Improved Region-Growing and Combinatorial Algorithms for k-Route Cut Problems

Guru Guruganesh[*] Laura Sanita[†] Chaitanya Swamy[‡]

Abstract

We study the *k-route* generalizations of various cut problems, the most general of which is *k-route multicut* (k-MC) problem, wherein we have r source-sink pairs and the goal is to delete a minimum-cost set of edges to reduce the edge-connectivity of every source-sink pair to below k. The k-route extensions of multiway cut (k-MWC), and the minimum s-t cut problem (k-(s,t)-Cut), are similarly defined. We present various approximation and hardness results for k-MC, k-MWC, and k-(s,t)-Cut that improve the state-of-the-art for these problems in several cases. Our contributions are threefold.

- For *k-route multiway cut*, we devise simple, but surprisingly effective, combinatorial algorithms that yield bicriteria approximation guarantees that markedly improve upon the previous-best guarantees.
- For *k-route multicut*, we design algorithms that improve upon the previous-best approximation factors by roughly an $O(\sqrt{\log r})$-factor, when $k = 2$, and for general k and unit costs and any fixed violation of the connectivity threshold k. The main technical innovation is the definition of a new, powerful *region growing* lemma that allows us to perform region-growing in a recursive fashion even though the LP solution yields a *different metric* for each source-sink pair, and *without incurring an $O(\log^2 r)$ blow-up* in the cost that is inherent in some previous applications of region growing to k-route cuts. We obtain the same benefits as [15] do in their divide-and-conquer algorithms, and thereby obtain an $O(\ln r \ln\ln r)$-approximation to the cost. We also obtain some extensions to k-route node-multicut problems.
- We complement these results by showing that the *k-route s-t cut* problem is at least as hard to approximate as the *densest-k-subgraph* (DkS) problem on uniform hypergraphs. In particular, this implies that one cannot avoid a $\mathrm{poly}(k)$-factor if one seeks a unicriterion approximation, without improving the state-of-the-art for DkS on graphs, and proving the existence of a family of one-way functions. Previously, only *NP*-hardness of k-(s,t)-Cut was known.

1 Introduction

The problem of finding minimum size cuts for a given graph has a rich history in the field of combinatorial optimization, with a wide range of applications in logistics, transportation and telecommunication systems. One key problem of interest is that of disconnecting a given set of node pairs in a network by removing edges at minimum cost. Formally, in the *multicut* problem, we are given an undirected graph $G = (V,E)$ with nonnegative edge costs $\{c_e\}_{e\in E}$ and pairs of nodes $(s_1,t_1),\ldots,(s_r,t_r)$ called source-sink pairs or commodities, and we seek a minimum-cost set of edges whose removal disconnects every s_i-t_i pair. Two special cases of this problem have by themselves attracted widespread attention: (i) the celebrated *minimum*

[*] `ggurugan@cs.cmu.edu`. Department of Computer Science, Carnegie Mellon University, Pittsburgh, USA. Work done while the author was an undergraduate research assistant at the University of Waterloo under Prof. Sanita.

[†] `sanita@uwaterloo.ca`. Dept. of Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1.

[‡] `cswamy@math.uwaterloo.ca`. Dept. of Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1. Supported in part by NSERC grant 327620-09, an NSERC Discovery Accelerator Supplement Award and an Ontario Early Researcher Award.

s-t *cut* problem, which is the special case when $r = 1$; and (ii) the *multiway cut* problem [12], where every pair of nodes from a given set $T \subseteq V$ of terminals forms a commodity. These cut problems and their variants have been widely studied in terms of hardness and approximation (see, e.g., [36, 37]), have numerous direct applications (e.g., identifying bottlenecks in a network), and algorithms for them serve as important primitives in the design of divide-and-conquer algorithms (see, e.g., [15, 34]) and find application in diverse settings such as image segmentation, VLSI design and network routing (see, e.g., [27, 29, 6, 32]).

We study a natural generalization of the above cut problems motivated by the fact that in various settings, we are not interested in a complete disconnection of our terminals but rather in reducing their connectivity below a certain threshold. Specifically, in the *k-route multicut* (k-MC) problem, the input is a multicut instance and an integer $k \geq 1$; the goal is to find a minimum-cost set $F \subseteq E$ of edges so that there are at most $(k-1)$ edge-disjoint s_i-t_i paths in $(V, E \setminus F)$ for all $i = 1, \ldots, r$. We define the *k-route multiway cut* (k-MWC), and the *k-route (s,t)-cut* (k-(s,t)-Cut) problems analogously.

The study of k-route cut problems can be motivated from various perspectives. One motivation comes from the fact that k-route cuts are dual objects to *k-route flows* [21], which can be seen as a robust or fault-tolerant version of flows where we seek to send traffic along tuples of k edge-disjoint paths. A k-route cut establishes an upper bound on the value (suitably defined) of the maximum k-route flow, and can thus be seen as identifying the bottleneck in a network when we seek a certain level of robustness. k-route cut problems can also be directly motivated as abstracting the problem of an attacker who seeks to reduce connectivity in a given network while incurring minimum cost. Viewed from this perspective, k-route cut problems are closely related to *network interdiction* problems, which typically consider the complementary objective of minimizing source-sink connectivity subject to a budget constraint on the edge-removal cost [30, 38, 39].

The k-route cut problems are at least as hard as their 1-route counterparts. Multicut and multiway cut are *APX*-hard [12], with the former not admitting any constant-factor approximation assuming the unique-games conjecture [9], and k-(s,t)-Cut is *NP*-hard; hence, we focus on approximation algorithms. Moreover, as highlighted in [10, 5, 23, 11, 22], k-route cut problems turn out to be much more challenging than their 1-route counterparts, especially for non-constant k, so (as in [11]) we consider bicriteria approximation guarantees. (This is further justified by our hardness result for k-(s,t)-Cut in Section 4.) We say that a solution $F \subseteq E$ is an (α, β)-approximation for the given k-MC instance if $\sum_{e \in F} c_e$ is at most β times the optimal value, and $(V, E \setminus F)$ contains at most $\alpha(k-1)$ edge-disjoint s_i-t_i paths for all $i = 1, \ldots, r$.

Our results. We develop various approximation and hardness results for k-MC, k-MWC and k-(s,t)-Cut that improve upon the current-best approximation and hardness results in several cases.

In Section 2, we consider the k-route multiway cut problem. We devise an $\big(O(1), O(1)\big)$-approximation for k-MWC with unit costs (Theorem 2.2), and an $\big(O(1), O(\log r)\big)$-approximation with general costs (Theorem 2.4), where $r = |T|$. The previous-best guarantees for k-MWC (for general k) are those that follow from the results of Chuzhoy et al. [11] for k-MC, namely, an $\big(O(1), O(\log^{1.5} r)\big)$-approximation for unit costs and an $\big(O(\log r), O(\log^3 r)\big)$-approximation for general costs. Thus, our guarantees constitute a significant improvement in the state-of-the-art for k-MWC. We also show that the special case where $T = V$, which we call *k-route all-pairs cut*, is *APX*-hard for $k \geq 3$ (Appendix A). (For $k = 1, 2$, it is easy to see that all-pairs k-route cut is polytime solvable.)

In Section 3, we design algorithms for the k-route multicut problem. We achieve approximation ratios of $O(\ln r \ln\ln r)$ for 2-MC, and $\big(\gamma, O(\frac{\gamma}{(\sqrt{\gamma}-1)^2} \ln r \ln\ln r)\big)$ for k-MC with unit costs. In contrast, Chuzhoy et al. [11] obtain approximation ratios of $O(\log^{1.5} r)$ for 2-MC, and $\big(\gamma, O(\frac{\log^{1.5} r}{\min\{1, \gamma-1\}})\big)$ for k-MC with unit costs. Thus, for any fixed γ (i.e., independent of k and r), our results improve upon the previous-best guarantees for these cases in [11] by roughly an $O(\sqrt{\log r})$-factor. (Setting $\gamma = \frac{k}{k-1}$, our guarantee and the one in [11] become unicriterion approximations that are incomparable.) In contrast to the algorithms in [11], which rely on approximations to suitable variants of sparsest cut, we devise *rounding* algorithms for a natural LP-relaxation for k-MC, and our guarantees therefore also translate to integrality-gap results. In

Section 5, we consider some extensions to k-route *node-multicut* problems.

Complementing the above results, we show in Section 4 that k-(s,t)-Cut is at least as hard as the densest k-subgraph (DkS) problem: a ρ-approximation for k-(s,t)-Cut yields a $(2\rho^{\lambda})$-approximation for DkS on λ-uniform hypergraphs (Theorem 4.2). The latter problem is hard to approximate within an n^{ϵ_0}-factor, for some constant ϵ_0, for all $\lambda \geq 3$, unless a certain family of one-way functions exists [3]. This implies that obtaining a unicriterion $O\big(k^{\epsilon_0}\,\text{polylog}(n)\big)$-approximation (even) for k-(s,t)-Cut for some constant ϵ_0 would improve the state-of-the-art for the notoriously hard densest k-subgraph problem on graphs, and imply the existence of certain one-way functions. Previously, only *NP*-hardness of k-(s,t)-Cut was known, as a consequence of the fact that certain *NP*-hard unbalanced graph partitioning problems [20, 28] can be cast as special cases of k-(s,t)-Cut.

Our techniques. Our algorithms for k-MWC are combinatorial, and rely on the following simple, but quite useful observation: if $F \subseteq E$ is feasible, then $\overline{G} = (V, E \setminus F)$ has a multiway cut with at most $(k-1)(r-1)$ edges (Claim 2.1). Using this, we show that we can identify a terminal $t_i \in T$, a t_i-isolating cut, and a set of edges of cost $O\big(\frac{OPT}{|T|}\big)$ whose removal causes the t_i-isolating cut to have $O(k-1)$ edges. We include these edges, drop t_i from T, and repeat, which naturally yields an $O(\log r)$-approximation in the cost. The improvement for unit costs stems from the stronger property that either the minimum multiway-cut in G has cost $O(OPT)$, or there is some t_i-isolating cut of value $O(k-1)$; thus, we may now drop terminals incurring *zero cost*, which results in an improved $O(1)$ cost-approximation.

Interestingly, [5] use a similar approach to obtain an $\big(O(1), O(1)\big)$-approximation for single-source k-MC with unit costs and they remark that such an approach is unlikely to work for k-MWC because there are examples where every pair of terminals is $2(k-1)$-edge connected but the optimal multiway cut value is $\Omega(r) \cdot OPT$. Thus, a useful insight to emerge from our work is that whereas a 2-factor violation in the pairwise terminal connectivity does not ensure that the multiway cut value is $O(OPT)$, a $(2+\varepsilon)$-factor violation in connectivity does, for any $\varepsilon > 0$.

Our algorithms for k-MC are based on rounding an optimal solution to a natural LP-relaxation of the problem. This is technically the most sophisticated part of the paper. The main technique that we use is *region growing*. The idea is to view the LP solution as a metric, grow a suitable ball in this metric and prove a region-growing lemma showing that the cost of the ball-boundary edges can be charged to the ball-volume, where volume measures the contribution to the LP objective from the edges inside the ball. This was introduced by [27, 19] in the context of the sparsest cut and multicut problems, and Even et al. [15], building upon the work of Seymour [33], extended the technique to obtain improved guarantees for various divide-and-conquer algorithms that involve recursive applications of region growing. However, in contrast with various applications of region growing considered in [27, 19, 16, 15], the difficulty in the k-route multicut problem stems from the fact that an LP-solution yields a *different metric for each source-sink pair* instead of a single common metric that can be applied in the region-growing process. (In particular, k-MC does not fall into the divide-and-conquer framework of Even et al. [15].) Although [5, 23, 22] adapted the region-growing lemma in [27, 19] to the 2-route, 3-route, and the k-route single-source settings, their approach seems incapable of obtaining any thing better than an $O(\log^2 r)$-approximation—one loses one $\log$-factor due to region growing and another due to recursion—which is worse than the guarantees in [11]. (In fact, [11] abandoned the region-growing approach and used a greedy set-cover strategy to obtain their improvements over [5, 23, 22].)

Our chief technical novelty is to prove a region-growing lemma (see Lemmas 3.1 and 3.3) applicable to settings with different metrics, that is inspired by, but more general, than the analogous lemma in [15], and much more sophisticated than the one used in [5, 23, 22]. This lemma, coupled with a subtle insight about the metrics derived from the LP solution, allows us to obtain the same kind of savings in our recursive region-growing algorithm that Even et al. [15] obtain (via their region-growing lemma) in their divide-and-conquer algorithms; this yields our improved approximation guarantees. We believe that our region-growing lemma and its application in the context of different metrics are tools of independent interest that will find further application in the study of cut problems.

The hardness proof for k-(s,t)-Cut dovetails the hardness proof in [11] for the *vertex-connectivity* version of k-(s,t)-Cut (where we want to decrease the *s-t vertex connectivity* to below k), who reduce from the *small-set vertex expansion* (SSVE) problem, which they show is DkS-hard. We observe that this reduction immediately implies the same hardness for k-(s,t)-Cut on a *directed graph*, and combine this with a useful trick from [8] that allows us to move from digraphs to undirected graphs. The idea is to take the digraph used in the hardness proof, remove edge directions, and add some extra nodes and expensive edges so that the *residual digraph* obtained after sending a partial s-t flow along the expensive edges essentially coincides with the digraph used in the hardness proof.

Related work. Standard (i.e., 1-route) cut problems have been extensively studied; we refer the reader to the textbooks [2, 36, 37] for more information.

The study of k-route flow and k-route cut problems was initiated by Kishimoto [21], and has since received much attention in the theoretical Computer Science community [7, 10, 5, 23, 22, 11]. Bruhn et al. [7] gave a $2(k-1)$-approximation for single-source k-MC with unit costs, whereas [10, 5, 23] obtained efficient polylogarithmic approximation results for k-MC with small values of k. Subsequently, Chuzhoy et al. [11] obtained the first non-trivial results for k-MC with arbitrary k in the form of bicriteria approximation guarantees. Independently, Kolman and Scheideler [22] obtained an $O\big(\exp(k)\,\mathrm{polylog}(r)\big)$-approximation for single-source k-MC (with general costs). As shown by our hardness result for k-(s,t)-Cut in Section 4, the move to bicriteria approximations is necessary unless one incurs a $\mathrm{poly}(k)$-factor in the approximation.

As noted earlier, k-route cut problems and *network interdiction* problems (see, e.g., [30, 38, 39, 13] and the references therein) can be viewed as complementary problems. For instance, in the *maximum-flow interdiction problem* (MFIP) we are given *edge capacities* in addition to edge costs, and the goal is to minimize the maximum s-t flow in the graph remaining after removing edges of total cost at most a given budget. MFIP with unit capacities is thus complementary to k-(s,t)-Cut, and bicriteria guarantees for one translate to the other. Unit-capacity MFIP is known to be polytime solvable for planar graphs [30, 39]. Dinitz and Gupta [13] propose a general framework for attacking *packing interdiction* problems. However, their results do not quite apply to MFIP (since phrasing max-flow in terms of edge-flows destroys the packing property, and phrasing it in terms of path-flows yields an interdiction problem where one removes paths).

2 A simple combinatorial algorithm for k-route multiway cut

Recall that in the k-route multiway cut (k-MWC) problem, we are given a set $T = \{t_1, \ldots, t_r\} \subseteq V$ of terminals and we seek to remove a minimum-cost set of edges so that the edge-connectivity between any two terminals is less than k. The case $k = 1$ is the multiway cut problem, which is known to be *APX*-hard [12] even with unit edge costs. We devise an $\big(O(1), O(1)\big)$-approximation for k-MWC with unit costs, and an $\big(O(1), O(\log r)\big)$-approximation with general costs. These improve upon the previous-best guarantees (for general k) of $\big(O(1), O(\log^{1.5} r)\big)$ for unit costs, and $O\big(O(\log r), O(\log^3 r)\big)$ for general costs due to [11]. Remark 2.5 shows that our guarantees also translate to integrality-gap bounds for a suitable LP-relaxation.

Let O^* denote the optimal set of edges, and let $\overline{G} = (V, E \setminus O^*)$ be the remainder graph. Let $k' = k-1$. Our algorithms are quite easy to describe and analyze. We first prove a simple claim about $\overline{G}$.

