

Improved Approximation for Two-Edge-Connectivity*

Mohit Garg[†]

Fabrizio Grandoni[‡]

Afrouz Jabal Ameli[§]

Abstract

The basic goal of survivable network design is to construct low-cost networks which preserve a sufficient level of connectivity despite the failure or removal of a few nodes or edges. One of the most basic problems in this area is the 2-Edge-Connected Spanning Subgraph problem (2-ECSS): given an undirected graph G , find a 2-edge-connected spanning subgraph H of G with the minimum number of edges (in particular, H remains connected after the removal of one arbitrary edge).

2-ECSS is NP-hard and the best-known (polynomial-time) approximation factor for this problem is $4/3$. Interestingly, this factor was achieved with drastically different techniques by [Hunkenschroder, Vempala and Vetta '00,'19] and [Sebö and Vygen, '14]. In this paper we present an improved $\frac{118}{89} + \varepsilon < 1.326$ approximation for 2-ECSS.

The key ingredient in our approach (which might also be helpful in future work) is a reduction to a special type of structured graphs: our reduction preserves approximation factors up to $6/5$. While reducing to 2-vertex-connected graphs is trivial (and heavily used in prior work), our structured graphs are “almost” 3-vertex-connected: more precisely, given any 2-vertex-cut $\{u, v\}$ of a structured graph $G = (V, E)$, $G[V \setminus \{u, v\}]$ has exactly 2 connected components, one of which contains exactly one node of degree 2 in G .

*Partially supported by the SNSF Excellence Grant 200020B.182865 / 1 and the SNSF Grant 200021.200731 / 1.

[†]University of Bremen, University of Hamburg, Germany. E-mail: garg@uni-bremen.de.

[‡]IDSIA, USI-SUPSI, Switzerland. E-mail: fabrizio@idsia.ch.

[§]Eindhoven University of Technology, Netherlands. E-mail: a.jabal.ameli@tue.nl.

1 Introduction

Real-world networks are prone to failures. For this reason it is important to design them so that they are still able to support a given traffic despite a few (typically temporary) failures of nodes or edges. The basic goal of survivable network design is to construct cheap networks which are resilient to such failures. Most problems in this family are NP-hard, therefore it makes sense to develop approximation algorithms to address them. In this paper we study one of the most fundamental survivable network design problems, the 2-edge-connected¹ (2EC) spanning subgraph problem (2-ECSS): given an n -node undirected 2EC graph $G = (V, E)$, find a minimum-cardinality subset of edges $S \subseteq E$ such that $G' = (V, S)$ is 2EC².

2-ECSS is known to be MAX-SNP-hard [13, 17], in particular it does not admit a PTAS assuming $P \neq NP$. It is easy to compute a 2 approximation for this problem. For example it is sufficient to compute a DFS tree and augment it greedily. Khuller and Vishkin [27] found the first non-trivial $3/2$ -approximation algorithm. Cheriyan, Sebö and Szigeti [9] improved the approximation factor to $17/12$. The current best approximation factor for 2-ECSS is $4/3$, and this is achieved with two drastically different techniques by Hunkenschröder, Vempala and Vetta [24, 37]³ and Sebö and Vygen [34] (we will tell more about these approaches later on).

Several failed or incomplete attempts were made to improve the $4/3$ -approximation factor for 2-ECSS (see Section F for more details). This makes $4/3$ a natural approximation barrier to breach. In this paper we achieve this result (though with a small improvement).

THEOREM 1.1. *For any constant $\varepsilon > 0$, there is a deterministic polynomial-time $(\frac{118}{89} + \varepsilon)$ -approximation algorithm for 2-ECSS.*

An overview of our approach will be discussed in Section 2.

1.1 Related Work The k -edge connected spanning subgraph problem (k -ECSS) is the natural generalization of 2-ECSS to any connectivity $k \geq 2$ (see, e.g., [10, 20]).

A problem related to k -ECSS is the k -Connectivity Augmentation problem (k -CAP): given a k -edge-connected undirected graph G and a collection of extra edges L (*links*), find a minimum cardinality subset of links L' whose addition to G makes it $(k + 1)$ -edge-connected. It is known [14] that k -CAP can be reduced to the case $k = 1$, a.k.a. the Tree Augmentation problem (TAP), for odd k and to the case $k = 2$, a.k.a. the Cactus Augmentation problem (CacAP), for even k . Several approximation algorithms better than 2 are known for TAP [1, 7, 8, 16, 18, 23, 28, 29, 31], culminating with the current best 1.393 approximation [4]. Till recently no better than 2 approximation was known for CacAP (excluding the special case where the cactus is a single cycle [21]): the first such algorithm was described in [3], and later improved to 1.393 by [4]. A related connectivity augmentation problem is the Matching Augmentation problem (MAP). Some of the techniques used in this paper (in particular the credit scheme in the analysis) are closely related to recent work on MAP [2, 5, 6].

