \frac{1}{100 \log m}$ from Definition V.5.)*

Definition V.4 (Conductance)**.** *Given a graph H, a set $S :$ $\emptyset \subsetneq S \subsetneq V(H)$ has* conductance

$$\frac{|\partial_H S|}{\min\{\mathbf{vol}(S), \mathbf{vol}(V(H) \setminus S)\}}.$$

Definition V.5 (γ, γ-expander)**.** *Define $\gamma := \frac{1}{100 \log m}$ throughout this entire section. A graph H is a γ-expander if every set $S \subseteq V(H)$ has conductance at most γ.*

Definition V.6 (Trim)**.** *Let H be a subgraph of $\overline{G}$. By* trimming *H, we mean iteratively removing from H any vertex $v \in V(H)$ satisfying $\deg_H(v) < \frac{2}{5}\deg_{\overline{G}}(v)$, until no such vertices are left.*

Definition V.7 (Loose vertex)**.** *Let C be a subgraph of $\overline{G}$. A vertex $v \in V(C)$ is* loose *if it is a regular vertex and $d_C(v) \leq d_{\overline{G}}(v)/2$.*

Definition V.8 (Shave)**.** *Let C be a subgraph of $\overline{G}$. By* shaving *C, we mean (simultaneously) removing all loose vertices in C. (Unlike trimming, this operation is not iterative.)*

Definition V.9 (Scrap)**.** *Let C be a subgraph of $\overline{G}$. Suppose we shaved C into a subgraph $A \subseteq H$. By* scrapping *A, we mean removing all vertices from A if $\mathbf{vol}_{\overline{G}[C]}(A) \leq \mathbf{vol}_{\overline{G}}(C)/4$ (and doing nothing otherwise).*

Definition V.10 (Core)**.** *Let C be a subgraph of $\overline{G}$. The* core *is the subgraph left over after shaving C into A and then scrapping A (that is, replacing A with $\emptyset$ if $\mathbf{vol}_{\overline{G}[C]}(A) \leq \mathbf{vol}_{\overline{G}}(C)/4$.)*

We first begin with a sparsification algorithm that makes calls to an *exact* minimum conductance cut algorithm. Of course, since conductance is an NP-hard problem, this algorithm is not polynomial-time. Then, we explain at the end how to make the algorithm approximate. We choose this approach for two reasons. First, the algorithm utilizing exact conductance cuts is cleaner to state and analyze, and transitioning to the approximate conductance case is straightforward. Second, for potential future applications which do not require a polynomial time algorithm (and

---

[9]Namely, "if we travel along the branch from v_j^* to u_j^*, then we encounter a vertex in B before or at the same time as we encounter a vertex in $\{u_1^*, \ldots, u_r^*\} \setminus \{u_j^*\}$."

mainly existence of the sparsification), the exact algorithm would suffice and give better bounds, so nothing needs to be reproven.

Algorithm 2 KT-Sparsification($G = (V, E), \alpha$)

1: $\overline{G} \leftarrow G$; G has min degree δ
2: **while** $< \frac{1}{20}$ fraction of the edges in $\overline{G}$ are incident to passive supervertices **do**
3:     $H \leftarrow \overline{G}$
4:     Remove passive supervertices from H and trim H
5:     **while** there exists a connected component C in H such that $H[C]$ is not a γ-expander **do**
6:         Compute a γ-conductance cut in $H[C]$, remove these edges from H, and trim H
7:     Take each connected component of H and contract its core (if nonempty) to a supervertex in $\overline{G}$
8: **return** $\overline{G}$

For the rest of this section, let us state the assumption on δ in Theorem III.2 as a formal assumption below:

Assumption V.11. δ *is larger than some absolute constant, and*

$$\delta > \omega(\max\{\alpha \log n, \alpha k\}).$$

For the rest of this section, we divide the proof of Theorem III.2 into three parts. In Section V-A, we show that the graph $\overline{G}$ returned by Algorithm 3 satisfies Condition 1 of Theorem III.2, the "correctness" guarantee. In Section V-B, we show that $\overline{G}$ satisfies Condition 2, the "quality" guarantee. Finally, in Section V-C, we show that Algorithm 3 indeed terminates, rather than looping indefinitely at line 2.

A. Correctness

In this section, we will prove the following lemma, which argues that the algorithm correctly preserves nontrivial k-cuts of small enough size (Condition 1 of Theorem III.2):

Lemma V.12 (Condition 1 of Theorem III.2)**.** *Suppose that the optimal k-cut has size $\leq \alpha\delta$ in G. Then, every non-optimal nontrivial k-cut is preserved in $\overline{G}$. That is, no edge of the cut is contracted in $\overline{G}$.*

The structure of the proof follows that of [12], except adapted to the approximate and k-cut settings.