Claim 2.1 *There is a set $\overline{E}$ of edges of $\overline{G}$ with $|\overline{E}| \le k'(r-1)$ such that $O^* \cup \overline{E}$ is a multiway cut in G.*

Proof : Compute a minimum t_1-t_2 cut F in $\overline{G}$, where F is a set of edges. By the definition of $\overline{G}$, $|F| \le k'$. Removing F from $\overline{G}$ creates at least two components. We can now recurse in each connected component, and after computing at most $r-1$ min cuts, each terminal will be in a different connected component. ■

The idea behind the algorithm for unit costs is the following. Claim 2.1 shows that the optimal multiway cut in G would be a good approximation to k-MWC if $|\overline{E}| = O\big(|O^*|\big)$. Otherwise, there is a multiway cut in G of cost $O(k'r)$, and so there is some terminal (in fact $\Omega(r)$ terminals) that has an isolating cut in G of size $O(k')$; we simply remove this terminal from T and repeat this process.

Theorem 2.2 *There is a $\big(\gamma, \frac{2\gamma}{\gamma-2}\big)$-approximation algorithm for k-MWC with unit costs for any $\gamma > 2$.*

Proof : For all i, compute a minimum t_i-isolating cut F_i. It is well known that, even with non-unit costs, $\sum_{i=1}^{r} c(F_i)$ is a 2-approximation to the minimum multiway cut [12]. In particular, $C = \sum_{i=1}^{r} |F_i| \leq 2|O^*| + 2|\overline{E}|$ (by Claim 2.1). If $C \geq \gamma k'r$, then we have $|O^*| \geq \frac{C}{2} - k'r \geq C\big(\frac{1}{2} - \frac{1}{\gamma}\big)$, so taking the union of the F_is yields a $\frac{2\gamma}{\gamma-2}$-approximation. Otherwise, there is some t_i such that $|F_i| < \gamma k'$, so we can simply remove t_i from T and decrease r, and repeat.

We remark that the number of iterations can be reduced to $\log_2 r$ at the expense of increasing the connectivity to $2\gamma k'$, since must be at least $r/2$ terminals such that $|F_i| \leq 2\gamma k'$. ∎

Remark 2.3 The condition $\gamma > 2$ above is tight. This follows from an example in [5] where every pair of terminals is $2k'$-edge connected but the minimum multiway cut yields an $\Omega(r)$-approximation.

To generalize this algorithm to general edge costs, assume for now that we know $OPT = c(O^*)$. Unlike in the unit edge-cost case where we could make progress by dropping terminals while incurring zero cost, here we will need to incur cost $O\big(\frac{OPT}{r}\big)$ to drop a terminal (or incur cost $O(OPT)$ to drop $r/2$ terminals). This naturally leads to an $O(\log r)$-approximation in the cost. Let $H_r := 1 + \frac{1}{2} + \ldots + \frac{1}{r} = O(\log r)$.

Theorem 2.4 *There is a $\big(\gamma, \frac{2\gamma}{\gamma-2} H_r\big)$-approximation algorithm for k-MWC with general edge costs, for any $\gamma > 2$.*

Proof : Let T' initialized to T denote the current terminal set and $r' \leftarrow r$. Let F initialized to $\emptyset$ denote the set of edges removed. Let $\alpha = \frac{2}{\gamma-2}$. While $|T'| > 1$, we do the following. Set $c'_e = \min\{c_e, \frac{\alpha OPT}{k'r'}\}$ for every edge e. Note that the c'-cost of the minimum multiway cut is at most $c'(O^* \cup \overline{E}) \leq OPT + k'r' \cdot \frac{\alpha OPT}{k'r'}$. For every terminal $t \in T'$, compute a minimum c'-cost t-isolating cut F_t. Then, we have $\sum_{t \in T'} c'(F_t) \leq 2(1+\alpha)OPT$. So there is some $t \in T'$ such that $c'(F_t) \leq \frac{2(1+\alpha)OPT}{r'}$. The number of edges in F_t with $c_e > \frac{\alpha OPT}{k'r'}$ is less than $\frac{2(1+\alpha)}{\alpha} k' = \gamma k'$. We add edges in F_t with $c_e \leq \frac{\alpha OPT}{k'r'}$ to F. This incurs cost at most $c'(F_t) \leq \frac{2(1+\alpha)OPT}{r'}$, and ensures that t is less than $\gamma k'$ connected to every other terminal in T' in the remaining graph. We now set $T' \leftarrow T' \setminus \{t\},\ r' \leftarrow r' - 1$, and repeat the above process.

Clearly, every pair of terminals is at most $\gamma k'$ connected in $(V, E \setminus F)$. Also, $c(F) \leq \sum_{r'=r}^{1} \frac{2(1+\alpha)OPT}{r'} = OPT \cdot \frac{2\gamma}{\gamma-2} \cdot H_r$.

Finally, we can eliminate the need for knowing OPT as follows. Given a guess C of OPT, if at some iteration we have $\sum_{t \in T'} c'(F_t) > 2(1+\alpha)C$ then we know that $C < OPT$; otherwise, we obtain a solution of cost at most $C \cdot \frac{2\gamma}{\gamma-2} \cdot H_r$. So we can try powers of $(1+\epsilon)$ to find the smallest C such that the latter case happens; this blows up the approximation in cost by at most a $(1+\epsilon)$-factor. ∎

Remark 2.5 (LP-relative bounds) The guarantees in Theorems 2.2 and 2.4 also translate to integrality-gap bounds for the following LP-relaxation of k-MWC. Let $\mathcal{P}_{ij}$ be the collection of all t_i-t_j paths.

$$\min\ c^T x \quad \text{s.t.} \quad \sum_{e \in P} (x_e + y_e) \geq 1 \quad \forall t_i, t_j \in T,\ P \in \mathcal{P}_{ij}; \qquad \sum_{e} y_e \leq k'(r-1); \qquad x, y \geq 0. \quad \text{(P')}$$

Claim 2.1 implies that (P') is indeed a valid relaxation of k-MWC. Let (x, y) be an optimal solution to (P') and $OPT_{\text{P'}}$ be its value. Then, for any $\lambda \geq 0$, for the cost function $c'_e = \min\{c_e, \lambda\}$, there is a fractional multiway-cut of c'-cost at most $OPT_{\text{P'}} + \lambda \sum_e y_e$. Also, if F_i is a minimum c'-cost t_i-isolating cut then we have $\sum_i c'(F_i) \leq 2\big(OPT_{\text{P'}} + \lambda \sum_e y_e\big)$. (This follows since an optimal solution to the multiway-cut LP is known to be half-integral (see, e.g., [36]); this implies that 2(cost of an optimal solution) is at least $\sum_i$ (cost of a minimum t_i-isolating cut).) This implies that we can replace $|O^*|$ and $|\overline{E}|$ in the proof of Theorem 2.2 by $OPT_{\text{P'}}$ and $\sum_e y_e$, and OPT in the proof of Theorem 2.4 by $OPT_{\text{P'}}$, and all the arguments go through.

The all-pairs case. This is the special case of k-MWC where $T = V$. To our knowledge, this *k-route all-pairs cut* problem has not been explicitly studied before. When $k = 1$, the all-pairs problem is trivial as the remainder graph cannot contain any edge. When $k = 2$, this problem is still in *P* as the remainder graph is a maximum-cost spanning forest. We prove that the problem is *APX*-hard for all $k \geq 3$ (see Appendix A), thus resolving the complexity (with respect to polytime solvability) of k-route all-pairs cut. The all-pairs problem can also be stated in terms of properties required of the remainder graph. For example, in 3-route all-pairs cut, we seek a minimum-cost edge set such that the remainder graph does not contain a *diamond* as a minor. Interestingly, this is equivalent to requiring that the remainder graph be a maximum-weight *cactus*, which is a graph where every edge lies in at most one cycle. As noted above, this problem is *APX*-hard. But we observe that this problem admits an $O(1)$-approximation as a consequence of the results of Fiorini et al. [17]; see Appendix A.

3 A region-growing algorithm for k-route multicut

We now consider general k-route multicut (k-MC): given source-sink pairs/commodities $(s_1, t_1), \ldots, (s_r, t_r)$, we want to find a minimum-cost set of edges whose removal reduces the s_i-t_i edge connectivity to less than k for all $i = 1, \ldots, r$. We consider the following LP-relaxation of the problem, which was also considered by Barman and Chawla [5]. Throughout e indexes the edges in E, and i indexes the commodities. Let $\mathcal{P}_i$ denote the collection of all (simple) s_i-t_i paths in G.

$$\min \sum_e c_e x_e \quad \text{s.t.} \quad \sum_{e \in P} (x_e + y^i_e) \geq 1 \quad \forall i, \forall P \in \mathcal{P}_i; \qquad \sum_e y^i_e \leq k - 1 \quad \forall i; \qquad x, y \geq 0. \tag{P}$$

Let $\big(x, \{y^i\}\big)$ denote an optimal solution to (P), and OPT be its value. We show that $\big(x, \{y^i\}\big)$ can be rounded to yield an $O(\ln r \ln\ln r)$-approximation when $k = 2$, and a bicriteria $\big(\gamma, O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2} \ln r \ln\ln r\big)\big)$-approximation for k-MC with unit edge costs, for any $\gamma > 1$. Notably, our cost-approximation is with respect to OPT, and so they translate to integrality-gap upper bounds for (P). Our results improve upon (for any fixed γ) the previous-best guarantees for these cases in [11] by roughly a $\sqrt{\log r}$-factor.

3.1 Region-growing lemmas

The main tool that we leverage is *region growing* [27, 19, 15]. The idea is to view the LP solution as a metric, grow a suitable ball in this metric and prove that the cost of the ball-boundary edges can be charged to the ball-volume, where volume measures the contribution to the LP objective from the edges inside the ball. The main difficulty in applying this idea to (P) is that, unlike most applications of region growing [27, 19, 16, 15], the LP solution yields a *different metric* for each commodity. The key technical ingredient and novelty is a new region-growing lemma (see Lemmas 3.1 and 3.3) that is analogous to, but more general, than the one in [15], and much more sophisticated than the one used in [5, 23, 22]. Roughly speaking, we prove that given a current set S of nodes, one can construct a ball around any s_i in the $(x + y^i)$-metric such that the cost of the "boundary x-edges" can be charged to (x-volume of S) $\cdot \ln\big(x\text{-volume of } S / x\text{-volume of the ball}\big) \cdot$

$\ln\ln r$ (Lemma 3.3). A subtle insight that helps deal with the complication that different applications of region growing involve different commodities and therefore different metrics is the following. Since the x-contribution is *common* to all $(x+y^i)$-metrics, even though we consider different commodity-metrics we can leverage the above guarantee and obtain the same kind of savings that [15] obtain in their divide-and-conquer algorithms (see Lemma 3.5); this leads to our improved approximation guarantees.

We now state our region-growing lemmas in a general form and then apply these to the optimal solution $\big(x,\{y^i\}\big)$ to obtain various useful corollaries. In Section 5, we extend our arguments to prove region-growing lemmas in settings that involve both edge and node lengths. Let $n=|V|$, $m=|E|$. Let $\ell : V\times V \mapsto \mathbb{R}_{\geq 0}$ be a metric on $V\times V$. Our algorithm will iteratively focus on certain regions of the graph G. Let $S\subseteq V$, which is intended to represent the node-set of the current region. Let $F\subseteq E$, which is intended to represent the edges that contribute to the volume, and whose cost we care about. Let $\beta\geq 0$. Let $z\in V$, and $\rho\geq 0$.

- Define $B_\ell(z,\rho) := \{v\in V : \ell_{zv}\leq \rho\}$ to be the ball of radius ρ around z.
- Let $B^S_\ell(z,\rho) := B_\ell(z,\rho)\cap S$ and $\overline{B^S_\ell}(z,\rho) := S\setminus B_\ell(z,\rho)$.
- Define the following volumes:

$$\mathcal{V}^{S,F}_\ell(\beta;z,\rho) \;:=\; \beta + \sum_{(u,v)\in F: u,v\in B^S_\ell(z,\rho)} c_{uv}\ell_{uv} + \sum_{\substack{(u,v)\in F: u\in B^S_\ell(z,\rho)\\ v\in \overline{B^S_\ell}(z,\rho)}} c_{uv}\big(\rho-\ell_{zu}\big)$$

$$\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,\rho) \;:=\; \beta + \sum_{(u,v)\in F: u,v\in \overline{B^S_\ell}(z,\rho)} c_{uv}\ell_{uv} + \sum_{\substack{(u,v)\in F: u\in B^S_\ell(z,\rho)\\ v\in \overline{B^S_\ell}(z,\rho)}} c_{uv}\big(\ell_{zv}-\rho\big)$$

- For a subset $T\subseteq S$, let $\delta^S_F(T)$ denote $\{(u,v)\in F : u\in T, v\in S\setminus T\}$.
- Define $\partial^{S,F}_\ell(z,\rho) := \delta^S_F\big(B^S_\ell(z,\rho)\big) = \delta^S_F\big(\overline{B^S_\ell}(z,\rho)\big)$.

When $F=E$, we drop F from the above pieces of notation (e.g., $\delta^S_E(T)$ is shortened to $\delta^S(T)$). For $H\subseteq E$, we use $\ell(H)$ to denote $\sum_{e\in H}\ell_e$.

Lemma 3.1 (Region-growing lemma) *Let $F\subseteq E$, $S\subseteq V$, $z\in V$, and $0\leq a<b$. Let ρ be chosen uniformly at random from $[a,b)$. Then,*

$$\mathrm{E}_\rho\left[\frac{c\big(\partial^{S,F}_\ell(z,\rho)\big)}{\mathcal{V}^{S,F}_\ell(\beta;z,\rho)\ln\Big(\frac{e\mathcal{V}^{S,F}_\ell(\beta;z,b)}{\mathcal{V}^{S,F}_\ell(\beta;z,\rho)}\Big)}\right] \;\leq\; \frac{1}{b-a}\cdot\ln\ln\left(\frac{e\mathcal{V}^{S,F}_\ell(\beta;z,b)}{\mathcal{V}^{S,F}_\ell(\beta;z,a)}\right) \quad \textit{and} \tag{1}$$

$$\mathrm{E}_\rho\left[\frac{c\big(\partial^{S,F}_\ell(z,\rho)\big)}{\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,\rho)\ln\Big(\frac{e\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,a)}{\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,\rho)}\Big)}\right] \;\leq\; \frac{1}{b-a}\cdot\ln\ln\left(\frac{e\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,a)}{\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,b)}\right) \tag{2}$$

Proof : We abbreviate $c\big(\partial^{S,F}_\ell(z,\rho)\big)$ to $c(\rho)$, $\mathcal{V}^{S,F}_\ell(\beta;z,\rho)$ to $\mathcal{V}(\rho)$ and $\overline{\mathcal{V}^{S,F}_\ell}(\beta;z,\rho)$ to $\overline{\mathcal{V}}(\rho)$. Let x^- be a value infinitesimally smaller than x. Let $I=\{\ell(s_i,v) : v\in V\}$. Note that $\mathcal{V}(\rho)$ and $\overline{\mathcal{V}}(\rho)$ are differentiable at all $\rho\in[a,b)\setminus I$ and for each such ρ, we have $\frac{d\mathcal{V}(\rho)}{d\rho}=c(\rho)$ and $\frac{d\overline{\mathcal{V}}(\rho)}{d\rho}=-c(\rho)$. Let $a_0=a$, $a_k=b$, and

$\{a_1, \ldots, a_{k-1}\} = (a, b) \cap I$ with $a_1 < \ldots < a_{k-1}$. Then,

$$
\begin{aligned}
(b-a) \cdot \mathrm{E}_\rho \left[\frac{c(\rho)}{\mathcal{V}(\rho) \ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(\rho)}\right)} \right] &= \sum_{i=1}^{k} \int_{a_{i-1}}^{a_i^-} \frac{d\mathcal{V}(\rho)}{\mathcal{V}(\rho) \ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(\rho)}\right)} = \sum_{i=1}^{k} \int_{a_{i-1}}^{a_i^-} \frac{d(\ln \mathcal{V}(\rho))}{\ln\left(e\mathcal{V}(b)\right) - \ln(\mathcal{V}(\rho))} \\
&= \sum_{i=1}^{k} - \ln\ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(\rho)}\right)\Big|_{a_{i-1}}^{a_i^-} = \sum_{i=1}^{k} \left[\ln\ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(a_{i-1})}\right) - \ln\ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(a_i^-)}\right) \right] \leq \ln\ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(a)}\right).
\end{aligned}
$$