For all the mentioned problems one can define a natural weighted version. Here a general result by Jain [25] gives a 2 approximation, and in most cases this is the best known. The most remarkable exception is a very recent breakthrough 1.694 approximation for the Weighted version of TAP in [35] (later improved to $1.5 + \varepsilon$ in [36]). Partial results in this direction were achieved earlier in [1, 12, 18, 23, 32]. Achieving a better than 2 approximation for the weighted version of 2-ECSS is a major open problem closely related to this paper. The 2-approximation barrier was very recently breached [22] for the special case of weighted 2-ECSS where the edge weights are 0 or 1, also known as the Forest Augmentation Problem (FAP).

2 Overview of our Approach

Our basic approach is similar in spirit to the one of [24]. The high-level idea in their approach is to perform some approximation preserving reductions to obtain an input graph G which is conveniently structured. Then the authors compute an optimal (i.e. minimum-size) 2-edge-cover⁴. H of G , and then gradually transform H into a feasible solution S by adding and removing edges. This is done in such a way that the total number of edges of

¹A graph is k -edge-connected (k EC) if it is connected after the removal of up to $k - 1$ edges. A k -vertex-connected (k VC) graph is defined similarly w.r.t. the removal of $k - 1$ nodes.

² G' is a spanning subgraph, meaning that it connects all the nodes.

³[37] contains some bugs which were later fixed in [24].

⁴A 2-edge-cover F of G is a subset of edges such that each node has at least 2 edges of F incident to it. A minimum-size 2-edge-cover can be computed in polynomial time (see e.g., [33]).

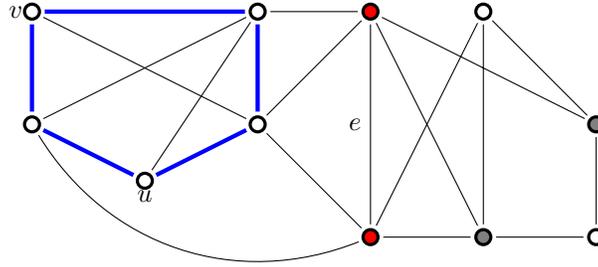


Figure 1: The 5-cycle C induced by blue edges is a $\frac{5}{4}$ -contractible subgraph: in particular, every 2EC spanning subgraph must contain at least 4 distinct edges with endpoints in $V(C)$. The red (resp., gray) nodes form a non-isolating (resp. isolating) 2-vertex-cut. The edge e is irrelevant.

the initial H grows at most by a factor of $4/3$. The claimed approximation ratio then derives from the fact that $|H|$ is a lower bound on the optimal solution size opt of a 2EC spanning subgraph (since any such solution must be a 2-edge-cover). Their analysis might also be rephrased in terms of a credit-based charging scheme analogously to [5, 6]: one can assign $\alpha = \frac{1}{3}$ credits to each edge of the initial H , and then use these credits to pay for the increase of the size of H .

We improve on the above approach mostly by exploiting more powerful reduction steps. This will allow us to use a smaller (average) value of α in the later stages, hence leading to a better approximation factor. Our construction can be also combined with the ear-decomposition based construction in [34], leading to a better approximation factor. We remark that we do not see a way to directly improve on $4/3$ by combining the ideas in [24] and [34] (namely, without passing through our structural results). We next provide some more details.

2.1 A Reduction to Structured Graphs Let $\varepsilon > 0$ be some constant which will appear in the approximation factor. W.l.o.g. we will always assume that the number of nodes of the input graph G is $\Omega(1/\varepsilon)$, otherwise we can solve 2-ECSS optimally by brute force in constant time. Analogously to [24] (and some prior work), we can reduce to the case that the input graph G is simple⁵ and 2VC⁶.

The authors of [24] define the notion of contractible cycles. We conveniently generalize this notion as follows (see Figure 1).