Lemma V.13. *Consider a point in the algorithm's execution, just after it trimmed H. Let C be a connected component of H. Take a subset $S \subseteq V(C)$ containing only regular vertices, and satisfies $|\partial S| \leq \alpha\delta$. Then, either $|S| \leq 3\alpha$ or $|S| \geq \delta/5$.*

Proof: Since every vertex $v \in S$ is regular, we have $\deg_G(v) \geq \delta$, and moreover, since v is not trimmed from

H, we have $\deg_H(v) \geq \frac{2}{5}\deg_G(v) \geq \frac{2}{5}\delta$. Moreover, since all other vertices in S are regular, at most $|S| - 1$ of v's edges can go to another vertex in S. Since $\deg_H(v) \geq \frac{2}{5}\delta$, this leaves at least $\frac{2}{5}\delta - (|S| - 1)$ edges to a vertex outside S. Altogether, the $|S|$ regular vertices in S are responsible for at least $|S| \cdot (\frac{2}{5}\delta - (|S| - 1))$ many edges in ∂S. Therefore,

$$|S| \cdot (\frac{2}{5}\delta - (|S| - 1)) \leq |\partial S| \leq \alpha\delta$$
$$\iff -|S|^2 + (\frac{2}{5}\delta + 1)|S| \leq \alpha\delta$$
$$\iff |S|^2 - (\frac{2}{5}\delta + 1)|S| + \alpha\delta \geq 0.$$

The solution to $x^2 - (\frac{2}{5}\delta + 1)x + \alpha\delta$ is

$$x = \frac{(\frac{2}{5}\delta + 1) \pm \sqrt{(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta}}{2}.$$

By Assumption V.11, $(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta \geq 0$, so x has real solutions, and $|S|$ must satisfy

$$|S| \leq \frac{(\frac{2}{5}\delta + 1) - \sqrt{(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta}}{2} \quad \text{or}$$
$$|S| \geq \frac{(\frac{2}{5}\delta + 1) + \sqrt{(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta}}{2}.$$

For the second scenario, we clearly have $|S| \geq \delta/5$. Now consider the first scenario. We claim that

$$(\frac{2}{5}\delta + 1) - \sqrt{(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta} \leq 6\alpha,$$

which would imply that $|S| \leq 3\alpha$. This can be seen from

$$(\frac{2}{5}\delta + 1) - 6\alpha \leq \sqrt{(\frac{2}{5}\delta + 1)^2 - 4\alpha\delta}$$
$$\iff (\frac{2}{5}\delta + 1)^2 - 12\alpha(\frac{2}{5}\delta + 1) + 36\alpha^2 \leq (\frac{2}{5}\delta + 1)^2 - 4\alpha\delta$$
$$\iff -0.8\alpha\delta - 12\alpha + 36\alpha^2 \leq 0,$$

the last of which follows from Assumption V.11. ∎

Lemma V.14. *Consider a point in the algorithm's execution, just after it trimmed H. Let C be a connected component of H that is a γ-expander. Suppose that $H[C]$ is cut by $\partial_{H[C]}S$ for some $S \subseteq C$ satisfying $|\partial_{H[C]}S| \leq \alpha\delta$ and $|S| \leq |C|/2$. Then, S has at most 3α regular vertices and no supervertex in C.*

Proof: First, suppose for contradiction that S contains a supervertex $s \in C$. Then, since s is active, $\deg_{\overline{G}}(s) \geq 3\alpha\delta/\gamma$, and since it's not trimmed, $\deg_H(s) \geq \frac{2}{5}\deg_{\overline{G}}(s) \geq \frac{2}{5} \cdot 3\alpha\delta/\gamma$, so in particular, $\mathbf{vol}_{H[C]}(S) \geq \deg_{H[C]}(s) = \deg_H(s) \geq 1.2\alpha\delta/\gamma$. Therefore, the conductance of the set $S \cap C$ inside C is at most

$$\frac{|\partial_{H[C]}S|}{\mathbf{vol}_{H[C]}(S)} \leq \frac{\alpha\delta}{1.2\alpha\delta/\gamma} < \gamma,$$

so the set S contradicts the assumption that $H[C]$ is a γ-expander.