The final inequality follows since $\mathcal{V}(\rho)$ increases with ρ, and $\ln\ln\left(\frac{e\mathcal{V}(b)}{\mathcal{V}(\rho)}\right)$ decreases with ρ. Similarly, we obtain that $(b-a) \cdot \mathrm{E}_\rho \left[\frac{c(\rho)}{\overline{\mathcal{V}}(\rho) \ln\left(\frac{e\overline{\mathcal{V}}(a)}{\overline{\mathcal{V}}(\rho)}\right)} \right] = \sum_{i=1}^{k} \ln\ln\left(\frac{e\overline{\mathcal{V}}(a)}{\overline{\mathcal{V}}(\rho)}\right)\Big|_{a_{i-1}}^{a_i^-} \leq \ln\ln\left(\frac{e\overline{\mathcal{V}}(a)}{\overline{\mathcal{V}}(b)}\right)$. ■

Corollary 3.2 *Let $F, H \subseteq E$, $S \subseteq V$, $z \in V$, and $0 \leq a < b$. For any $\alpha \in (0,1)$, we can efficiently find a radius $\rho_1 \in [a, b)$ such that*

$$
\begin{aligned}
c\left(\partial_\ell^{S,F}(z, \rho_1)\right) &\leq \frac{2}{(1-\alpha)(b-a)} \cdot \mathcal{V}_\ell^{S,F}(\beta; z, \rho_1) \cdot \ln\left(\frac{e\mathcal{V}_\ell^{S,F}(\beta;z,b)}{\mathcal{V}_\ell^{S,F}(\beta;z,\rho_1)}\right) \cdot \ln\ln\left(\frac{e\mathcal{V}_\ell^{S,F}(\beta;z,b)}{\mathcal{V}_\ell^{S,F}(\beta;z,a)}\right) && (3) \\
c\left(\partial_\ell^{S,F}(z, \rho_1)\right) &\leq \frac{2}{(1-\alpha)(b-a)} \cdot \overline{\mathcal{V}_\ell^{S,F}}(\beta; z, \rho_1) \cdot \ln\left(\frac{e\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,a)}{\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,\rho_1)}\right) \cdot \ln\ln\left(\frac{e\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,a)}{\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,b)}\right) && (4) \\
\left|\partial_\ell^{S,H}(z, \rho_1)\right| &< \frac{\ell(H)}{\alpha(b-a)}. && (5)
\end{aligned}
$$

Proof : Suppose we pick ρ uniformly at random from $[a, b)$. Define the following events.

$$
\begin{aligned}
\Omega &:= \left\{\rho \in [a,b) : \left|\partial_\ell^{S,H}(z,\rho)\right| \geq \frac{\ell(H)}{\alpha(b-a)}\right\} \\
\Omega_1 &:= \left\{\rho \in [a,b) : c\left(\partial_\ell^{S,F}(z,\rho)\right) > \frac{2}{(1-\alpha)(b-a)} \cdot \mathcal{V}_\ell^{S,F}(\beta; z, \rho) \ln\left(\frac{e\mathcal{V}_\ell^{S,F}(\beta;z,b)}{\mathcal{V}_\ell^{S,F}(\beta;z,\rho)}\right) \cdot \ln\ln\left(\frac{e\mathcal{V}_\ell^{S,F}(\beta;z,b)}{\mathcal{V}_\ell^{S,F}(\beta;z,a)}\right)\right\} \\
\Omega_2 &:= \left\{\rho \in [a,b) : c\left(\partial_\ell^{S,F}(z,\rho)\right) > \frac{2}{(1-\alpha)(b-a)} \cdot \overline{\mathcal{V}_\ell^{S,F}}(\beta; z, \rho) \ln\left(\frac{e\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,a)}{\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,\rho)}\right) \cdot \ln\ln\left(\frac{e\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,a)}{\overline{\mathcal{V}_\ell^{S,F}}(\beta;z,b)}\right)\right\}
\end{aligned}
$$

For an edge $(u,v) \in E$, $\Pr\left[(u,v) \in \partial_\ell^{S,H}(z,\rho)\right] \leq \frac{\ell_{uv}}{b-a}$. Hence, $\mathrm{E}_\rho\left[\left|\partial_\ell^{S,H}(z,\rho)\right|\right] \leq \frac{\ell(H)}{b-a}$, and therefore $\Pr[\Omega] \leq \alpha$. By Lemma 3.1 and Markov's inequality, we have that $\Pr[\Omega_1], \Pr[\Omega_2] < (1-\alpha)/2$. Conditioning on $\Omega^c := [a,b) \setminus \Omega$ increases the probability of an event by at most a factor $\frac{1}{1-\Pr[\Omega]} \leq \frac{1}{1-\alpha}$, so $\Pr[\Omega_1|\Omega^c], \Pr[\Omega_2|\Omega^c] < 1/2$. Thus, $\Pr[\Omega^c \cap \Omega_1^c \cap \Omega_2^c] > 0$.

We argue that Ω^c, Ω_1^c, and Ω_2^c are all unions of at most n subintervals of $[a,b)$, and we can find these efficiently. Since $\Pr[\Omega^c \cap \Omega_1^c \cap \Omega_2^c] > 0$, we can then efficiently find an interval contained in $\Omega^c \cap \Omega_1^c \cap \Omega_2^c$ (in fact, a non-singleton interval), and hence, find $\rho_1 \in [a,b]$ satisfying (3)–(5) (in fact, there are infinitely many such ρ).

There are at most n distinct sets $B_\ell^S(z,\rho)$ that one may encounter as ρ varies in $[a,b)$. For each such set A, there is an interval $[lb, ub)$ such that $A = B_\ell^S(z,\rho)$ for all $\rho \in [lb, ub)$. Note that the right-hand-sides (RHSs) of (3) and (4) are continuous and differentiable in (lb, ub), and are monotonic (increasing and decreasing, respectively) functions of ρ. We call $[lb, ub)$ a smooth subinterval of $[a,b)$. By definition, the left-hand-sides (LHSs) of (3)–(5) are invariant over a smooth subinterval. Hence, Ω^c is the union of some smooth subintervals. Consider a smooth subinterval $[lb, ub)$. By continuity, if some $\rho \in [lb, ub)$ satisfies (3) or (4), then we can efficiently find the maximal subinterval of $[lb, ub)$ (which may be a singleton interval) such that all ρ in the subinterval satisfy the given bound. Hence, both Ω_1^c and Ω_2^c are the union of at most n subintervals of $[a, b)$. By trying out the at most $3n$ possible subintervals of $\Omega_1^c \cup \Omega_1^c \cup \Omega_2^c$, we can find some interval contained in $\Omega^c \cap \Omega_!^c \cap \Omega_2^c$ and hence obtain ρ_1 satisfying (3)–(5). ∎

Applications of the region-growing lemmas. To apply the above results to the metrics obtained from $\big(x, \{y^i\}\big)$, it will be convenient to modify G by subdividing every edge e into $r+1$ edges $e_0, e_1, \ldots, e_r$, and setting $x_f = x_e$ for $f = e_0$ and 0 otherwise, and $y_f^i = y_e^i$ if $f = e_i$ and 0 otherwise. We call e_0 an x-edge, and we call e_i, a y^i-edge for $i = 1, \ldots, r$. Clearly, any solution in G yields a solution in the subdivided graph of the same cost and vice versa, and this holds even for fractional solutions to (P). In the sequel, we work with the subdivided graph. To keep notation simple, we continue to use $G = (V, E)$ to denote the subdivided graph, and $\big(x, \{y^i\}\big)$ to denote the above solution in the subdivided graph. Let F be the set of all x-edges, and H^i be the set of all y^i-edges for all $i = 1, \ldots, r$. Consider a commodity i. Let ℓ^i denote the shortest-path metric of G (i.e., the subdivided graph) induced by the $\{x_e + y_e^i\}$ edge lengths. Set $\beta = OPT/r$. To avoid cumbersome notation, we shorten:

$B_{\ell^i}^S(z,\rho)$ and $\overline{B_{\ell^i}^S}(z,\rho)$	to	$B_i^S(z,\rho)$ and $\overline{B_i^S}(z,\rho)$ respectively
$\mathcal{V}_{\ell^i}^{S,F}(\beta; z,\rho)$ and $\overline{\mathcal{V}_{\ell^i}^{S,F}}(\beta; z,\rho)$	to	$\mathcal{V}_i^{S,x}(z,\rho)$ and $\overline{\mathcal{V}_i^{S,x}}(z,\rho)$ respectively
$\partial_{\ell^i}^{S,F}(z,\rho)$, $\partial_{\ell^i}^{S,H^i}(z,\rho)$, and $\partial_{\ell^i}^{S,E}(z,\rho)$	to	$\partial_i^{S,x}(z,\rho)$, $\partial_i^{S,y}(z,\rho)$, and $\partial_i^S(z,\rho)$ respectively

Also, define $\mathcal{V}^x(S) := \beta + \sum_{e \in E(S)} c_e x_e$, where $E(S)$ is the set of edges having both endpoints in S. Finally, for an integer $q \geq 1$ and a set A of edges let $c^q(A)$ be the cost of all but the $q-1$ most expensive edges of A (so $c^q(A) = 0$ if $|A| < q$).

Lemma 3.3 *Let $S \subseteq V$, $z \in V$, and i be some commodity. Let $\alpha \in (0,1)$ and $q = \lceil \frac{k-1}{\alpha} \rceil$. We can efficiently find $\rho_1 \in [0,1)$ such that*

$$c^q\big(\partial_i^S(z,\rho_1)\big) \leq \frac{2}{1-\alpha} \cdot \mathcal{V}_i^{S,x}(z,\rho_1) \ln\Big(\frac{e\mathcal{V}^x(S)}{\mathcal{V}_i^{S,x}(z,\rho_1)}\Big) \ln\ln\big(e(r+1)\big) \tag{6}$$

$$c^q\big(\partial_i^S(z,\rho_1)\big) \leq \frac{2}{1-\alpha} \cdot \overline{\mathcal{V}_i^{S,x}}(z,\rho_1) \ln\Big(\frac{e\mathcal{V}^x(S)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)}\Big) \ln\ln\big(e(r+1)\big). \tag{7}$$

Proof : We apply Corollary 3.2 taking $\ell = \ell^i$, $H = H^i$, and $[a, b) = [0, 1)$ (and S, z, α as given by the statement of the lemma, and F to be the set of x-edges). Note that

$$\mathcal{V}_i^{S,x}(z,1) \leq \mathcal{V}^x(S), \qquad \overline{\mathcal{V}_i^{S,x}}(z,0) \leq \mathcal{V}^x(S), \qquad \frac{\mathcal{V}_i^{S,x}(z,1)}{\mathcal{V}_i^{S,x}(z,0)} \leq r+1, \qquad \frac{\overline{\mathcal{V}_i^{S,x}}(z,0)}{\overline{\mathcal{V}_i^{S,x}}(z,1)} \leq r+1.$$

Thus, we obtain $0 \leq \rho_1 < 1$ such that $c\big(\partial_i^{S,x}(z,\rho_1)\big)$ satisfies the bounds given by the RHS of (6) and (7) respectively. Moreover, since $\ell^i(H^i) \leq k-1$, we have that $\big|\partial_i^{S,y}(z,\rho_1)\big| < \frac{k-1}{\alpha}$ due to (5), and so $\big|\partial_i^{S,y}(z,\rho_1)\big| < q$. Finally, since edges not in $F \cup H^i$ have zero ℓ^i-length, $\partial_i^{S,x}(z,\rho)$ and $\partial_i^{S,y}(z,\rho)$ partition $\partial_i^S(z,\rho)$ for all ρ. Therefore, $c\big(\partial_i^{S,x}(z,\rho_1)\big) \geq c^q\big(\partial_i^S(z,\rho_1)\big)$, and the lemma follows. ∎

Corollary 3.4 *Let $S \subseteq V$. Suppose that $s_i, t_i \in S$ and there are $\gamma(k-1)$ edge-disjoint s_i-t_i paths internal to S, where $\gamma > 1$. Suppose that $c_e = 1$ for all edges e. We can efficiently find $\rho_1 \in [0,1)$ such that*

$$c\big(\partial_i^S(s_i,\rho_1)\big) \leq \frac{2\gamma}{(\sqrt{\gamma}-1)^2} \cdot \mathcal{V}_i^{S,x}(s_i,\rho_1) \ln\Big(\frac{e\mathcal{V}^x(S)}{\mathcal{V}_i^{S,x}(s_i,\rho_1)}\Big) \ln\ln\big(e(r+1)\big)$$

$$c\big(\partial_i^S(s_i,\rho_1)\big) \leq \frac{2\gamma}{(\sqrt{\gamma}-1)^2} \cdot \overline{\mathcal{V}_i^{S,x}}(s_i,\rho_1) \ln\Big(\frac{e\mathcal{V}^x(S)}{\overline{\mathcal{V}_i^{S,x}}(s_i,\rho_1)}\Big) \ln\ln\big(e(r+1)\big)$$

Proof : We apply Lemma 3.3 with $\alpha \in (0, 1)$, whose value we will fix later, to obtain $\rho_1 \in [0, 1)$. Note that $B_i^S(s_i, \rho_1)$ is an s_i-t_i cut. Let $q = \lceil \frac{k-1}{\alpha} \rceil$. For any s_i-t_i cut $A \subseteq S$, we know that $c\big(\delta^S(A)\big) \geq \gamma(k-1)$. Therefore, since we have unit edge costs, $c\big(\delta^S(A)\big) \leq c^q\big(\delta^S(A)\big) + q - 1 \leq c^q\big(\delta^S(A)\big) \cdot \frac{\gamma}{\gamma - 1/\alpha}$. Plugging in the bounds for $c^q(.)$ from (6), (7), we see that the constant factor multiplying the volume terms on the RHS is $\frac{2\gamma}{(1-\alpha)(\gamma-1/\alpha)}$. This factor is minimized by setting $\alpha = \gamma^{-1/2}$, which yields the constant factor $\frac{2\gamma}{(\sqrt{\gamma}-1)^2}$ and completes the proof. ■

3.2 The rounding algorithms and their analyses

The case $k = 2$. The algorithm for $k = 2$ follows a similar template as the algorithm in [5] for 2-MC. However, its analysis resulting in our improved guarantee relies crucially on Lemma 3.3 which is derived from our stronger region-growing lemma. The algorithm proceeds as follows. Given a current set U of nodes and a current set of N source-sink pairs, we repeatedly use Lemma 3.3 to "carve out" disjoint regions $A_1, \ldots, A_h \subseteq U$ and build a set Z of edges until there is no 2-edge-connected source-sink pair in $U \setminus (A_1 \cup \ldots \cup A_h)$. Given $A_1, \ldots, A_{p-1}$, we obtain A_p as follows. We choose an s_i-t_i pair that is at least 2-edge connected in $S = U \setminus \bigcup_{q=1}^{p-1} A_q$, and use Lemma 3.3 with center s_i, $\alpha = 0.5$ and set S. Note that $2(k-1) = 2 = k$. We set A_p to be $B_i^S(s_i, \rho_1)$ or $\overline{B_i^S}(s_i, \rho_1)$, so as to ensure there are at most $N/2$ source-sink pairs inside A_p. We add the edges corresponding to $c^2\big(\delta^S(A_p)\big)$ to Z; Lemma 3.3 ensures that the cost of these edges can be bounded in terms of the volume contained in A_p. Having obtained $A_1, \ldots, A_h$ this way, we now recurse on each set A_p and the source-sink pairs contained in A_p, to obtain edge-sets $Z_1, \ldots, Z_h$. The solution we return is $Z \cup (Z_1 \cup \ldots \cup Z_h)$. A more formal description follows.