DEFINITION 1. (α -CONTRACTIBLE SUBGRAPH) Let $\alpha \geq 1$ be a fixed constant. A 2EC subgraph C of a 2EC graph G is α -contractible if every 2EC spanning subgraph of G contains at least $\frac{1}{\alpha}|E(C)|$ edges with both endpoints in $V(C)$.

Intuitively, an α -contractible subgraph C is a 2EC subgraph that we can safely contract with the goal of computing an α approximation for 2-ECSS. The first extra property that we enforce is that G does not contain α -contractible subgraphs of size at most $1/\varepsilon$.

The authors of [24] also show that it is possible to delete any edge uv incident to the two neighbours of a node w of degree two (*beta node*) provided that $G[V \setminus \{u, v\}]$ has at least 3 connected components (including w itself). We generalize their approach by introducing the notion of irrelevant edges and non-isolating 2-vertex-cuts (see Figure 1).

DEFINITION 2. (IRRELEVANT EDGE) Given a graph G and $e = uv \in E(G)$, we say that e is irrelevant if $\{u, v\}$ is a 2-vertex-cut of G .

We show that irrelevant edges can be deleted w.l.o.g. Hence in a structured instance we can assume that there are no such edges.

⁵Namely, there are no loops nor parallel edges

⁶A k -vertex cut is a subset of k nodes whose removal from G creates 2 or more connected components. A graph is k -vertex-connected (kVC) if there is no vertex cut of size at most $k - 1$.

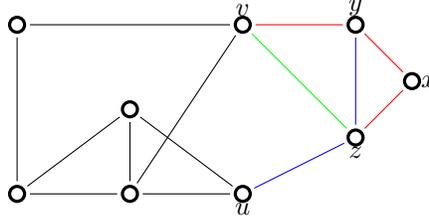


Figure 2: The black edges, red and blue edges, and red and green edges induce a subgraph of type A, B, and C, resp., w.r.t. the 2-vertex-cut $\{u, v\}$.

LEMMA 2.1. *Let $e = uv$ be an irrelevant edge of a 2VC simple graph $G = (V, E)$. Then there exists a minimum-size 2EC spanning subgraph of G not containing e .*

Proof. Assume by contradiction that all optimal 2EC spanning subgraphs of G contain e , and let OPT be one such solution. Define $\text{OPT}' = \text{OPT} \setminus \{e\}$. Clearly OPT' cannot be 2EC since this would contradict the optimality of OPT . Let (V_1, V_2) be any partition of $V \setminus \{u, v\}$, $V_1 \neq \emptyset \neq V_2$, such that there are no edges between V_1 and V_2 . Notice that at least one between $\text{OPT}'_1 := \text{OPT}'[V_1 \cup \{u, v\}]$ and $\text{OPT}'_2 := \text{OPT}'[V_2 \cup \{u, v\}]$, say OPT'_2 , needs to be connected otherwise OPT would not be 2EC. We also have that OPT'_1 is disconnected otherwise OPT' would be 2EC. More precisely, OPT'_1 consists of precisely two connected components $\text{OPT}'_1(u)$ and $\text{OPT}'_1(v)$ containing u and v , resp. Assume w.l.o.g. that $|V(\text{OPT}'_1(u))| \geq 2$. Observe that there must exist an edge f between $V(\text{OPT}'_1(u))$ and $V(\text{OPT}'_1(v))$: indeed otherwise u would be a 1-vertex-cut separating $V(\text{OPT}'_1(u)) \setminus \{u\}$ from $V \setminus (V(\text{OPT}'_1(u)) \cup \{u\})$. Thus $\text{OPT}'' := \text{OPT}' \cup \{f\}$ is an optimal 2EC spanning subgraph of G not containing e , a contradiction. \square

DEFINITION 3. (ISOLATING 2-VERTEX-CUT) *Given a graph G , and a 2-vertex-cut $\{u, v\} \subseteq V(G)$ of G , we say that $\{u, v\}$ is isolating if $G \setminus \{u, v\}$ has exactly two connected components, one of which is of size 1. Otherwise it is non-isolating.*