Now suppose for contradiction that S contains more than 3α regular vertices and no supervertex. By Lemma V.13, since $|S| > 3\alpha$, it must be that $|S| \geq \delta/5$. Since each vertex in S is regular and is not trimmed, it has degree at least $\frac{2}{5}\delta$ in H (Observation V.2), so $\mathbf{vol}_{H[C]}(S) \geq |S| \cdot \frac{2}{5}\delta \geq \frac{2}{25}\delta^2$. Again, the conductance of the set $S \cap C$ inside $H[C]$ is at most

$$\frac{|\partial_{H[C]}S|}{\mathbf{vol}_{H[C]}(S)} \leq \frac{\alpha\delta}{\frac{2}{25}\delta^2} \overset{(AV.11)}{<} o\left(\frac{1}{\log n}\right) < \gamma, \qquad (7)$$

so the set S contradicts the assumption that $H[C]$ is a γ-expander. ∎

Corollary V.15. *Consider a point in the algorithm's execution, just after it trimmed H. Let E^* be a k-cut of size $\leq \alpha\delta$ in $\overline{G}$, and let C be a connected component of H that is a γ-expander and is cut by E^*. Then, all but one component of E^* (in $\overline{G}$) have at most 3α regular vertices and no supervertices in C.*

Proof: Suppose not: there exist two such components that either contain more than 3α regular vertices or one supervertex in C. Let S^* be one such component with $|S^* \cap C| \leq |C|/2$. Observe that $\partial_{H[C]}(S^* \cap C) \subseteq \partial_{\overline{G}}S^*$, since every edge going across $S^* \cap C$ in $H[C]$ must also go across S^* in $\overline{G}$. Therefore,

$$|\partial_{H[C]}(S^* \cap C)| \leq |\partial_{\overline{G}}S^*| \leq \alpha\delta,$$

so we can apply Lemma V.14 on the set $S \cap X$, proving the statement. ∎

Lemma V.16. *Suppose the algorithm is at line 7 of an iteration. Let A be a core of a connected component C that we contract. Let E^* be a nontrivial k-cut of size $\leq \alpha\delta$ in $\overline{G}$. Then, A cannot be cut by E^*.*

Proof: Suppose not: $G[A] - OPT$ splits $G[A]$ into more than one component. Let S_0^* be the single component of OPT with $|S_0^* \cap C| > |C|/2$, if it exists, and let $S_1^*, \ldots, S_r^*$ be the components of OPT which intersect A and are not S_0^* (by assumption, one must exist). Since the last action of the algorithm before line 7 was trim H, by Corollary V.15, each of $S_1^*, \ldots, S_r^*$ has at most 3α regular vertices in C and no supervertex. Let $v \in S_1^*$ be arbitrary; we have $|E[\{v\}, S_i^* \cap C]| \leq 3\alpha$ for each $i \geq 1$, which means that $|E[\{v\}, (S_1^* \cup \cdots \cup S_r^*) \cap C]| \leq 3\alpha k$. Since v was not shaved, we have

$$\deg_C(v) \geq 0.51 \deg_{\overline{G}}(v)$$
$$\geq \frac{1}{2} \deg_{\overline{G}}(v) + 0.01\delta$$
$$\overset{(AV.11)}{>} \frac{1}{2} \deg_{\overline{G}}(v) + 3\alpha k$$
$$\geq \frac{1}{2} \deg_{\overline{G}}(v) - |E[\{v\}, (S_1^* \cup \cdots \cup S_r^*) \cap C]|,$$

so in particular, more than $\frac{1}{2}\deg_{\overline{G}}(v)$ edges of v go to vertices in $C \setminus (S_1^* \cup \cdots \cup S_r^*)$. This means that S_0^* must exist, and $|E[v, S_0^*]| \geq |E[v, S_0^* \cap C]| > \frac{1}{2}\deg_{\overline{G}}(v)$. Hence, we also have $|E[v, S_1^*]| < \frac{1}{2}\deg_{\overline{G}}(v)$.

Now consider another k-cut formed by moving v from S_1^* to S_0^*. This is still a k-cut, since the old k-cut E^* is nontrivial. Moreover, the value of the new k-cut is

$$|OPT| + |E[v, S_1^*]| - |E[v, S_0^*]|$$
$$< OPT + \frac{1}{2}\deg_{\overline{G}}(v) - \frac{1}{2}\deg_{\overline{G}}(v)$$
$$= OPT,$$

contradicting the choice of OPT. ∎

Finally, Lemma V.12 easily follows from Lemma V.16, since the only way the lemma can break is if we contract a set of vertices that OPT cuts in line 7, but this cannot happen by Lemma V.16. This concludes Lemma V.12.