Algorithm 2-MCAlg$(U, \mathcal{T} = \{(s_1, t_1), \ldots, (s_N, t_N)\})$
Input: A subset $U \subseteq V$, and a collection $\mathcal{T} = \{(s_i, t_i)\}_{i=1}^N$ of source-sink pairs, where $s_i, t_i \in U$ for all $i = 1, \ldots, N$.
Output: A set $Z \subseteq E(U)$ such that s_i and t_i are at most 1-edge-connected in $(U, E(U) \setminus Z)$ for all $i = 1, \ldots, N$.

A1. Set $S = U$, $Z = \emptyset$, $\mathcal{S} = \emptyset$, and $\mathcal{T}' = \big\{(s_i, t_i) \in \mathcal{T} : s_i$ and t_i are at least 2-edge-connected in $(S, E(S))\big\}$.

A2. If $\mathcal{T}' = \emptyset$, return Z.

A3. While $\mathcal{T}' \neq \emptyset$, we do the following.

- A3.1 Pick some $(s_i, t_i) \in \mathcal{T}'$.
- A3.2 Apply Lemma 3.3 with $z = s_i$, $\alpha = 0.5$ and the set S to find a radius $0 \leq \rho_1 < 1$ satisfying (6), (7).
- A3.3 If $B_i^S(s_i, \rho_1)$ contains at most $N/2$ pairs from $\mathcal{T}$ then set $A = B_i^S(s_i, \rho_1)$, else set $A = \overline{B_i^S}(s_i, \rho_1)$.
- A3.4 Set $\mathcal{S} \leftarrow \mathcal{S} \cup \{A\}$. Add the edges contributing to $c^2\big(\delta^S(A)\big)$ (i.e., all edges of $\delta^S(A)$ except the most-expensive one) to Z.
- A3.5 Set $S \leftarrow S \setminus A$. Update $\mathcal{T}'$ to be the s_i-t_i pairs from $\mathcal{T}$ that are at least 2-edge-connected in $(S, E(S))$.

A4. For every set $A \in \mathcal{S}$, set $Z \leftarrow Z \cup$ 2-MCAlg$\big(A, \{(s_i, t_i) \in \mathcal{T} : s_i, t_i \in A\}\big)$.

A5. Return Z.

The initial call to 2-MCAlg, which computes the solution we return, is 2-MCAlg$\big(V, \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$.

Let $Z :=$ 2-MCAlg$\big(V, \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$. The feasibility of Z follows from the same arguments as in [5]; Lemma 3.7 gives a self-contained proof. We focus on showing that $c(Z) \leq O(\ln r \ln\ln r) \cdot OPT$. Consider the recursion tree associated with the execution of 2-MCAlg, where each node is labeled with arguments passed in the current invocation of 2-MCAlg. Define the depth of a subtree of this recursion tree to be the maximum number of edges on a root to leaf path of the subtree. Recall that $\beta = \frac{OPT}{r}$.

Lemma 3.5 *Let d be the depth of a subtree of the recursion tree rooted at $\big(U \subseteq V, \mathcal{T} \subseteq \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$. Let $Z_U =$ 2-MCAlg$(U, \mathcal{T})$. We have $c(Z_U) \leq 4\big(\beta|\mathcal{T}| + \mathcal{V}^x(U)\big) \ln\big(\frac{e^d r \mathcal{V}^x(U)}{OPT}\big) \ln\ln\big(e(r+1)\big)$.*

Proof : The proof is by induction on d. If $d = 0$, there is no recursive call; so $Z^U = \emptyset$, which satisfies the stated bound. Otherwise, suppose that we make the recursive calls 2-$\mathsf{MCAlg}(A_1, \mathcal{T}_1), \ldots,$ 2-$\mathsf{MCAlg}(A_h, \mathcal{T}_h)$ in step A4 to obtain edge-sets $Z_1, \ldots, Z_h$ respectively. For $p = 1, \ldots, h$, let S_p be the current set S when set A_p was added to $\mathcal{S}$ in step A3.4 (so $A_p \subseteq S_p$), and let E_p be the edge-set added to Z in this step. Then, $Z_U = \bigcup_{p=1}^{h}(E_p \cup Z_p)$. By the induction hypothesis, $c(Z_p) \leq 4\Big(\beta|\mathcal{T}_p| + \mathcal{V}^x(A_p)\Big)\ln\Big(\frac{e^{d-1}r\mathcal{V}^x(A_p)}{OPT}\Big)\ln\ln\big(e(r+1)\big)$. Let $\mathrm{vol}_p = \mathcal{V}_i^{S_p,x}(s_i,\rho_1)$ if $A_p = B_i^{S_p}(s_i,\rho_1)$ and $\mathrm{vol}_p = \overline{\mathcal{V}_i^{S_p,x}}(s_i,\rho_1)$ if $A_p = \overline{B_i^{S_p}}(s_i,\rho_1)$. Note that $\mathcal{V}^x(A_p) \leq \mathrm{vol}_p \leq \mathcal{V}^x(S_p) \leq \mathcal{V}^x(U)$. By Lemma 3.3 and the above upper bounds, we have

$$
\begin{aligned}
c(E_p) &= c^2\big(\delta^{S_p}(A_p)\big) \leq 4\mathrm{vol}_p \ln\Big(\frac{e\mathcal{V}^x(U)}{\mathcal{V}^x(A_p)}\Big)\ln\ln\big(e(r+1)\big), \qquad \text{and therefore} \\
c(E_p) + c(Z_p) &\leq 4\Big(\beta|\mathcal{T}_p| + \mathrm{vol}_p\Big)\ln\Big(\frac{e^d r\mathcal{V}^x(U)}{OPT}\Big)\ln\ln\big(e(r+1)\big). \qquad (8)
\end{aligned}
$$

Note that $\sum_{p=1}^{h}\mathrm{vol}_p \leq \mathcal{V}^x(U) + \beta(h-1)$ and $\sum_{p=1}^{h}|\mathcal{T}_p| \leq N - h$ (since each time we create a child of $(U,\mathcal{T})$ we remove at least one new (s_i,t_i) pair from $\mathcal{T}$). So adding (8) over all $p = 1, \ldots, h$, we obtain that $c(Z_U) \leq 4\big(\beta|\mathcal{T}| + \mathcal{V}^x(U)\big)\ln\big(\frac{e^d r\mathcal{V}^x(U)}{OPT}\big)\ln\ln\big(e(r+1)\big)$. ■

Theorem 3.6 *Algorithm* 2-MCAlg *returns a feasible solution of cost at most* $O(\ln r \ln\ln r)\cdot OPT$.

Proof : Feasibility of Z is shown in Lemma 3.7. Each time we make a recursive call to 2-MCAlg, the number of source-sink pairs involved decreases by at least a factor of 2, so the depth d of the overall recursion tree is $O(\log_2 r)$. Since $\beta r + \mathcal{V}^x(V) = \big(2 + \frac{1}{r}\big)OPT$ and $\frac{r\mathcal{V}^x(V)}{OPT} = r+1$, by Lemma 3.5, this implies that $c(Z) \leq O(OPT)\cdot\big(\ln(r+1) + O(\log_2 r)\big)\ln\ln\big(e(r+1)\big)$. ■

Lemma 3.7 *The solution* Z *returned by* 2-MCAlg *is feasible.*

Proof : Suppose for a contradiction that there is some s_i-t_i pair such that there are (at least) 2 edge-disjoint s_i-t_i paths P_1, P_2 in $(V, E\setminus Z)$. Consider the recursion tree of 2-MCAlg, and let $(Y,\mathcal{T}_Y)$ be the node furthest from the root such that $P_1, P_2 \subseteq E(Y)$. (Such a node must exist since the root satisfies this property.)

Note that there is at least one child $(X,.)$ of $(Y,\mathcal{T}_Y)$ such that $\delta^Y_{P_1\cup P_2}(X) \neq \emptyset$. If not and both P_1 and P_2 are contained in some child of $(Y,\mathcal{T}_Y)$ then this contradicts the definition of $(Y,\mathcal{T}_Y)$. Otherwise, we have $P_1, P_2 \subseteq E(A)$, where $A = Y \setminus \bigcup_{\text{children } (X,.) \text{ of } (Y,\mathcal{T}_Y)} X$. But then we would have processed A in step A3 and created at least one child $(A',.)$ for some $A' \subseteq A$.

We claim that if $\delta^Y_{P_1\cup P_2}(X) \neq \emptyset$ for a child $(X,.)$ of $(Y,\mathcal{T}_Y)$, then $|\delta^Y_{P_1\cup P_2}(X)| \geq 2$. This is true if both s_i and t_i are in X or if neither of them are in X since then a path crossing X must cross it at least twice. Otherwise, X is an s_i-t_i cut, and since P_1 and P_2 are edge-disjoint s_i-t_i paths in $E(Y)$, we again have $|\delta^Y_{P_1\cup P_2}(X)| \geq 2$. Among all the children $(X,.)$ of $(Y,\mathcal{T}_Y)$ such that $\delta^Y_{P_1\cup P_2}(X) \neq \emptyset$, let $(W,.)$ be the child that was added to $\mathcal{S}$ earliest in step A3.4 of 2-$\mathsf{MCAlg}(Y,\mathcal{T}_Y)$; let $S' \subseteq Y$ be the current set S when W was added. Then, $P_1 \cup P_2 \subseteq E(S')$, and so $|\delta^{S'}_{E\setminus Z}(W)| \geq |\delta^{S'}_{P_1\cup P_2}(W')| \geq 2$. But this is a contradiction, since we include in Z all but at most one edge of $\delta^{S'}(W)$. ■

General k and unit costs. The algorithm, which we denote by k-MCAlg, leading to our bicriteria guarantee is quite similar to 2-MCAlg. The only changes are the following:

- In steps A1 and A3.5, we set $\mathcal{T}'$ to be the s_i-t_i pairs from $\mathcal{T}$ that are at least $\gamma(k-1)$-edge-connected in $(S, E(S))$.

- In step A3.2, we apply Corollary 3.4 with the set S to find the radius $\rho_1 \in [0, 1)$.
- In step A3.4, we add *all* edges of $\delta^S(A)$ to Z. (Unlike 2-MC, if we only include the edges contributing to $c^q\big(\delta^S(A)\big)$ for some suitable q, then we cannot necessarily argue that the final solution satisfies the stated connectivity guarantee.)
- Of course, in step A4, we now recursively call k-MCAlg (with the same arguments).

Theorem 3.8 *For any $\gamma > 1$, algorithm k-MCAlg returns a solution Z such that $c(Z) \le O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2}\ln r \ln\ln r\big) \cdot OPT$ and every s_i-t_i pair is less than $\gamma(k-1)$-edge-connected in $(V, E \setminus Z)$.*

Thus, taking $\gamma = \frac{k}{k-1}$, we obtain a feasible solution of cost at most $O\big((k-1)^2 \ln r \ln\ln r\big) \cdot OPT$.

Proof : Let Z be the output of k-MCAlg$\big(V, \{(s_1,t_1),\dots,(s_r,t_r)\}\big)$. It is clear that Z is feasible: every s_i-t_i pair that is at least $\gamma(k-1)$-edge-connected in $(U, E(U))$, where $(U, .)$ is a node of the recursion tree is either taken care of (i.e., rendered less than $\gamma(k-1)$-edge-connected) by the edges added to Z in step A3, or, by induction, is taken care of by a recursive call.

Mimicking the proof of Lemma 3.5, and using Corollary 3.4 in place of Lemma 3.3 in the proof, one can easily show that if d is the depth of the recursion tree rooted $(U, \mathcal{T})$ and $Z_U = k$-MCAlg$(U, \mathcal{T})$, then

$$c(Z_U) \le \frac{2\gamma}{(\sqrt{\gamma}-1)^2}\Big(\beta|\mathcal{T}| + \mathcal{V}^x(U)\Big)\ln\Big(\frac{e^d r \mathcal{V}^x(U)}{OPT}\Big)\ln\ln\big(e(r+1)\big).$$

Since the depth of the overall recursion tree is $O(\log_2 r)$, as argued in the proof of Theorem 3.6, we obtain that $c(Z) \le O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2}\ln r \ln\ln r\big) \cdot OPT$. ■

4 Improved hardness result for k-(s,t)-Cut

Theorems 4.1 and 4.2 together prove that k-(s,t)-Cut is at least as hard as the densest-k-subgraph problem (DkS) problem. In DkS on hypergraphs, we seek a set of k nodes containing the maximum the number of hyperedges. Our hardness result implies that obtaining a unicriterion $O\big(k^{\epsilon_0}\,\text{polylog}(n)\big)$-approximation for some constant ϵ_0 would improve the current-best guarantee for DkS on graphs, and imply the existence of a certain family of one-way functions. Our reduction is from *small set vertex expansion* (SSVE), wherein we have a bipartite graph $G = (U \cup V, E)$ and a parameter $0 < \alpha \le 1$, and we seek a subset $S \subseteq U$ with $|S| \ge \alpha|U|$ that minimizes the number of neighbors, $\Gamma(S)$. Chuzoy et al. [11] show that SSVE reduces to the minimization version of DkS, MinDkS, wherein we seek a minimum number of nodes that contain at least k hyperedges. They also show that a ρ-approximation for MinDkS on λ-uniform hypergraphs yields a $(2\rho^\lambda)$-approximation for DkS on λ-uniform hypergraphs.

For a graph $H = (V_H, E_H)$, subset $S \subseteq V_H$, and $v \in V_H \setminus S$, we use $\delta_H(S,v) = \delta_H(v,S)$ to denote the edges between v and S in H, and $\Gamma_H(S)$ to denote the set of neighbors of S in H. As is standard, we abbreviate $\delta_H(\{v\}, V_H \setminus \{v\})$ to $\delta_H(v)$.

Theorem 4.1 ([11]) *For any $\lambda \ge 2$, there is a polytime approximation-preserving reduction that given a MinDkS-instance on a λ-uniform hypergraph with n nodes and m edges, creates an SSVE-instance with $m+n$ nodes and λm edges. Hence, a $\rho(m,n)$-approximation for SSVE yields, for λ-uniform hypergraphs, a $\rho(\lambda m, m+n)$-approximation for MinDkS and a $\Big(2\big(\rho(\lambda m, m+n)\big)^\lambda\Big)$-approximation for D$k$S.*

Theorem 4.2 *There is a polytime approximation-preserving reduction that given an SSVE-instance with n nodes and m edges, creates a k-(s,t)-Cut-instance with $O(n^3)$ nodes, $O(mn^2+n^5)$ edges, and $k = O(mn^2)$. Hence, a $\rho(k,m,n)$-approximation for k-(s,t)-Cut yields a $\rho\big(O(mn^2), O(mn^2+n^5), O(n^3)\big)$-approximation for SSVE.*

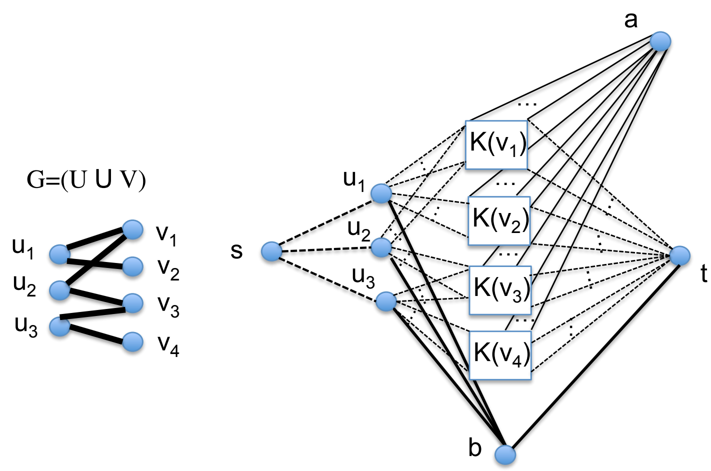

Figure 4.1: To the left, a graph of a given SSVE instance. To the right, the graph for the k-(s,t)-Cut instance created by our reduction. The edges incident into t have unit cost, while all other edges have infinite cost. Each $K(v_i)$ is a clique with $N = 25$ vertices. Dashed edges have unit capacity. The other edges have the following capacities: edge $\{s,a\}$ and edge $\{b,t\}$ have capacity 150. Each edge $\{b,u_i\}$ $(i = 1,2,3)$ has capacity 50. Each edge $\{a,x\}$ for $x \in K(v_1)$ and $x \in K(v_3)$ has capacity 2. Each edge $\{a,x\}$ for $x \in K(v_2)$ and $x \in K(v_4)$ has capacity 1.