We show how to get rid of non-isolating 2-vertex-cuts and hence can assume that a structured instance does not contain such cuts. The construction here is slightly complex (details in Section A). The intuition is as follows: consider a non-isolating 2-vertex-cut $\{u, v\}$. W.l.o.g. we will assume $uv \notin E(G)$ since we can remove irrelevant edges. Let (V_1, V_2) be a partition of $V \setminus \{u, v\}$ with $|V_2| \geq |V_1| \geq 2$ and such that there are no edges between V_1 and V_2 . If $|V_1| \geq \Omega(1/\varepsilon)$, we contract $\{u, v\}$ into a single node: this leads to two independent problems that we can solve recursively. Then by adding up to 2 edges one can obtain a solution for the input problem. The cost of the extra two edges can be conveniently amortized over the cost of the optimal solution restricted to $V_1 \cup \{u, v\}$. Otherwise, consider the restriction OPT_i of the optimal solution OPT w.r.t. $G[V_i \cup \{u, v\}]$. Let $C(u)$ and $C(v)$ be the 2EC components of OPT_i containing u and v , resp. We observe that OPT_i must be of one of the following three types (see Figure 2) after collapsing each 2EC component C into a single super-node C (see also Definition 8 and Lemma A.4): (A) one super-node $C(u) = C(v)$ (i.e. OPT_i is 2EC); (B) a path of length at least 1 between $C(u)$ and $C(v)$; (C) two isolated super-nodes $C(u)$ and $C(v)$. Observe also that if OPT_1 is of type C then OPT_2 must be of type A and vice versa. Since V_1 has constant size, we can compute the cheapest subgraphs of type A, B, and C in $G[V_1 \cup \{u, v\}]$. Depending on the relative sizes of these 3 subgraphs and based on a careful case analysis (critically exploiting the fact that there are no α -contractible subgraphs of small size), we recurse on different subproblems over the nodes $V_2 \cup \{u, v\}$.

Altogether, we obtain the following definition and lemma (proof in Section A).

DEFINITION 4. ((α, ε)-STRUCTURED GRAPH) *Given $\alpha \geq 1$ and $\varepsilon > 0$, an undirected graph G is (α, ε) -structured if it is simple and 2VC, it contains at least $2/\varepsilon$ nodes and it does not contain:*

1. α -contractible subgraphs of size at most $1/\varepsilon$;
2. irrelevant edges;

3. non-isolating 2-vertex-cuts.

LEMMA 2.2. For all $\alpha \geq \frac{6}{5}$ and $\varepsilon \in (0, \frac{1}{24}]$, if there exists a deterministic polynomial-time α -approximation algorithm for 2-ECSS on (α, ε) -structured graphs, then there exists a deterministic polynomial-time $(\alpha + 2\varepsilon)$ -approximation algorithm for 2-ECSS.

When α and ε are clear from the context, we will simply talk about contractible subgraphs and structured graphs. Notice that we will present an algorithm with an approximation factor worse than $6/5$ (though better than $4/3$), hence by Lemma 2.2 it is sufficient to consider (α, ε) -structured graphs for some $\alpha < 4/3$. In particular, we will consider $(5/4, \varepsilon)$ -structured graphs for a sufficiently small constant $\varepsilon > 0$ ⁷. Lemma 2.2 could be a valid starting point for future research on 2-ECSS approximation, and therefore might be of independent interest.

One of the key properties of a structured graph (which we will exploit multiple times) is the existence of the following 3-matchings:

LEMMA 2.3. (3-MATCHING LEMMA) Let $G = (V, E)$ be a 2VC simple graph without irrelevant edges and without non-isolating 2-vertex-cuts. Consider any partition (V_1, V_2) of V such that for each $i \in \{1, 2\}$, $|V_i| \geq 3$ and if $|V_i| = 3$, then $G[V_i]$ is a triangle. Then, there exists a matching of size 3 between V_1 and V_2 .

Proof. Consider the bipartite graph F induced by the edges with exactly one endpoint in V_1 . Let M be a maximum (cardinality) matching of F . Assume by contradiction that $|M| \leq 2$. By König-Egeváry theorem⁸ there exists a vertex cover C of F of size $|M|$. We distinguish 3 subcases:

(1) $C = \{u\}$. Assume w.l.o.g. $u \in V_1$. Since C is a vertex cover of F , there are no edges in F (hence in G) between the non-empty sets $V_1 \setminus \{u\}$ and V_2 . Hence C is a 1-vertex-cut, a contradiction.

(2) $C = \{u, v\}$ where C is contained in one side of F . Assume w.l.o.g. $C \subseteq V_1$. Since $|V_1| \geq 3$ and C is a vertex cover of F , C is a 2-vertex-cut separating $V_1 \setminus C$ from V_2 . This implies that $uv \notin E(G)$ (otherwise uv would be an irrelevant edge). Since $G[V_1]$ is not a triangle, $|V_1 \setminus C| \geq 2$. This implies that C is a non-isolating 2-vertex-cut of G , a contradiction.