B. Quality

In this section, we prove the lemma below (Condition 2 of Theorem III.2):

Lemma V.17 (Condition 2 of Theorem III.2)**.** *At the end of the algorithm, $\overline{G}$ has $\tilde{O}(\alpha m/\delta)$ edges and $\tilde{O}(\alpha m/\delta^2)$ vertices.*

We first introduce two lemmas, one directly from [12], and one we reprove:

Lemma V.18 (Lemma 17 of [12])**.** *There are $\Omega(\delta^2)$ edges from G contracted in each supervertex of $\overline{G}$.*

Lemma V.19 (Lemma 18 of [12], reproven)**.** *The total number of edges leaving passive supervertices is $O(\alpha \log n \cdot m/\delta)$.*

Proof: By Lemma V.18, every passive supervertex has $\Omega(\delta^2)$ edges contracted to it, and since there are m edges total, there are $O(m/\delta^2)$ passive supervertices. By definition, a passive supervertex has degree $\leq 3\alpha\delta/\gamma$, which is at most $O(m/\delta^2) \cdot 3\alpha\delta/\gamma = O(\alpha \log m \cdot m/\delta)$. ∎

We now prove Lemma V.17. Since `KT-Sparsification` terminates when a constant fraction of the edges of $\overline{G}$ are incident to passive supervertices, we conclude that when the algorithm terminates, $\overline{G}$ has $O(\alpha \log m \cdot m/\delta)$ edges. We now focus on the vertex bound of Lemma V.17. Since each regular vertex in $\overline{G}$ must have degree $\geq \delta$, there are at most $O(\alpha \log m \cdot m/\delta^2)$ many of them. From the proof of Lemma V.19, there are $O(m/\delta^2)$ supervertices in $\overline{G}$, so altogether, $\overline{G}$ has $O(\alpha \log m \cdot m/\delta^2)$ vertices. This concludes the proof of Lemma V.17.

C. Termination

In [12], since they aimed at a near-linear time algorithm, they needed the graph $\overline{G}$ to shrink by a constant factor per

iteration. Here, all we need is *some* progress in $\overline{G}$, i.e., one single contraction, so that the algorithm does not loop indefinitely. Nevertheless, we will still prove that the number of edges of $\overline{G}$ decreases by a constant factor:

Lemma V.20 (Lemma A.5 of [27]). *In each except for the last iteration of the outer loop, i.e., the repeat loop, the number of edges in the graph $\overline{G}$ is decreased by a factor of at least $7/10$.*

The following lemma, which relates the total number of edges cut during an iteration to the low-conductance cuts in line 6, is mostly unchanged.

Lemma V.21 (Lemma A.4 of [27]). *If the total number of edges cut in line 6 during an iteration of the outer loop is c, then the total number of edges lost from all clusters due to trimming, shaving, and scrapping during this iteration is $6c$.*

The only part of its proof that is different is the scrapping part, since we defined loose vertices slightly differently from [12]. We prove our version of this part below, which is conveniently captured as Lemma A.3 in [27]. (The use of k in Lemma A.3 clashes with our notion of k (in k-cut), so we changed it to k' instead.)

Lemma V.22 (Lemma A.3 of [27], reproven). *If a component C has k' edges leaving it in $\overline{G}$ and the core of C is scrapped, then $\mathbf{vol}_{\overline{G}}(C) \le 4k'$.*

Proof: Call an edge with at least one endpoint in C *internal to the core* if both of its endpoints are inside the core. We first prove that there are always at most $3k'$ edges in C that are not internal to the core. There are two types of edges incident to C and not internal to the core: the edges incident to loose vertices in C and the edges in $E[C, \overline{G} - C]$. For the first type, since every loose vertex $v \in C$ has $|E[v, V(\overline{G}) \setminus C]| \ge 0.49 \deg_{\overline{G}}(v)$, the number of edges of the first type, denoted n_1, is at most

$$\sum_{\text{loose } v} \deg_{\overline{G}}(v) \le \frac{1}{0.49} \sum_{\text{loose } v} |E[v, V(\overline{G}) \setminus C]|.$$

But the number of second type edges, denoted n_2, is exactly $k' - \sum_{\text{loose } v} |E[v, V(\overline{G}) \setminus C]|$, so altogether, there are

$$n_1 + n_2 \le \frac{1}{0.49} \sum_{\text{loose } v} |E[v, V(\overline{G}) \setminus C]|$$
$$+ \left(k - \sum_{\text{loose } v} |E[v, V(\overline{G}) \setminus C]| \right)$$
$$\le \frac{1}{0.49} k' \le 2.1k' < 3k'$$

many edges not internal to the core.