Proof : Given an instance $\big(G = (U \cup V, E), \alpha\big)$ of SSVE, we construct the following instance of k-(s,t)-Cut; see Fig. 4.1. Let $N = 2|U||V| + 1$. Below, an infinite-cost edge (u,v) of capacity b_{uv} is simply a shorthand for b_{uv} parallel infinite-cost edges. Also, unless otherwise specified, an edge has unit capacity.

(i) We replace each vertex $v \in V$ with a clique $K(v)$ of size N. All edges in the clique have infinite cost. For each edge $(u,v) \in E$, we add an edge between u and every vertex in $K(v)$.

(ii) We add the source s and connect it to all vertices in U; we add the sink t and connect it to all vertices in $K(v)$ for every $v \in V$.

(iii) Finally, we add a vertex b, an edge (b,t) with capacity $|E| \cdot N$, and edges (b,u) with capacity $|\delta_G(u)| \cdot N$ for all $u \in U$. We also add a vertex a, an edge (s,a) with capacity $|E| \cdot N$, and for all $v \in V$, we add edges (a,x) for all $x \in K(v)$ with capacity $|\delta_G(v)|$.

All edges have infinite cost except for the edges between $\bigcup_{v \in V} K(v)$ and t, which have unit cost. We set $k = |U|(1-\alpha) + N|E| + 1$.

We claim that there exists a solution of value at most C for the SSVE instance iff there is a solution of value at most CN for the k-(s,t)-Cut instance. Note that a solution F consisting of unit-cost edges is feasible if the maximum s-t flow in the capacitated remainder graph $G' \setminus F$ has value at most $k-1$. The intuition is that if we send $N|E|$ units of flow along the paths $s{-}a{-}x{-}u{-}b{-}t$ for all $(u,v) \in E, x \in K(v)$, then in the residual digraph, all arcs between U and $\bigcup_{v \in V} K(v)$ leave U. Given this, one can mimic the arguments in [11] to show the desired claim.

Suppose the SSVE instance has a solution $S \subseteq U$ of value at most C. Construct a k-(s,t)-Cut-solution by removing the (v,t) edges for all $v \in \Gamma_G(S)$. Clearly the cost of this set is at most CN. We now argue feasibility. Consider the s-t cut induced by $A = \{s, a\} \cup S \cup \bigcup_{v \in \Gamma_G(S)} K(v)$. Then $|\delta_{G'}(A)|$ is equal to

$$\underbrace{|U \setminus S|}_{\substack{\text{edges between}\\ s \text{ and } U \setminus S}} + \underbrace{\sum_{u \in S} N|\delta_G(u)|}_{\substack{\text{edges between}\\ b \text{ and } S}} + \underbrace{\sum_{v \in V \setminus \Gamma_G(S)} N|\delta_G(v)|}_{\substack{\text{edges between } a \text{ and}\\ \bigcup_{v \in V \setminus \Gamma_G(s)} K(v)}} + \underbrace{N\big(\#\text{ of edges in } G \text{ between } U \setminus S \text{ and } \Gamma_G(S)\big)}_{\text{edges between } U \setminus S \text{ and } \bigcup_{v \in \Gamma_G(S)} K(v)}.$$

The sum of the last two terms is $\sum_{u \in U \setminus S} N|\delta_G(u)|$ and $|U \setminus S| \leq |U|(1-\alpha)$, so the size of the s-t cut is at most $|U|(1-\alpha) + \sum_{u \in U} N|\delta_G(u)| \leq |U|(1-\alpha) + N|E| \leq k-1$.

For the other direction, suppose G' has a solution F of value at most CN. Clearly, F can consist of only unit-cost edges (incident to t). We first argue that we may convert F into a structured feasible solution F' of cost at most CN where $|F' \cap \delta_{G'}(K(v))| \in \{0, N\}$ for all $v \in V$.

Fix $v \in V$. If $|\delta_{G'}(K(v), t) \setminus F| \leq k' := |U|(1-\alpha)$, then we add all edges of $\delta_{G'}(K(v), t) \setminus F$ to F. Now suppose $|\delta_{G'}(K(v), t) \setminus F| > k'$ and let $w_1, \ldots, w_{k'+1}$ be vertices in $K(v)$ such that $(w_i, t) \notin F$ for all $i = 1, \ldots, k'+1$. We claim that $F \setminus \delta_{G'}(K(v), t)$ is also feasible. Suppose to the contrary that we now have $k'+1$ edge-disjoint s-t-paths. We may assume that each such path contains at most one vertex from $K(v)$ since we can always shortcut the path to t. Consider a path P that contains a vertex $w \in K(v)$ where $w \notin \{w_1, \ldots, w_{k'+1}\}$. Then we can construct another path P' by switching w with a distinct vertex w_j for some $j \in \{1 \ldots, k'+1\}$. Note that P' is an s-t path that avoids edges in F. If we repeat this argument for all such paths, we obtain $k'+1$ edge-disjoint s-t-paths in $G' \setminus F$, contradicting the feasibility of F.

If we perform the above transformation for all $v \in V$, then we obtain a feasible solution F' of cost at most $|F| + |V|k' < (C+1)N$. But by construction $|F'|$ must be an integer multiple of N, so $|F'| \leq CN$.

Consider the residual network $\tilde{G}$ that is obtained from $G' \setminus F'$ as follows. We first bidirect the edges of $G' \setminus F'$, giving each resulting arc the same capacity as that of the corresponding edge of G'. $\tilde{G}$ is the residual network obtained after we send one unit of flow along the path s-a-x-u-b-t for every edge $(u, v) \in E$ and every $x \in K(v)$. Note that these paths are edge disjoint, so we send $N|E|$ units of flow. By flow theory [2], we know that the value of the maximum s-t-flow in $G' \setminus F'$ is at most $k - 1 = k' + N|E|$ iff the value of maximum s-t-flow in $\tilde{G}$, which equals the capacity of the minimum s-t cut in $\tilde{G}$, is at most k'. It follows there is an s-t cut in $\tilde{G}$ of capacity at most k'.

Let A be the vertices that are on the s-side of this cut. Let $S := U \cap A$. Then $|S| \geq \alpha|U|$, otherwise the cut would have capacity more than $|U|(1-\alpha)$ due to the arcs (s, u) for $u \in U \setminus S$. Consider $v \in \Gamma_G(S)$. We must have $K(v) \subseteq A$: if $K(v) \cap A = \emptyset$, then considering $u \in S$ such that $(u, v) \in E$, the cut has capacity at least $N > k'$ due to the edges between u and $K(v)$; otherwise, since $K(v)$ is split between the s- and t-sides, the cut has capacity at least $N - 1 > k'$. Finally, $\delta_{G'}(K(v), t) \subseteq F'$, otherwise $\delta_{G'}(K(v), t) \cap F' = \emptyset$, and again the cut has capacity at least $N > k'$. Thus, $|\Gamma_G(S)| \leq C$, so S is an SSVE-solution. ■

5 Extensions to node-connectivity versions of k-MC

We now consider variants of k-MC where we seek to delete edges or nodes so as to reduce the *node connectivity* of each s_i-t_i pair to at most $k-1$. Formally, as before, we are given an undirected graph $G = (V, E)$, r source-sink pairs $(s_1, t_1), \ldots, (s_r, t_r)$, and an integer $k \geq 1$. In the *edge-deletion k-route node-multicut* (ED-k-NMC) problem, we have nonnegative edge-costs $\{c_e\}_{e \in E}$ and we seek a minimum-cost set $F \subseteq E$ of edges such that the remainder graph $\overline{G} = (V, E \setminus F)$ contains at most $k-1$ node-disjoint s_i-t_i paths for all $i = 1, \ldots, r$. In the *node-deletion k-route node-multicut* (ND-k-NMC) problem, we have nonnegative node costs $\{c_v\}_{v \in V}$ and we seek a minimum-cost set $A \subseteq V \setminus \{s_1, t_1, \ldots, s_r, t_r\}$ of nodes such that the remainder graph $\overline{G} = \big(V \setminus A, E(V \setminus A)\big)$ contains at most $k-1$ node-disjoint s_i-t_i paths for all $i = 1, \ldots, r$.

The LP-relaxations of these k-route node-multicut problems induce both edge and node lengths, so to round these we develop region-growing lemmas that also incorporate node lengths. To keep notation simple, instead of proving a cumbersome overly-general region-growing lemma and obtaining the lemmas required for ED-k-NMC and ND-k-NMC as corollaries, we specifically focus on ED-k-NMC (Section 5.1) and ND-k-NMC (Section 5.2) and prove suitable region-growing lemas. We use these to obtain an $O(\ln r \ln\ln r)$-approximation for ED-k-NMC with $k = 2$, and a bicriteria $\big(\gamma, O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2}\ln r \ln\ln r\big)\big)$-approximation for ND-$k$-NMC with general k and unit node costs.

5.1 Edge-deletion k-route node-multicut (ED-k-NMC) with $k = 2$

The LP-relaxation for ED-k-NMC is as follows.

$$\begin{aligned}
\min \quad & \sum_e c_e x_e && && \text{(P2)}\\
\text{s.t.} \quad & \sum_{e\in E(P)} x_e + \sum_{v\in V(P)} y^i_v \geq 1 && \forall i, \forall P \in \mathcal{P}_i\\
& \sum_v y^i_v \leq k-1, \quad y^i_{s_i} = y^i_{t_i} = 0 && \forall i\\
& x, y \geq 0.
\end{aligned}$$

Region-growing lemma. Let $\big(x, \{y^i\}\big)$ be an optimal solution to (P2), and OPT be its value. Let $S \subseteq V$ represent the node-set of the current region. For $T \subseteq S \subseteq V$, recall that $E(S)$ is the set of edges with both endpoints in S and $\delta^S(T)$ is the set of boundary edges of T in $E(S)$. Set $\beta = OPT/r$. As before, define $\mathcal{V}^x(S) := \beta + \sum_{e\in E(S)} c_e x_e$. Let $\rho \geq 0$. Let $z \in V$. Fix a commodity i.

- Define $\ell^i(u; v) = \min_{P: P \text{ is a } u\text{-}v \text{ path}}\big(\sum_{e\in E(P)} x_e + \sum_{w\in V(P): w\neq u} y^i_w\big)$, where $E(P)$ and $V(P)$ denote respectively the set of edges and nodes of P. Note that ℓ^i defines an *asymmetric metric* on $V \times V$.
- Define $B_i(z,\rho) := \{v \in V : \ell^i(z; v) \leq \rho\}$ to be the ball of radius ρ around z. Let $B^S_i(z,\rho) := B_i(z,\rho) \cap S$.
- Define the edge-boundary $\partial^{S,x}_i$, and node-boundary $\Gamma^{S,y}_i$, of $B^S_i(z,\rho)$ in S as follows.

$$\begin{aligned}
\partial^{S,x}_i(z,\rho) &:= \{(u,v) \in E : u, v \in S,\ \ell^i(z;u) \leq \rho,\ \ell^i(z;v) - y^i_v > \rho\}\\
\Gamma^{S,y}_i(z,\rho) &:= \{v \in S : \rho < \ell^i(z;v) \leq \rho + y^i_v\}.
\end{aligned}$$

  Let $\overline{B^S_i}(z,\rho) := S \setminus \big(B^S_i(z,\rho) \cup \Gamma^{S,y}_i(z,\rho)\big)$.
- Define the following volumes:

$$\begin{aligned}
\mathcal{V}^{S,x}_i(z,\rho) &:= \beta + \sum_{\substack{(u,v)\in E: u\in B^S_i(z,\rho)\\ v\in B^S_i(z,\rho)\cup\Gamma^{S,y}_i(z,\rho)}} c_{uv}x_{uv} + \sum_{\substack{(u,v)\in\partial^{S,x}_i(z,\rho):\\ u\in B^S_i(z,\rho)}} c_{uv}\big(\rho - \ell^i(z;u)\big)\\
\overline{\mathcal{V}^{S,x}_i}(z,\rho) &:= \beta + \sum_{\substack{(u,v)\in E: u\in \overline{B^S_i}(z,\rho)\\ v\in \overline{B^S_i}(z,\rho)\cup\Gamma^{S,y}_i(z,\rho)}} c_{uv}x_{uv} + \sum_{\substack{(u,v)\in\partial^{S,x}_i(z,\rho):\\ u\in B^S_i(z,\rho)}} c_{uv}\big(\ell^i(z;v) - \rho - y^i_v\big)
\end{aligned}$$

Lemma 5.1 *Let $S \subseteq V$, $z \in V$, i be some commodity, and $0 \le a < b$. Let ρ be chosen uniformly at random from $[a, b)$. Then,*

$$\mathrm{E}_\rho\left[\frac{c\big(\partial_i^{S,x}(z,\rho)\big)}{\mathcal{V}_i^{S,x}(z,\rho)\ln\Big(\frac{e\mathcal{V}_i^{S,x}(z,b)}{\mathcal{V}_i^{S,x}(z,\rho)}\Big)}\right] \le \frac{1}{b-a}\cdot\ln\ln\Big(\frac{e\mathcal{V}_i^{S,x}(z,b)}{\mathcal{V}_i^{S,x}(z,a)}\Big) \quad \textit{and} \tag{9}$$

$$\mathrm{E}_\rho\left[\frac{c\big(\partial_i^{S,x}(z,\rho)\big)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho)\ln\Big(\frac{e\overline{\mathcal{V}_i^{S,x}}(z,a)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho)}\Big)}\right] \le \frac{1}{b-a}\cdot\ln\ln\Big(\frac{e\overline{\mathcal{V}_i^{S,x}}(z,a)}{\overline{\mathcal{V}_i^{S,x}}(z,b)}\Big) \tag{10}$$

Proof : We abbreviate $c\big(\partial_i^{S,x}(z,\rho)\big)$ to $c(\rho)$, $\mathcal{V}_i^{S,x}(z,\rho)$ to $\mathcal{V}(\rho)$ and $\overline{\mathcal{V}_i^{S,x}}(z,\rho)$ to $\overline{\mathcal{V}}(\rho)$. Let $I = \{\ell^i(z;v), \ell^i(z;v) - y_v^i : v \in V\}$. Note that $\mathcal{V}(\rho)$ and $\overline{\mathcal{V}}(\rho)$ are differentiable at all $\rho \in [a,b) \setminus I$ and for each such ρ, we have $\frac{d\mathcal{V}(\rho)}{d\rho} = c(\rho)$ and $\frac{d\overline{\mathcal{V}}(\rho)}{d\rho} = -c(\rho)$. The proof now follows from exactly the same arguments as in the proof of Lemma 3.1. ■

Corollary 5.2 *Let $S \subseteq V$, $z \in V$, and i be some commodity. Let $\alpha \in (0,1)$ and $q = \lceil\frac{k-1}{\alpha}\rceil$. We can efficiently find $\rho_1 \in [0,1)$ such that*

$$\begin{aligned}
c\big(\partial_i^{S,x}(z,\rho_1)\big) &\le \frac{2}{1-\alpha}\cdot\mathcal{V}_i^{S,x}(z,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\mathcal{V}_i^{S,x}(z,\rho_1)}\Big)\ln\ln\big(e(r+1)\big)\\
c\big(\partial_i^{S,x}(z,\rho_1)\big) &\le \frac{2}{1-\alpha}\cdot\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)}\Big)\ln\ln\big(e(r+1)\big)\\
\big|\Gamma_i^{S,y}(z,\rho_1)\big| &< q.
\end{aligned}$$

Proof : If we choose ρ uniformly at random from $[0,1)$ then $\mathrm{E}_\rho\Big[\big|\Gamma_i^{S,y}(z,\rho)\big|\Big] \le \sum_i y_v^i \le k-1$. Taking $[a,b) = [0,1)$, the arguments in Corollary 3.2 readily generalize to show that we can efficiently find $\rho_1 \in [0,1)$ such that $c\big(\partial_i^{S,x}(z,\rho_1)\big)/\mathcal{V}_i^{S,x}(z,\rho_1)\ln\big(\frac{e\mathcal{V}_i^{S,x}(z,1)}{\mathcal{V}_i^{S,x}(z,\rho_1)}\big)$, and $c\big(\partial_i^{S,x}(z,\rho_1)\big)/\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)\ln\big(\frac{e\overline{\mathcal{V}_i^{S,x}}(z,0)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)}\big)$ are at most $\frac{2}{(1-\alpha)}$ times the right-hand-sides of (9) and (10) respectively, and $\big|\Gamma_i^{S,y}(z,\rho)\big| < \frac{k-1}{\alpha}$. The lemma now follows by noting that

$$\mathcal{V}_i^{S,x}(z,1) \le \mathcal{V}^x(S), \qquad \overline{\mathcal{V}_i^{S,x}}(z,0) \le \mathcal{V}^x(S), \qquad \frac{\mathcal{V}_i^{S,x}(z,1)}{\mathcal{V}_i^{S,x}(z,0)} \le r+1, \qquad \frac{\overline{\mathcal{V}_i^{S,x}}(z,0)}{\overline{\mathcal{V}_i^{S,x}}(z,1)} \le r+1. \quad ■$$

Algorithm and analysis. The algorithm and analysis dovetail the one in Section 3.2 for $k = 2$.

Algorithm ED-2-NMCAlg$(U, \mathcal{T} = \{(s_1,t_1),\ldots,(s_N,t_N)\})$
Input: A subset $U \subseteq V$, and a collection $\mathcal{T} = \{(s_i,t_i)\}_{i=1}^N$ of source-sink pairs, where $s_i, t_i \in U$ for all $i = 1,\ldots,N$.
Output: A set $Z \subseteq E(U)$ such that s_i and t_i are at most 1-node-connected in $(U, E(U) \setminus Z)$ for all $i = 1,\ldots,N$.

B1. Set $S = U$, $Z = \emptyset$, $\mathcal{S} = \emptyset$, and $\mathcal{T}' = \big\{(s_i,t_i) \in \mathcal{T} : s_i$ and t_i are at least 2-node-connected in $(S, E(S))\big\}$.

B2. If $\mathcal{T}' = \emptyset$, return Z.

B3. While $\mathcal{T}' \neq \emptyset$, we do the following.

B3.1 Pick some $(s_i, t_i) \in \mathcal{T}'$.

B3.2 Apply Corollary 5.2 with $z = s_i$, $\alpha = 0.5$ and the set S (and $k = 2$) to find a radius $0 \le \rho_1 < 1$.

B3.3 If $B_i^S(s_i, \rho_1) \cup \Gamma_i^{S,y}(s_i, \rho_1)$ contains at most $N/2$ pairs from $\mathcal{T}$ then set $A = B_i^S(s_i, \rho_1)$, else set $A = \overline{B_i^S}(s_i, \rho_1)$.

B3.4 Set $\mathcal{S} \leftarrow \mathcal{S} \cup \{A \cup \Gamma_i^{S,y}(s_i, \rho_1)\}$. Add the edges in $\partial_i^{S,x}(s_i, \rho_1)$ to Z.

B3.5 Set $S \leftarrow S \setminus A$. Update $\mathcal{T}'$ to be the s_i-t_i pairs from $\mathcal{T}$ that are at least 2-node-connected in $(S, E(S))$.

B4. For every set $A \in \mathcal{S}$, set $Z \leftarrow Z \cup \textsf{ED-2-NMCAlg}\big(A, \{(s_i, t_i) \in \mathcal{T} : s_i, t_i \in A\}\big)$.

B5. Return Z.

The initial call to ED-2-NMCAlg, which computes the solution we return, is $\textsf{ED-2-NMCAlg}\big(V, \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$.

Let $Z := \textsf{ED-2-NMCAlg}\big(V, \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$. Define the depth of a subtree of the recursion tree corresponding to the execution of ED-2-NMCAlg to be the maximum number of edges on a root to leaf path of the subtree. The following claim will be useful to prove feasibility of Z.

Claim 5.3 *Let $S, T \subseteq V$ with $|S \cap T| \le 1$. Let $E_S \subseteq E(S)$ and $E_T \subseteq E(T)$. Let $u, v \in S$ be such that u and v are at most 1-node-connected in (S, E_S). Then, u and v are at most 1-node-connected in $(S \cup T, E_S \cup E_T)$.*

Proof : If $S \cap T = \emptyset$, this clearly holds. So assume otherwise. Suppose P_1, P_2 are two simple node-disjoint u-v paths in $G' = (S \cup T, E_S \cup E_T)$. At least one of P_1 and P_2 does not lie completely in (S, E_S); suppose P_2 is this path. But since all edges of $\delta_{G'}(S)$ are incident to a single node, and P_2 both exits and leaves S, P_2 contains a repeated node, which is a contradiction. ∎

Lemma 5.4 *The solution Z returned is feasible.*

Proof : Suppose for a contradiction that there is some s_i-t_i pair that is (at least) 2-node-connected in $(V, E \setminus Z)$. Consider the recursion tree of ED-2-NMCAlg, and let $(Y, \mathcal{T}_Y)$ be the node furthest from the root such that s_i and t_i are at least 2-node-connected in the subgraph $(Y, E(Y))$ induced by Y. Suppose that the loop in step B3 of $\textsf{ED-2-NMCAlg}(Y, \mathcal{T}_Y)$ runs for h iterations. Note that $h \ge 1$ since s_i and t_i are at least 2-node-connected in $(Y, E(Y))$. Let X_p be the set added to $\mathcal{S}$ in step B3.4 in the p-th iteration of the loop. Let $X_{h+1} \subseteq Y$ be the set S at the termination of the loop. Let $S_p = \bigcup_{q=p}^{h+1} X_q$ (so $S_1 = Y$). Notice that $|X_p \cap S_{p+1}| \le 1$, since $X_p \cap S_{p+1} \subseteq \Gamma_p^{S_p,y}(s_p, \rho_p)$, where s_p-t_p is the source-sink pair and ρ_p is the radius chosen in the p-th iteration, and $|\Gamma_p^{S_p,y}(s_p, \rho_p)| < 2$ by Lemma 5.2.

Let p be the highest index such that s_i and t_i are at least 2-node-connected in $(S_p, E(S_p))$. Note that $p \le h$, otherwise the loop in step B3 would not have terminated with $S = X_{h+1}$. If $s_i, t_i \in S_{p+1}$, they are at most 1-node-connected in $(S_{p+1}, E(S_{p+1}))$. Since $|X_p \cap S_{p+1}| \le 1$, we have $E(S_p) = E(X_p) \cup E(S_{p+1})$, and by Claim 5.3 it follows that s_i and t_i are at most 1-node-connected in $(S_p, E(S_p))$, which is a contradiction. If $s_i, t_i \in X_p$, they are at most 1-node-connected in $(X_p, E(X_p))$ due to the definition of $(Y, \mathcal{T}_Y)$, and so we arrive at the same contradiction. So it must be that $\big|\{s_i, t_i\} \cap (X_p \setminus S_{p+1})\big| = 1$. But then all s_i-t_i paths in $(S_p, E(S_p))$ contain the singleton node in $X_p \cap S_{p+1}$. So s_i and t_i are at most 1-node-connected in $(S_p, E(S_p))$, and we have the same contradiction. ∎

Lemma 5.5 *Let d be the depth of a subtree of the recursion tree rooted at $\big(U \subseteq V, \mathcal{T} \subseteq \{(s_1, t_1), \ldots, (s_r, t_r)\}\big)$. Let $Z_U = \textsf{ED-2-NMCAlg}(U, \mathcal{T})$. Then $c(Z_U) \le 4\big(\beta|\mathcal{T}| + \mathcal{V}^x(U)\big) \ln\big(\frac{e^d r \mathcal{V}^x(U)}{OPT}\big) \ln\ln\big(e(r+1)\big)$.*

Proof : When $d = 0$, we have $Z_U = \emptyset$, so the statement holds. Suppose in step B3 of $\mathsf{ED\text{-}2\text{-}NMCAlg}(U, \mathcal{T})$, we add sets $A_1, \ldots, A_h$ to $\mathcal{S}$ (where $h \geq 1$), in that order. For $p = 1, \ldots, h$, let S_p be the current set S when A_p was added to $\mathcal{S}$ in step B3.4, and let E_p be the edge-set added to Z in this step. Let $Z_1, \ldots, Z_h$ be the edge-sets returned by the recursive calls to $\mathsf{ED\text{-}2\text{-}NMCAlg}(A_1, \mathcal{T}_1), \ldots, \mathsf{ED\text{-}2\text{-}NMCAlg}(A_h, \mathcal{T}_h)$ in step B4.

Let $\text{vol}_p = \mathcal{V}_i^{S_p,x}(s_i, \rho_1)$ if $A_p = B_i^{S_p}(s_i, \rho_1) \cup \Gamma_i^{S_p,y}(s_i, \rho_1)$ and $\text{vol}_p = \overline{\mathcal{V}_i^{S_p,x}}(s_i, \rho_1)$ if $A_p = \overline{B_i^{S_p}}(s_i, \rho_1) \cup \Gamma_i^{S_p,y}(s_i, \rho_1)$. The key thing to note is that we still have $\mathcal{V}^x(A_p) \leq \text{vol}_p \leq \mathcal{V}^x(S_p) \leq \mathcal{V}^x(U)$ and $\sum_{p=1}^{h} \text{vol}_p \leq \mathcal{V}^x(U) + \beta(h-1)$. The latter follows since an easy induction argument shows that $\sum_{p=q}^{h} \text{vol}_p \leq \mathcal{V}^x(S_q) + \beta(h-q)$ for all $q = 1, \ldots, h$. Given this, the rest of the proof is identical to that of Lemma 3.5. ■

Each recursive call to ED-2-NMCAlg, reduces the number of source-sink pairs involved by a factor of at least 2, so the depth d of the entire recursion tree is $O(\log_2 r)$. So we have shown the following.

Theorem 5.6 *Algorithm* ED-2-NMCAlg *returns a feasible solution of cost at most* $O(\ln r \ln\ln r) \cdot OPT$.

5.2 Node-deletion k-route node-multicut (ND-k-NMC) with unit costs

The LP-relaxation for ND-k-NMC is as follows.

$$\begin{aligned}
\min \quad & \sum_{v} c_v x_v && \text{(P3)}\\
\text{s.t.} \quad & \sum_{v \in V(P)} (x_v + y_v^i) \geq 1 && \forall i, \forall P \in \mathcal{P}_i\\
& \sum_{v} y_v^i \leq k-1, \quad y_{s_i}^i = y_{t_i}^i = 0 && \forall i\\
& x, y \geq 0, \qquad x_v = 0 && \forall v \in \{s_1, t_1, \ldots, s_r, t_r\}.
\end{aligned}$$

Region-growing lemma. Let $\big(x, \{y^i\}\big)$ be an optimal solution to (P3), and OPT be its value. As before, let $S \subseteq V$ represent the node-set of the current region. Set $\beta = OPT/r$. Let $z \in V$ and $\rho \geq 0$. Fix a commodity i.

- Define $\ell^i(u; v) = \min_{P: P \text{ is a } u\text{-}v \text{ path}} \sum_{w \in V(P): w \neq u} (x_w + y_w^i)$, where $V(P)$ is the set of nodes of P. As before, ℓ^i defines an asymmetric metric on $V \times V$.
- Define $B_i(z, \rho) := \{v \in V : \ell^i(z; v) \leq \rho\}$, and $B_i^S(z, \rho) := B_i(z, \rho) \cap S$.
- Define the x-boundary of $B_i^S(z, \rho)$ in S to be $\Gamma_i^{S,x}(z, \rho) := \{v \in S : \rho + y_v^i < \ell^i(z; v) \leq \rho + x_v + y_v^i\}$. Define the y-boundary of $B_i^S(z, \rho)$ in S to be $\Gamma_i^{S,y}(z, \rho) := \{v \in S : \rho < \ell^i(z; v) \leq \rho + y_v^i\}$. Note that $\Gamma_i^{S,x}(z, \rho)$ and $\Gamma_i^{S,y}(z, \rho)$ partition $\Gamma_i^S(z, \rho) := \{v \in S \setminus B_i^S(z, \rho) : \exists u \in B_i^s(z, \rho) \text{ s.t. } (u, v) \in E\}$. Let $\overline{B_i^S}(z, \rho) := S \setminus \big(B_i^S(z, \rho) \cup \Gamma_i^S(z, \rho)\big)$.
- Define the following volumes:

$$\begin{aligned}
\mathcal{V}_i^{S,x}(z, \rho) \quad &:= \quad \beta + \sum_{u \in B_i^S(z,\rho) \cup \Gamma_i^{S,y}(z,\rho)} c_u x_u + \sum_{u \in \Gamma_i^{S,x}(z,\rho)} c_u\big(\rho - (\ell^i(z; u) - x_u - y_u^i)\big)\\
\overline{\mathcal{V}_i^{S,x}}(z, \rho) \quad &:= \quad \beta + \sum_{u \in \overline{B_i^S}(z,\rho) \cup \Gamma_i^{S,y}(z,\rho)} c_u x_u + \sum_{u \in \Gamma_i^{S,x}(z,\rho)} c_u(\ell^i(z; u) - y_u^i - \rho)
\end{aligned}$$

The following lemma is analogous to Lemma 5.1 and follows from the same reasoning.

Lemma 5.7 *Let $S \subseteq V$, $z \in V$, i be some commodity, and $0 \le a < b$. Let ρ be chosen uniformly at random from $[a, b)$. Then,*

$$\mathrm{E}_\rho\left[\frac{c\big(\Gamma_i^{S,x}(z,\rho)\big)}{\mathcal{V}_i^{S,x}(z,\rho)\ln\Big(\frac{e\mathcal{V}_i^{S,x}(z,b)}{\mathcal{V}_i^{S,x}(z,\rho)}\Big)}\right] \le \frac{1}{b-a}\cdot\ln\ln\Big(\frac{e\mathcal{V}_i^{S,x}(z,b)}{\mathcal{V}_i^{S,x}(z,a)}\Big) \quad \textit{and}$$

$$\mathrm{E}_\rho\left[\frac{c\big(\Gamma_i^{S,x}(z,\rho)\big)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho)\ln\Big(\frac{e\overline{\mathcal{V}_i^{S,x}}(z,a)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho)}\Big)}\right] \le \frac{1}{b-a}\cdot\ln\ln\Big(\frac{e\overline{\mathcal{V}_i^{S,x}}(z,a)}{\overline{\mathcal{V}_i^{S,x}}(z,b)}\Big)$$

Corollary 5.8 *Let $S \subseteq V$. Suppose that $s_i, t_i \in S$ and there are $\gamma(k-1)$ node-disjoint s_i-t_i paths internal to S, where $\gamma > 1$. Suppose that $c_v = 1$ for all nodes v. We can efficiently find $\rho_1 \in [0, 1)$ such that*

$$c\big(\Gamma_i^S(s_i,\rho_1)\big) \le \frac{2\gamma}{(\sqrt{\gamma}-1)^2}\cdot\mathcal{V}_i^{S,x}(s_i,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\mathcal{V}_i^{S,x}(s_i,\rho_1)}\Big)\ln\ln\big(e(r+1)\big)$$
$$c\big(\Gamma_i^S(s_i,\rho_1)\big) \le \frac{2\gamma}{(\sqrt{\gamma}-1)^2}\cdot\overline{\mathcal{V}_i^{S,x}}(s_i,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\overline{\mathcal{V}_i^{S,x}}(s_i,\rho_1)}\Big)\ln\ln\big(e(r+1)\big)$$

Proof : Let $\alpha \in (0, 1)$, whose value we will fix later. Take $[a, b) = [0, 1)$. We have

$$\mathcal{V}_i^{S,x}(z,1) \le \mathcal{V}^x(S), \qquad \overline{\mathcal{V}_i^{S,x}}(z,0) \le \mathcal{V}^x(S), \qquad \frac{\mathcal{V}_i^{S,x}(z,1)}{\mathcal{V}_i^{S,x}(z,0)} \le r+1, \qquad \frac{\overline{\mathcal{V}_i^{S,x}}(z,0)}{\overline{\mathcal{V}_i^{S,x}}(z,1)} \le r+1.$$