Now suppose the core A is scraped. By definition, this means that $\mathbf{vol}_{H[C]}(A) \le \mathbf{vol}_{\overline{G}}(C)/4$. Now observe that

$\mathbf{vol}_{\overline{G}}(C)$ is exactly $\mathbf{vol}_{H[C]}(A) + n_1 + n_2$, so

$$\mathbf{vol}_{\overline{G}}(C) = \mathbf{vol}_{H[C]}(A) + n_1 + n_2 \le \mathbf{vol}_{\overline{G}}(C)/4 + 3k$$

$$\implies \mathbf{vol}_{\overline{G}}(C) \le 4k'.$$

$\blacksquare$

We now bound the number of edges cut by low-conductance cuts (line 6). Instead of using the more complicated procedures in [27], [12], we resort to simply iteratively computing low-conductance cuts. As for why the low-conductance cuts cut a small number of edges in total, the reasoning is the same as the one in [12], and we sketch it here for convenience.

Lemma V.23. *For an appropriate γ in the definition of γ, the total number of edges cut in line 6 during an iteration of the outer loop is at most $0.04|E(\overline{G})|$.*

Proof: We set up a charging scheme as follows: every time we compute a low-conductance cut E' in a connected component C of H, we charge a total cost of $|E'|$ uniformly to the edges of the smaller side S of C (the side with smaller $\mathbf{vol}_{\overline{G}}(S)$). Since E' has conductance $\le \gamma$, we have

$$\frac{|E'|}{\mathbf{vol}_{\overline{G}}(S)} = \frac{|E'|}{2|E[S]| + |E'|} \le \gamma$$

$$\implies |E'| \le \frac{2\gamma}{1 - \gamma}|E[S]| \le 4\gamma|E[S]|,$$

so every edge in $E[S]$ is charged at most 4γ. Now observe that every time an edge is charged, since it belongs to the smaller side of the cut, the size of the component containing this edge halves, so an edge is charged at most $\log m$ times. In total, an edge is charged a cost of at most $4\gamma \log m$. Therefore, at most $4\gamma|E(\overline{G})| \log m = 0.04|E(\overline{G})|$ cost was charged in total, and this also upper bounds the total edges cut. $\blacksquare$

Therefore, by Lemmas V.21 and V.23, the algorithm cuts at most $7 \cdot 0.04|E(\overline{G})|$ edges in $\overline{G}$ on a given iteration. Since it contracts the rest, we have proven Lemma V.20.

D. Polynomial-time Algorithm

Next, we modify the algorithm `KT-Sparsification` to take an approximate conductance cut algorithm instead, making it run in polynomial time. We use the $O(\sqrt{\log n})$-approximation algorithm of Arora, Rao, and Vazirani below:

Theorem V.24 ([28]). *There exists a universal constant $C > 0$ and a polynomial-time $C\sqrt{\log m}$-approximation algorithm for minimum conductance cut.*

Algorithm 3 KT-Sparsification-Polytime($G = (V, E), \alpha$)

1: $\overline{G} \leftarrow G$; G has min degree δ
2: **while** $< \frac{1}{20}$ fraction of the edges in $\overline{G}$ are incident to passive supervertices **do**
3:   $H \leftarrow \overline{G}$
4:   Remove passive supervertices from H and trim H
5:   **while** there exists a connected component C in H such that $H[C]$ is not a γ-expander **do**
6:     Compute a $C\sqrt{\log m} \cdot \gamma$-conductance cut in $H[C]$, remove these edges from H, and trim H
7:   Take each connected component of H and contract its core (if nonempty) to a supervertex in $\overline{G}$
8: **return** $\overline{G}$

The entire analysis goes through without change, except for the following differences:

1) We re-define the parameter $\gamma := \frac{1}{100C\log^{1.5} m}$ throughout the entire algorithm and analysis.

2) Assumption V.11 is replaced by the assumption that

$$\delta > \omega(\max\{\alpha\log^{1.5} n, \alpha k\}) \qquad (8)$$

instead.

3) In the proof of Lemma V.14, Equation (7) is replaced with

$$\frac{|\partial_{H[C]}S|}{\mathbf{vol}_{H[C]}(S)} \leq \frac{\alpha\delta}{\frac{2}{25}\delta^2} \overset{(8)}{<} o\left(\frac{1}{\log^{1.5} n}\right) < \gamma,$$

and the rest of the proof of Lemma V.14 is identical.

4) In the proof of Lemma V.23, every instance of γ is replaced by $C\sqrt{\log m} \cdot \gamma$ (with the new value of γ). Since $C\sqrt{\log m} \cdot \gamma = \frac{1}{100\log m}$, which is exactly the old value of γ, the rest of the proof of Lemma V.23 is unchanged.