Given this, the arguments in Corollary 3.2 readily generalize to show that we can efficiently find $\rho_1 \in [0, 1)$ such that

$$c\big(\Gamma_i^{S,x}(s_i,\rho_1)\big) \le \frac{2}{1-\alpha}\cdot\mathcal{V}_i^{S,x}(z,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\mathcal{V}_i^{S,x}(z,\rho_1)}\Big)\ln\ln\big(e(r+1)\big) \tag{11}$$
$$c\big(\Gamma_i^{S,x}(s_i,\rho_1)\big) \le \frac{2}{1-\alpha}\cdot\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)\ln\Big(\frac{e\mathcal{V}^x(S)}{\overline{\mathcal{V}_i^{S,x}}(z,\rho_1)}\Big)\ln\ln\big(e(r+1)\big) \tag{12}$$
$$\big|\Gamma_i^{S,y}(z,\rho)\big| < \frac{k-1}{\alpha}. \tag{13}$$

Note that $t_i \notin \Gamma_i^S(s_i, \rho_1)$ since $\rho_1 < 1$. So removing $\Gamma_i^S(s_i, \rho_1)$ disconnects s_i and t_i, and hence, $|\Gamma_i^S(s_i, \rho_1)| \ge \gamma(k-1)$. Therefore, since we have unit node costs and $\Gamma_i^{S,x}(s_i, \rho_1)$ and $\Gamma_i^{S,y}(s_i, \rho_1)$ partition $\Gamma_i^S(s_i, \rho_1)$, we have $c\big(\Gamma_i^S(s_i, \rho_1)\big) \le c\big(\Gamma_i^{S,x}(s_i, \rho_1)\big)\cdot\frac{\gamma}{\gamma-1/\alpha}$. Plugging in the bounds from (11), (12), we see that the constant factor multiplying the volume terms on the RHS is minimized by setting $\alpha = \gamma^{-1/2}$, which yields the constant factor $\frac{2\gamma}{(\sqrt{\gamma}-1)^2}$. ■

Algorithm and analysis. The algorithm, ND-k-NMCAlg, for ND-k-NMC is quite similar to ED-2-NMCAlg. The only changes are the following:

- In steps B1 and B3.5, we set $\mathcal{T}'$ to be the s_i-t_i pairs from $\mathcal{T}$ that are at least $\gamma(k-1)$-node-connected in $(S, E(S))$.

- In step B3.2, we apply Corollary 5.8 with the set S to find the radius $\rho_1 \in [0,1)$.
- In step B3.4, we add A to $\mathcal{S}$, and add all nodes of $\Gamma_i^S(s_i, \rho_1)$ to Z.
- Of course, in step B4, we now recursively call ND-k-NMCAlg (with the same arguments).

Theorem 5.9 *For any $\gamma > 1$, algorithm* ND-k-NMCAlg *returns a solution Z such that $c(Z) \le O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2}\ln r \ln\ln r\big)\cdot$ OPT and every s_i-t_i pair is less than $\gamma(k-1)$-node-connected in $(V\setminus Z, E(V\setminus Z))$.*

Thus, taking $\gamma = \frac{k}{k-1}$, we obtain a feasible solution of cost at most $O\big((k-1)^2 \ln r \ln\ln r\big)\cdot OPT$.

Proof : Let Z be the output of ND-k-NMCAlg$\big(V, \{(s_1,t_1),\ldots,(s_r,t_r)\}\big)$. It is clear that Z is feasible. Mimicking the proof of Lemma 3.5, and using Corollary 5.8 in place of Lemma 3.3 in the proof, one can easily show that if d is the depth of the recursion tree rooted $(U,\mathcal{T})$ and $Z_U =$ ND-k-NMCAlg$(U,\mathcal{T})$, then

$$c(Z_U) \le \frac{2\gamma}{(\sqrt{\gamma}-1)^2}\Big(\beta|\mathcal{T}| + \mathcal{V}^x(U)\Big)\ln\Big(\frac{e^d r\mathcal{V}^x(U)}{OPT}\Big)\ln\ln\big(e(r+1)\big).$$

Since the depth of the recursion tree is $O(\log_2 r)$, we obtain that $c(Z) \le O\big(\frac{\gamma}{(\sqrt{\gamma}-1)^2}\ln r\ln\ln r\big)\cdot OPT$. ■

References

[1] S. Arora and C. Lund. Hardness of approximations. In D. Hochbaum, editor, *Approximation Algorithms for NP-hard Problems*, PWS Publishing co., Boston, MA, USA, 1997.

[2] R. Ahuja, T. Magnanti, J. Orlin, Network Flows: Theory, Algorithms, and Applications. *Prentice-Hall, Inc.* 1993.

[3] B. Applebaum, Pseudorandom generators with long stretch and low locality from random local one-way functions. In *Electronic Colloquium on Computational Complexity* (ECCC), 18:7, 2011.

[4] S. Arora, L. Babai, J. Stern, and Z. Swedyk, The hardness of approximate optima in lattices, codes, and systems of linear equations. *Journal of Computer System Science*, 54:317331, 1997

[5] S. Barman and S. Chawla, Region growing for multi-route cuts. In *Proceedings of the 21st SODA*, pages 404–418, 2010.

[6] S. Bhatt, and F. T. Leighton, A framework for solving VLSI graph layout problems. *Journal of Computer and System Sciences* 28.2 300-343, 1984.

[7] H. Bruhn, J. Cerny, A. Hall, P. Kolman, and J. Sgall. Single source multiroute flows and cuts on uniform capacity networks. *Theory of Computing*, 4(1):1–20, 2008.

[8] D. Chakrabarty, R. Krishnaswamy, S. Li, and S. Narayanan, Capacitated Network Design on Undirected Graphs. In *Proceedings of the 16th APPROX*, pages 71–80, 2013.

[9] S. Chawla, A. Gupta, and H. Räcke. Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. *ACM Transactions on Algorithms*, 4(2), 2008.

[10] C. Chekuri and S. Khanna, Algorithms for 2-Route Cut Problems. *Automata, Languages and Programming*, 472-484, 2008.

[11] J. Chuzhoy, Y. Makarychev, A. Viajayaraghavan, and Y. Zhou, Approximation algorithms and hardness of the k-route cut problem. In *Proceedings of the 23rd SODA*, pages 780–799, 2012.

[12] D. Dahlhaus, D. S. Johnson, C. H. Papadimitriou, P. D. Seymour, and M. Yannakakis, The complexity of multiway cuts. *Proceedings of the 24th STOC*, 241-251, 1992

[13] M. Dinitz and A. Gupta. Packing interdiction and partial covering problems. In *Proceedings of the 16th IPCO*, pages 157–168, 2013.

[14] M. Dinitz, G. Kortsarz, and R. Raz. Label cover instances with large girth and the hardness of approximating basic k-spanner, *Proceedings of the 39th international colloquium conference on Automata, Languages, and Programming* 290-301, 2012.

[15] G. Even, J. Naor, S. Rao, and B. Schieber, Divide-and-conquer approximation algorithms via spreading metrics. *Journal of the ACM*, 47(4):585–616, 2000.

[16] G. Even, J. Naor, B. Schieber, and M. Sudan, Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20:151–174, 1998.

[17] S. Fiorini, G. Joret, U. Pietropaoli, Hitting Diamonds and Growing Cacti. *14th Conference on Integer Programming and Combinatorial Optimization*, pages 191-204, 2010.

[18] F. Fomin, D. Lokshtanov, N. Misra, and S. Saurabh, Planar F-Deletion: Approximation, Kernelization and Optimal FPT Algorithms. In *Proceedings of the 53rd FOCS*, pages 470–479, 2012.

[19] N. Garg, V. Vazirani, and M. Yannakakis, Approximate max-flow min-(multi)-cut theorems and their applications. *SIAM Journal on Computing*, 25:235–251, 1996.

[20] A. Hayrapetyan, D. Kempe, M. Pl, and Z. Svitkina, Unbalanced graph cuts. *In Proceedings of the 13th ESA*, pages 191–202, 2005.

[21] W. Kishimoto, A method for obtaining the maximum multi-route flow in a network. *Networks*, 27(4):279-291, 1996.

[22] P. Kolman, and C. Scheideler, Approximate Duality of Multicommodity Multiroute Flows and Cuts: Single Source Case. In *Proceedings of the 23rd SODA*, pages 800–810, 2012.

[23] P. Kolman, and C. Scheideler, Towards Duality of Multicommodity Multiroute Cuts and Flows: Multilevel Ball-Growing. In *Proceedings of the 28th STACS*, pages 129–140, 2011.

[24] G. Kortsarz, On the Hardness of Approximating Spanners. *Algorithmica*, 30, 1999.

[25] B. Laekhanukit, *Personal communication*, October 2013.

[26] B. Laekhanukit, Parameters of two-prover-one-round game and the hardness of connectivity problems. In *Proceedings of the 25th SODA*, pages 1626–1643, 2014.

[27] F. Leighton and S. Rao., Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 46:787-832, 1999.

[28] A. Li, and P. Zhang, Unbalanced graph partitioning. *Theory of Computing Systems*, 53:454-466, 2013.

[29] J. Shi, and J. Malik, Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 888–905, 2000.

[30] C. A. Phillips. The network inhibition problem. In *Proceedings of the 25th STOC*, pages 776–785, 1993.

[31] C. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[32] H. Räcke. Minimizing congestion in general networks. In *Proceedings of the 43rd FOCS*, pages 43–52, 2002.

[33] P. Seymour. Packing directed circuits fractionally. *Combinatorica*, 15:281–288, 1995.

[34] D. Shmoys Cut problems and their application to divide-and-conquer. In D. Hochbaum, editor, *Approximation Algorithms for NP-hard Problems*, PWS Publishing co., Boston, MA, USA, 1997.

[35] C. Tovey, A simplified NP-complete satisfiability problem. *Discrete Applied Mathematics* 8.1: 85-89, 1984.

[36] V. Vazirani, *Approximation Algorithms*. Springer Verlag, 2001.

[37] D. Williamson and D. Shmoys. *The Design of Approximation Algorithms*, Cambridge University Press, 2010.

[38] R. K. Wood. Deterministic network interdiction. *Mathematical and Computer Modeling*, 17(2):1–18, 1993.

[39] R. Zenklusen. Network flow interdiction on planar graphs. *Discrete Applied Mathematics*, 158:1441–1455, 2010.

A The k-route all-pairs cut problem

Theorem A.1 *The* 3*-route all-pairs cut problem is APX-hard.*

Proof : We give an L-reduction from vertex cover on bounded-degree graphs, which is known to be *APX*-hard [31]. Given a vertex-cover instance $\hat{G} = (\hat{V}, \hat{E})$, where $\hat{G}$ has maximum degree $\alpha = O(1)$, we construct an instance $G = (V, E)$ of the 3-route all-pairs cut problem. In the following, to avoid confusion, we will refer to the elements $(\hat{V}, \hat{E})$ of the vertex-cover instance $\hat{G}$ as *vertices* and *edges*, and to the elements (V, E) of the constructed 3-route all-pairs-cut instance as *nodes* and *links*.

Let the vertices in $\hat{V}$ be numbered $1, 2, \ldots, |\hat{V}|$. For every vertex $v \in \hat{V}$, we introduce a path P_v in G that contains as many links as the degree of v. That is, P_v has one link f_{e^v} for every edge $e \in \hat{E}$ incident to v in $\hat{G}$. We give infinite cost to such links. Let a_v be the first node of the path P_v and b_v be the last. We add a link (a_v, b_v) of unit cost in G. Note that P_v and (a_v, b_v) yields a cycle in G for every $v \in \hat{V}$. We also connect a_v to a_{v+1} through a cycle formed by 3 links with infinite cost. That is, we introduce $|\hat{V}| - 1$ triangles connecting all paths. For each edge $e = (u, v) \in \hat{E}$ we introduce a node σ_e and we connect σ_e to the endpoints of f_{e^u} and to the endpoints of f_{e^v}, with links of unit cost. We let $G = (V, E)$ be the resulting graph for our 3-route all-pairs cut instance (see Fig.A.2).

Let p^* and c^* be the cost of an optimal solution for the vertex-cover instance and the cost of an optimal solution for the 3-route all-pairs cut instance, respectively. We claim that:

(i) If there exists a vertex cover in $\hat{G}$ of size p, then there is a solution for the 3-route all-pairs cut instance of cost at most $2|\hat{E}| + p$. Note that this implies that $c^* \leq 2|\hat{E}| + p^* \leq (2\alpha + 1)p^*$.

(ii) For any feasible solution for the 3-route all-pairs cut instance of cost at most $2|\hat{E}| + p$ we can construct a cover of $\hat{G}$ of size at most p.

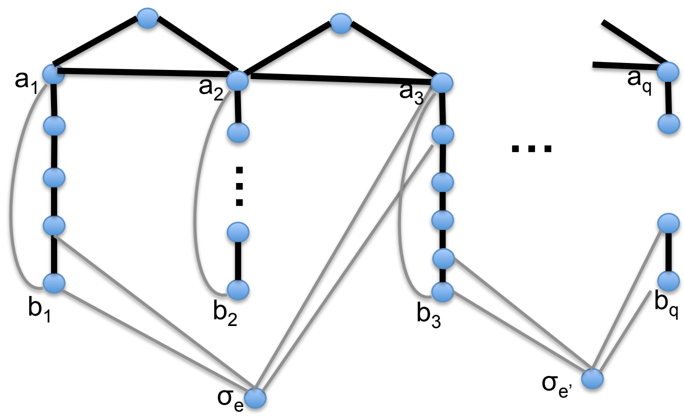

Figure A.2: The instance created by our reduction from a vertex-cover instance $\hat{G}$ on q vertices. Black edges have infinite cost and grey edges have unit cost. Node σ_e represents the edge $(1, 3)$ and node $\sigma_{e'}$ represents the edge $(3, q)$. Vertex 1 has degree 4 in $\hat{G}$ and vertex 3 has degree 5 in $\hat{G}$.

This implies that we have an L-reduction, and shows that if we have a β-approximation for 3-route all-pairs cut, then we can obtain a vertex-cover solution of size at most $\beta c^* - 2|\hat{E}| = O(\beta)p^*$, yielding the theorem.

In proving this, a useful observation is that a set F of edges is feasible for 3-route all-pairs cut problem iff the remainder graph $\overline{G} = (V, E \setminus F)$ has the property that every two simple cycles meet at most at one vertex. Such a graph is called a *cactus* graph.

For (i), suppose there exists a vertex cover of size p. For every v in the cover, we add (a_v, b_v) in F. Furthermore, for each edge $e = (u, w) \in \hat{E}$, we select one vertex between u and w that is in the cover (at least one of them is in the cover by definition), say u, and we add to F the links connecting σ_e to the endpoints of f_{e^w} for the other vertex w. It is not difficult to check that F yields a feasible solution for the 3-route all-pairs cut instance (using the relationship to cactus graphs) of the claimed cost.

For (ii), suppose we have a feasible solution F for the 3-route all-pairs cut instance, and consider the remainder graph obtained by removing F. Clearly, all links of infinite cost are still present. Note that each node σ_e can have at most two links incident to it in the remainder graph, and both these links must be incident to two nodes of the same path P_v for some v. If not, then we would have an infinite-cost link of some triangle that connects the vertices $\{a_v\}_{v \in \hat{V}}$ that is contained into another cycle other than the triangle, contradicting feasibility of our solution. We first argue that we may assume that each σ_e has exactly two links incident to it in the remainder graph (and hence, also in F). Let $e = (u, w)$ be the edge in $\hat{E}$ corresponding to σ_e. As argued above, F contains at least one pair of links that connect σ_e to the nodes of a path, say P_u. Suppose that F also contains some links connecting σ_e to P_w. If we remove such links from F, we create one additional cycle, containing the link f_{e^w}, in the remainder graph. Thus, the new remainder graph is not a cactus iff f_{e^w} is already contained in some cycle in $(V, E \setminus F)$. But there is only one possible cycle in $(V, E \setminus F)$ containing f_{e^w}, namely the cycle formed by P_w and (a_w, b_w). This observation implies that $F \cup \{(a_w, b_w)\} \setminus \{$the two links connecting σ_e to $P_w\}$ is a feasible solution to our 3-route all-pairs cut instance. Furthermore, this solution has no greater cost since we are adding at most one link of unit cost, and we removing at least one link of the same cost.