References

[1] O. Goldschmidt and D. S. Hochbaum, "A polynomial algorithm for the k-cut problem for fixed k," *Math. Oper. Res.*, vol. 19, no. 1, pp. 24–37, 1994. [Online]. Available: http://dx.doi.org/10.1287/moor.19.1.24

[2] D. R. Karger and C. Stein, "A new approach to the minimum cut problem," *Journal of the ACM (JACM)*, vol. 43, no. 4, pp. 601–640, 1996.

[3] A. Gupta, E. Lee, and J. Li, "The number of minimum k-cuts: Improving the karger-stein bound," in *STOC 2019, to appear*, 2019.

[4] Y. Kamidoi, N. Yoshida, and H. Nagamochi, "A deterministic algorithm for finding all minimum k-way cuts," *SIAM J. Comput.*, vol. 36, no. 5, pp. 1329–1341, 2006/07. [Online]. Available: http://dx.doi.org/10.1137/050631616

[5] M. Thorup, "Minimum k-way cuts via deterministic greedy tree packing," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*. ACM, 2008, pp. 159–166.

[6] C. Chekuri, K. Quanrud, and C. Xu, "Lp relaxation and tree packing for minimum k-cuts," *arXiv preprint arXiv:1808.05765*, 2018.

[7] A. Gupta, E. Lee, and J. Li, "Faster exact and approximate algorithms for k-cut," in *Foundations of Computer Science (FOCS), 2018 IEEE 59th Annual Symposium on*, 2018.

[8] F. Le Gall, "Powers of tensors and fast matrix multiplication," in *Proceedings of the 39th international symposium on symbolic and algebraic computation*. ACM, 2014, pp. 296–303.

[9] V. V. Williams, "Multiplying matrices faster than Coppersmith–Winograd," in *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. ACM, 2012, pp. 887–898.

[10] V. V. Williams and R. Williams, "Subcubic equivalences between path, matrix and triangle problems," in *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*. IEEE, 2010, pp. 645–654.

[11] A. Abboud, A. Backurs, and V. V. Williams, "If the current clique algorithms are optimal, so is Valiant's parser," in *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*. IEEE, 2015, pp. 98–117.

[12] K.-i. Kawarabayashi and M. Thorup, "Deterministic edge connectivity in near-linear time," *Journal of the ACM (JACM)*, vol. 66, no. 1, p. 4, 2018.

[13] N. Alon, R. Yuster, and U. Zwick, "Color-coding," *J. ACM*, vol. 42, no. 4, pp. 844–856, 1995.

[14] H. Saran and V. V. Vazirani, "Finding k-cuts within twice the optimal," *SIAM Journal on Computing*, vol. 24, no. 1, pp. 101–108, 1995.

[15] J. Naor and Y. Rabani, "Tree packing and approximating k-cuts," in *Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms (Washington, DC, 2001)*. SIAM, Philadelphia, PA, 2001, pp. 26–27.

[16] R. Ravi and A. Sinha, "Approximating k-cuts using network strength as a Lagrangean relaxation," *European J. Oper. Res.*, vol. 186, no. 1, pp. 77–90, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ejor.2007.01.040

[17] M. Xiao, L. Cai, and A. C.-C. Yao, "Tight approximation ratio of a general greedy splitting algorithm for the minimum k-way cut problem," *Algorithmica*, vol. 59, no. 4, pp. 510–520, 2011.

[18] S. Kapoor, "On minimum 3-cuts and approximating k-cuts using cut trees," in *Integer programming and combinatorial optimization (Vancouver, BC, 1996)*, ser. Lecture Notes in Comput. Sci. Springer, Berlin, 1996, vol. 1084, pp. 132–146. [Online]. Available: http://dx.doi.org/10.1007/3-540-61310-2_11

[19] L. Zhao, H. Nagamochi, and T. Ibaraki, "Approximating the minimum k-way cut in a graph via minimum 3-way cuts," *J. Comb. Optim.*, vol. 5, no. 4, pp. 397–410, 2001. [Online]. Available: http://dx.doi.org/10.1023/A:1011620607786

[20] P. Manurangsi, "Inapproximability of Maximum Edge Biclique, Maximum Balanced Biclique and Minimum k-Cut from the Small Set Expansion Hypothesis," in *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 80, 2017, pp. 79:1–79:14. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2017/7500

[21] R. G. Downey, V. Estivill-Castro, M. Fellows, E. Prieto, and F. A. Rosamund, "Cutting up is hard to do: The parameterised complexity of k-cut and related problems," *Electronic Notes in Theoretical Computer Science*, vol. 78, pp. 209–222, 2003.