Since the cost of our solution is at most $2|\hat{E}| + p$, it follows that there are at most p links in F of the form (a_v, b_v). We claim that these vertices v form a cover in $\hat{G}$. Suppose not. Then there is at least one edge $e \in \hat{E}$ that is not covered by these vertices. We know that the node σ_e is connected in $\overline{G}$ to the endpoints of the link f_{e^u} for one of the endpoints, say u, of e. The link f_{e^u} is therefore contained in the cycle formed by P_u and (a_u, b_u), since (a_u, b_u) is not in F, and is also contained in the triangle with the node σ_e, which contradicts feasibility of F. ■

Corollary A.2 *The k-route all-pairs cut problem is APX-hard for all $k \geq 3$.*

Proof : The reduction is very similar to the one in the proof of Theorem A.1. The only change is that in the graph G created from the given vertex-cover instance, we now have: (a) $k-2$ parallel links between every pair of consecutive nodes of every path P_v; and (b) $k-2$ parallel links between a_j, a_{j+1} for all $j = 1, \ldots, |\hat{V}| - 1$.

Suppose there exists a vertex cover of size p. As before, for each node σ_e we remove exactly one pair of links incident to σ_e, and in particular we choose the pair of links that connect σ_e to the endpoints of the edge f_{e^u} if $e = (u, v)$ and v is in the cover, and the other pair otherwise. We also remove all edges of the form (a_v, b_v) for v in the cover. We remove $2|\hat{E}| + p$ edges in total.

We claim that for every pair of nodes z, w of the remainder graph, we have at most $k-1$ edge-disjoint paths. Every node that is not a node of a path P_v has maximum degree two, and therefore this is clear. If z and w belongs to two different paths P_v and P_u with $u > v$, then every path connecting them must use either the link (a_v, a_{v+1}) (and there are at most $k-2$ such links) or the two upper links of the triangle formed with a_v and a_{v+1}. Therefore, we can have at most $k-1$ edge-disjoint such paths. Finally, if z and w belong to the same path P_u, we have $k-2$ paths given by the infinite cost links, plus at most one additional path that uses either the edge (a_u, b_u) or a sequence of pairs of links incident into the nodes $\{\sigma_e\}$ for the edges e that have u as an endpoint in $\hat{G}$. Note that, by construction, if (a_u, b_u) is still in the graph, then all the pairs of links incident into the nodes $\{\sigma_e\}$ for the edges e that have u as an endpoint have been removed, and therefore, once again we get at most $k-1$ edge-disjoint paths.

For the other direction, suppose we have a feasible solution F for the k-route all-pairs cut instance, and consider the remainder graph obtained by removing such set of links. Clearly, all links of infinite cost are still present. Once again, each node σ_e corresponding to an edge $e = (u, v)$ can have at most two links incident to it in the remainder graph, and both these links must be incident to two nodes of the same path. If not, then we would have a pair of nodes (a_v and a_u) that are connected by k edge-disjoint paths: $k-1$ given by the infinite-cost links not in the paths P_u and P_v, and one which uses the links in the paths P_u and P_v and two links incident to σ_e. Also, as before, we may assume that each σ_e has exactly two links incident into it in the remainder graph (and hence, in F), because otherwise for one endpoint u of e we get that $F \cup \{(a_u, b_u)\} \setminus \{$the two links connecting σ_e to $P_u\}$ is a feasible solution to our k-route all-pairs cut instance of no larger cost. So if the cost of F is at most $2|\hat{E}| + p$, it follows that we have at most p links in F of the form (a_v, b_v). We claim that these vertices v form a cover in $\hat{G}$.

Suppose not. Then there is at least one edge $e \in \hat{E}$ that is not covered by such vertices. We know that the node σ_e is connected in $\overline{G}$ to the endpoints of the link f_{e^u} for one of the endpoints, say u, of e. The endpoints of f_{e^u} are therefore connected by $k-2$ parallel paths using one single link, one path formed by the edge (a_u, b_u) and edges of $P_u \setminus \{f_{e^u}\}$, and one path contained in the triangle with the node σ_e. This contradicts feasibility of F. ■

On the positive side, the 3-route all-pairs cut problem admits an $O(1)$-approximation. This follows from: (1) the equivalence of 3-route all-pairs cut and the problem of removing a min-cost set of edges so that the remainder graph is a cactus; (2) the results of Fiorini et al. [17], who gave an $O(1)$-approximation for the problem of removing a minimum-weight node set so that the remaining graph is a cactus; and (3) the edge-removal version easily reduces to the node-removal version by subdividing edges, and setting the cost of the original vertices to ∞ and the cost of each vertex corresponding to an edge to be the cost of the edge.

Recently, Fomin et al. [18] developed an $O(1)$-approximation algorithm for the problem of removing the fewest number of nodes so that the remaining graph excludes a minor from a given list of graphs, at least one of which should be planar. While k-route all-pairs cut can be stated as excluding the planar graph with k parallel edges as a minor of the remainder graph, the result of [18] does not directly apply here. This is because our transformation of an edge-weighted instance to a node weighted one introduces non-uniform node weights, whereas the algorithm in [18] is for uniform node weights.

B Hardness of the edge-deletion k-route node-multicut problem

Recall that in the edge-deletion k-route node-multicut (ED-k-NMC) problem, we have an undirected graph $G = (V, E)$ with nonnegative edge costs $\{c_e\}_{e\in E}$, and r source-sink pairs $(s_1, t_1), \ldots, (s_r, t_r)$ and an integer $k \geq 1$. We seek a minimum-cost set $F \subseteq E$ of edges such that the remainder graph $\overline{G} = (V, E \setminus F)$ contains at most $k - 1$ node-disjoint s_i-t_i paths for all $i = 1, \ldots, r$.

Chuzhoy et al. [11] show that ED-k-NMC is hard to approximate within a factor $\Omega(k^{\epsilon})$. They present a reduction from 3-SAT(5), which is the variant of 3-SAT where each variable occurs in at most 5 clauses, coupled with the parallel-repetition theorem, which is essentially a reduction from (the minimization version) of *label cover*. However, Laekhanukit [26] pointed out some subtle (but fixable) errors in their proof and proposed a correction, but his reduction also suffers from some subtle (again fixable) errors [25]. We give a correct proof below via a somewhat simpler reduction than the ones in [11, 26].

Label cover was first introduced by Arora et al. [4] and has been subsequently used as a basis for many hardness reductions (see, e.g., [1]). Kortsarz [24] presented a minimization version of label cover (sometimes known as MinRep) with the same hardness guarantee, that has since found use in various network-design applications (see, e.g., [14]).

In the MinRep problem, we are given a bipartite graph $H = (U \cup W, F)$, two sets of labels L_1 (for vertices in U) and L_2 (for vertices in W), and a constraint function for each edge e defined as $\pi_e : L_1 \to L_2$. A *labeling* is given by specifying a set of labels $f(u) \subseteq L_1$ for every vertex $u \in U$ and a set of labels $f(w) \subseteq L_2$ for every vertex $w \in W$. We say that a labeling *covers* an edge $e = uw \in F$ if there exists $a \in f(u)$ and $b \in f(w)$ such that $\pi_e(a) = b$. *Min-Rep* asks for a labeling that covers all the edges while minimizing $\sum_{u\in U} |f(u)| + \sum_{w\in W} |f(w)|$.

Theorem B.1 (see, e.g., [37]) *There are constants $\epsilon_0, \delta_0 > 0$ such that there is no polytime algorithm for* MinRep *with approximation factor:*

- $O\big(q^{\epsilon_0}\big)$ *unless P=NP, where $q = |L_1| + |L_2|$ is the size of the label set;*
- $O\big(\Delta^{\delta_0}\big)$ *unless P=NP, where Δ is the maximum degree of the underlying graph;*
- $2^{\log^{1-\epsilon} m}$ *for any constant ϵ, unless NP is contained in deterministic quasipolynomial time, where m is the number of edges.*

Theorem B.2 *There is a polytime approximation-preserving reduction that given a* MinRep*-instance (H, π, L_1, L_2) with label-size $q = |L_1| + |L_2|$ and maximum-degree Δ, constructs an* ED-k-NMC*-instance $\big(G, \{c_e\}, \{s_1, t_1, \ldots, s_r, t_r\}, k\big)$ with $k = O(\Delta q)$, $r = |E_H|$, and $|E_G| = O(|E_H|q)$.*
Hence, there is no $O\big(k^{\epsilon_0}\big)$-approximation for ED-$k$-NMC*, for some constant $\epsilon_0 > 0$, unless P=NP, and no $2^{\log^{1-\epsilon} |E_G|}$-approximation for any $\epsilon > 0$ unless NP is contained in deterministic quasipolynomial time.*

Proof : We first describe the construction and then argue the approximation-preservation property.

The construction. For each vertex $u \in U$ and for each label in $a \in L_1$ we introduce two vertices $a^u, \bar{a}^u$ in G connected by an edge of unit cost. Intuitively, if we select this edge in an ED-k-NMC solution for the instance we construct, this implies that we are selecting label a for u. Similarly, for each $w \in W$ and for each label $b \in L_2$ we introduce two vertices $b^w, \bar{b}^w$ connected by an edge of unit cost. The above edges will be the only ones having unit cost. All subsequent edges added to this construction will have infinite cost.

Consider an edge $e = (u, w)$ of H. For each such edge, we construct the following gadget. We introduce two nodes s_e, t_e that will form a terminal pair in our new instance. s_e and t_e are connected as follows. For each $b \in L_2$, let $L_b^e \subseteq L_1$ be the labels of L_1 such that $a \in L_b^e$ implies $\pi_e(a) = b$. Clearly, the sets L_b^e,

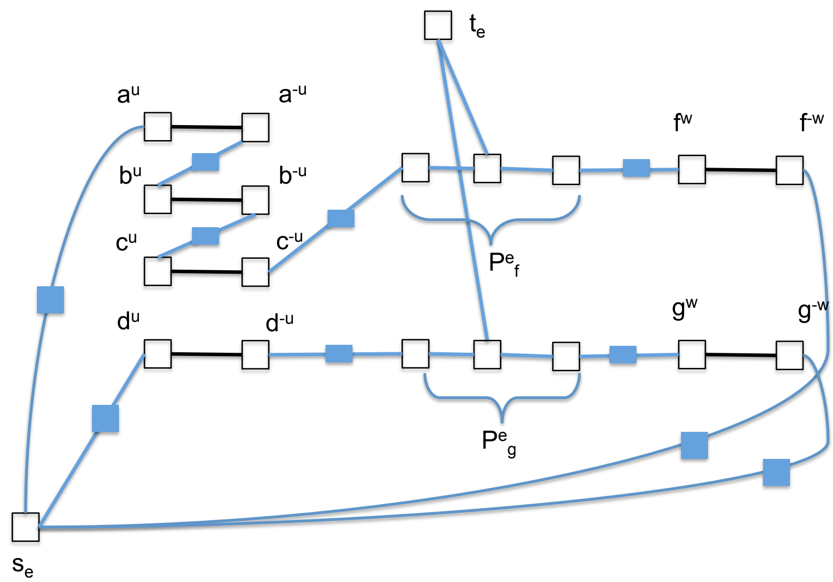

Figure B.3: The gadget $(N_e, E(N_e))$ introduced for an edge $e = (u, w)$. Here $L_1 = \{a, b, c, d\}, L_2 = \{f, g\}, L_f^e = \{a, b, c\}, L_g^e = \{d\}$. Each black edge has unit cost, while all other edges have ∞ cost. Blue (grey) rectangles indicate the nodes in $V(C_f^e) \cup V(C_g^e)$.

for all $b \in L_2$, form a partition of the labels L_1. For each non empty set L_b^e we add a path P_b^e of length 2 with the middle vertex connected to t_e. We then add edges to form a cycle C_b^e starting and ending at s_e, containing all the edges in L_b^e, the edges in the path P_b^e and the edge $b^w\bar{b}^w$. Finally we split each of these added edges into 2 by introducing a middle vertex. We let $V(C_b^e)$ we the the set of new vertices introduced by this operation, and let N_e be all the vertices participating in this gadget; see Fig. B.3.

Let $G' = \left(\bigcup_{e' \in E_H} N_{e'}, \bigcup_{e' \in E_H} E(N_{e'})\right)$ be the graph formed by the vertices and edges of all the edge gadgets. For all e, we are going to add other edges forming paths (of length 2) between s_e and t_e. We add edges (s_e, v), (v, t_e) for all vertices $v \in \Gamma_{G'}(N_e)$, that is, for all $v \notin N_e$ that are adjacent in some edge-gadget to some node in N_e. Note that $N_e \cap N_{e'} = \emptyset$ unless e and e' share an endpoint in H, say u, in which case, the two gadgets share the vertices $\{a^u, \bar{a}^u\}$ for all labels a of u. Thus, $|\Gamma_{G'}(N_e)| = O(\Delta q)$.

Define $k_e := |\{b \in L_2 : L_b^e \neq \emptyset\}| + |\Gamma_{G'}(N_e)|$. Finally, set $k := \max_e k_e$. For all edges e with $k_e < k$, we add $k - k_e$ new vertices and connect these to s_e and t_e (via ∞-cost edges). Let G be the resulting graph. This concludes our construction.

Approximation preservation. We now argue that any feasible solution to the ED-k-NMC-instance of finite cost yields a feasible solution to the MinRep-instance of no greater cost, and vice versa. This will complete the proof.

($\Rightarrow$) Let Z be a solution of finite cost for our ED-k-NMC instance. Consider a node $u \in V_H$ and let $L_u \in \{L_1, L_2\}$ be the label-set of u. Set $f(u) := \{a \in L_u : (a^u, \bar{a}^u) \in Z\}$. Clearly, the cost of the two solutions are the same. We now claim that the resulting labeling is feasible for the label-cover instance. Suppose not, then there is an edge $e \in E_H$ that is not covered. By our construction, this means that for each $b \in L_2$ with $L_b^e \neq \emptyset$, the subgraph of the remainder subgraph $\overline{G} = (V_G, E_G \setminus Z)$ induced by the nodes of the cycle C_b^e and t_e is connected. Each such cycle C_b^e, yields therefore one vertex-disjoint path in $\overline{G}$ between s_e and t_e. Also, all edges incident to s_e and t_e are present in $\overline{G}$ (since they have ∞ cost), so all length-2 paths in G between s_e and t_e are still present in $\overline{G}$. It follows that the vertex connectivity of s_e and t_e is at least k, a contradiction.

($\Leftarrow$) For the other direction, given a labeling for the label-cover instance, we construct $Z := \{(a^u, \bar{a}^u) : u \in V_H,\ a \in f(u)\}$. Clearly, the cost of the two solutions is the same. We claim that Z is a feasible solution to our ED-k-NMC instance. Suppose not. Then, for some $e = (u, w) \in E_H$, the s_e-t_e vertex connectivity in the remainder graph $\overline{G} = (V_G, E_G \setminus Z)$ is at least k. Therefore we can find a set of vertex-disjoint paths $\mathcal{P}$ between these vertices of size $|\mathcal{P}| \geq k$. Without loss of generality, we may assume that all the $k - |\{b \in L_2 : L_b^e \neq \emptyset\}|$ length-2 paths between s_e, t_e are in $\mathcal{P}$. If we remove the internal nodes on these length-2 paths from $\overline{G}$, the connected component containing s_e in the remaining portion of $\overline{G}$ is a subgraph of the gadget $(N_e, E(N_e))$ for edge e (shown in Fig. B.3). This means that this subgraph contains at least $|\{b \in L_2 : L_b^e \neq \emptyset\}|$ s_e-t_e vertex-disjoint paths. Clearly, this is only possible if, for every label b such that $L_b^e \neq \emptyset$, either $(b^w, \bar{b}^w) \notin Z$ or $(a^u, \bar{a}^u) \notin Z$ for every $a \in L_b^e$. But that means that the edge e is not covered by our labeling, a contradiction. ■