[22] K.-i. Kawarabayashi and M. Thorup, "The minimum k-way cut of bounded size is fixed-parameter tractable," in *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*. IEEE, 2011, pp. 160–169.

[23] R. Chitnis, M. Cygan, M. Hajiaghayi, M. Pilipczuk, and M. Pilipczuk, "Designing FPT algorithms for cut problems using randomized contractions," *SIAM J. Comput.*, vol. 45, no. 4, pp. 1171–1229, 2016. [Online]. Available: http://dx.doi.org/10.1137/15M1032077

[24] M. Cygan, F. V. Fomin, Ł. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized algorithms*. Springer, Cham, 2015. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-21275-3

[25] D. Marx, "Parameterized complexity and approximation algorithms," *The Computer Journal*, vol. 51, no. 1, pp. 60–78, 2007.

[26] H. Nagamochi and T. Ibaraki, "Computing edge-connectivity in multigraphs and capacitated graphs," *SIAM J. Discrete Math.*, vol. 5, no. 1, pp. 54–66, 1992. [Online]. Available: http://dx.doi.org/10.1137/0405004

[27] M. Henzinger, S. Rao, and D. Wang, "Local flow partitioning for faster edge connectivity," in *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2017, pp. 1919–1938.

[28] S. Arora, S. Rao, and U. Vazirani, "Expander flows, geometric embeddings and graph partitioning," *Journal of the ACM (JACM)*, vol. 56, no. 2, p. 5, 2009.

VI. Proof of Theorem I.2

Given a $(k-1)$-clique graph instance $G = (V, E)$, construct the following graph H: Let W be a clique of size $k^2 n$. Take the union of G and W, and then for each vertex $v \in V$, add $n - \deg(v)$ edges to arbitrary vertices in W. This is the graph H. Note that a k-cut of size $\leq (k-1)n$ can be formed by isolating any $(k-1)$ vertices in G. We now claim that G has a $(k-1)$-clique iff H has minimum k-cut $(k-1)n - \binom{k-1}{2}$.

Fix a minimum k-cut $S_1^*, \ldots, S_k^*$ in H. First, if W is not entirely contained in one component, then the cut has size $\geq k^2 n - 1$ already, so we can assume that this does not happen. Let us assume that W is contained in S_k^*. Next, if

$|S_1^*| + |S_2^*| + \cdots + |S_{k-1}^*| > k - 1$, then this k-cut will have size $\geq kn - \binom{k-1}{2} > (k-1)n$ (we can assume $n \gg k$), so this cannot happen either. Therefore, S_i^* is a singleton vertex $v_i^* \in V$ for all $i \in [k-1]$. It follows that the minimum k-cut has cost exactly $(k-1)n - |E[\{v_1^*\}, \ldots, \{v_{k-1}^*\}]|$. This is exactly $(k-1)n - \binom{k-1}{2}$ iff G has a $(k-1)$-clique.

References

[1] O. Goldschmidt and D. S. Hochbaum, "A polynomial algorithm for the k-cut problem for fixed k," *Math. Oper. Res.*, vol. 19, no. 1, pp. 24–37, 1994. [Online]. Available: http://dx.doi.org/10.1287/moor.19.1.24

[2] D. R. Karger and C. Stein, "A new approach to the minimum cut problem," *Journal of the ACM (JACM)*, vol. 43, no. 4, pp. 601–640, 1996.

[3] A. Gupta, E. Lee, and J. Li, "The number of minimum k-cuts: Improving the karger-stein bound," in *STOC 2019, to appear*, 2019.

[4] Y. Kamidoi, N. Yoshida, and H. Nagamochi, "A deterministic algorithm for finding all minimum k-way cuts," *SIAM J. Comput.*, vol. 36, no. 5, pp. 1329–1341, 2006/07. [Online]. Available: http://dx.doi.org/10.1137/050631616

[5] M. Thorup, "Minimum k-way cuts via deterministic greedy tree packing," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*. ACM, 2008, pp. 159–166.

[6] C. Chekuri, K. Quanrud, and C. Xu, "Lp relaxation and tree packing for minimum k-cuts," *arXiv preprint arXiv:1808.05765*, 2018.

[7] A. Gupta, E. Lee, and J. Li, "Faster exact and approximate algorithms for k-cut," in *Foundations of Computer Science (FOCS), 2018 IEEE 59th Annual Symposium on*, 2018.

[8] F. Le Gall, "Powers of tensors and fast matrix multiplication," in *Proceedings of the 39th international symposium on symbolic and algebraic computation*. ACM, 2014, pp. 296–303.

[9] V. V. Williams, "Multiplying matrices faster than Coppersmith–Winograd," in *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. ACM, 2012, pp. 887–898.

[10] V. V. Williams and R. Williams, "Subcubic equivalences between path, matrix and triangle problems," in *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*. IEEE, 2010, pp. 645–654.

[11] A. Abboud, A. Backurs, and V. V. Williams, "If the current clique algorithms are optimal, so is Valiant's parser," in *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*. IEEE, 2015, pp. 98–117.

[12] K.-i. Kawarabayashi and M. Thorup, "Deterministic edge connectivity in near-linear time," *Journal of the ACM (JACM)*, vol. 66, no. 1, p. 4, 2018.

[13] N. Alon, R. Yuster, and U. Zwick, "Color-coding," *J. ACM*, vol. 42, no. 4, pp. 844–856, 1995.

[14] H. Saran and V. V. Vazirani, "Finding k-cuts within twice the optimal," *SIAM Journal on Computing*, vol. 24, no. 1, pp. 101–108, 1995.

[15] J. Naor and Y. Rabani, "Tree packing and approximating k-cuts," in *Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms (Washington, DC, 2001)*. SIAM, Philadelphia, PA, 2001, pp. 26–27.

[16] R. Ravi and A. Sinha, "Approximating k-cuts using network strength as a Lagrangean relaxation," *European J. Oper. Res.*, vol. 186, no. 1, pp. 77–90, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ejor.2007.01.040

[17] M. Xiao, L. Cai, and A. C.-C. Yao, "Tight approximation ratio of a general greedy splitting algorithm for the minimum k-way cut problem," *Algorithmica*, vol. 59, no. 4, pp. 510–520, 2011.

[18] S. Kapoor, "On minimum 3-cuts and approximating k-cuts using cut trees," in *Integer programming and combinatorial optimization (Vancouver, BC, 1996)*, ser. Lecture Notes in Comput. Sci. Springer, Berlin, 1996, vol. 1084, pp. 132–146. [Online]. Available: http://dx.doi.org/10.1007/3-540-61310-2_11

[19] L. Zhao, H. Nagamochi, and T. Ibaraki, "Approximating the minimum k-way cut in a graph via minimum 3-way cuts," *J. Comb. Optim.*, vol. 5, no. 4, pp. 397–410, 2001. [Online]. Available: http://dx.doi.org/10.1023/A:1011620607786

[20] P. Manurangsi, "Inapproximability of Maximum Edge Biclique, Maximum Balanced Biclique and Minimum k-Cut from the Small Set Expansion Hypothesis," in *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 80, 2017, pp. 79:1–79:14. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2017/7500

[21] R. G. Downey, V. Estivill-Castro, M. Fellows, E. Prieto, and F. A. Rosamund, "Cutting up is hard to do: The parameterised complexity of k-cut and related problems," *Electronic Notes in Theoretical Computer Science*, vol. 78, pp. 209–222, 2003.

[22] K.-i. Kawarabayashi and M. Thorup, "The minimum k-way cut of bounded size is fixed-parameter tractable," in *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*. IEEE, 2011, pp. 160–169.

[23] R. Chitnis, M. Cygan, M. Hajiaghayi, M. Pilipczuk, and M. Pilipczuk, "Designing FPT algorithms for cut problems using randomized contractions," *SIAM J. Comput.*, vol. 45, no. 4, pp. 1171–1229, 2016. [Online]. Available: http://dx.doi.org/10.1137/15M1032077

[24] M. Cygan, F. V. Fomin, Ł. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized algorithms*. Springer, Cham, 2015. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-21275-3

[25] D. Marx, "Parameterized complexity and approximation algorithms," *The Computer Journal*, vol. 51, no. 1, pp. 60–78, 2007.

[26] H. Nagamochi and T. Ibaraki, "Computing edge-connectivity in multigraphs and capacitated graphs," *SIAM J. Discrete Math.*, vol. 5, no. 1, pp. 54–66, 1992. [Online]. Available: http://dx.doi.org/10.1137/0405004

[27] M. Henzinger, S. Rao, and D. Wang, "Local flow partitioning for faster edge connectivity," in *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2017, pp. 1919–1938.

[28] S. Arora, S. Rao, and U. Vazirani, "Expander flows, geometric embeddings and graph partitioning," *Journal of the ACM (JACM)*, vol. 56, no. 2, p. 5, 2009.