(3) $C = \{u, v\}$ where u and v belong to different sides of F . Assume w.l.o.g. $u \in V_1$ and $v \in V_2$. Consider the sets $V'_1 := V_1 \setminus \{u\}$ and $V'_2 := V_2 \setminus \{v\}$, both of size at least 2. Notice that there are no edges in F (hence in G) between V'_1 and V'_2 (otherwise C would not be a vertex cover of F). This implies that C is a non-isolating 2-vertex-cut of G , a contradiction. \square

2.2 A Canonical 2-Edge-Cover Let $G = (V, E)$ be an input instance of 2-ECSS which is $(\frac{5}{4}, \varepsilon)$ -structured for a sufficiently small constant $\varepsilon > 0$ to be fixed later (so that $|V(G)| \geq \frac{2}{\varepsilon}$ is large enough). In the following we will simply say structured instead of $(\frac{5}{4}, \varepsilon)$ -structured for shortness.

Analogously to [24], an initial step in our construction is to compute a minimum-size 2-edge-cover $H \subseteq E$. We call each (connected) component of H ⁹ which is 2EC, a *2EC component* of H . Given a component C of H which is *not* 2EC, we call any maximal 2EC subgraph of C containing at least 2 (hence 3) nodes a *block*, any node of C not contained in a block *lonely*, and any edge of C whose removal splits C into two components a *bridge* (see Figure 3). Notice that in our notation blocks and 2EC components of H are distinct. In particular the 2EC components, blocks, and bridges of H induce a partition of H . A block B of C is a *leaf block* if there exists precisely one bridge b of C incident to B , and an *inner block* otherwise.

Thanks to the fact that G is structured (and in particular exploiting the 3-matching Lemma 2.3), we can enforce that H is of the following canonical form (see Figure 3).

DEFINITION 5. [*canonical 2-edge-cover*] A 2-edge-cover H of G is canonical if the following properties hold:

(i) The 2EC components of H are i -cycles for $3 \leq i \leq 6$ or contain at least 7 edges.

⁷The reader may wonder why we set $\alpha = 5/4$ instead of $\alpha = 6/5$. The reason is that a larger value of α excludes more contractible subgraphs, hence making our analysis simpler. Since our approximation factor for $(5/4, \varepsilon)$ -structured instances is larger than $5/4$, this has no impact on the final approximation factor.

⁸This theorem states that, in a bipartite graph, the cardinality of a maximum matching equals the cardinality of a minimum vertex cover, see e.g. [33].

⁹Here and in similar cases in the rest of the paper, we mean the components of the corresponding subgraph (V, H) .

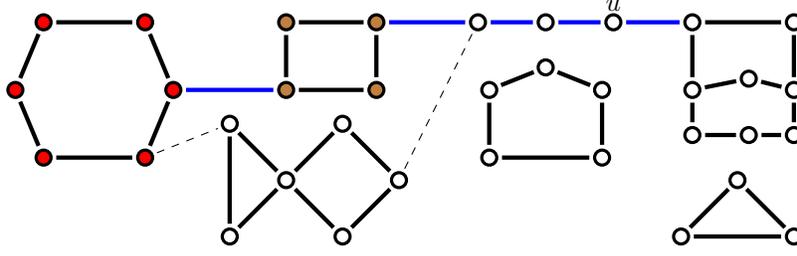


Figure 3: A canonical 2-edge-cover H . The blue edges are bridges. Node u is lonely. The red and brown nodes induce a leaf and inner block, resp., of the unique connected component C of S which is not 2EC. The dashed edges induce a (cheap) bridge-covering path P_C for C involving $bl = 2$ blocks of C , $br = 2$ bridges of C , and one connected component other than C .

- (ii) *The leaf blocks contain at least 6 edges and the inner ones at least 4 edges.*
- (iii) *(3-optimality) It is not possible to obtain from H a 2-edge-cover of the same size with fewer connected components by adding up to 3 edges and removing the same number of edges.*

LEMMA 2.4. *Given a 2-edge-cover H of a $(5/4, \varepsilon)$ -structured graph G , $\varepsilon \leq 1/5$, in polynomial time one can compute a 2-edge-cover H' of no larger size which is canonical.*

The proof of Lemma 2.4 is given in Section B. Enforcing property (iii) is trivial. Property (i) does not hold only if a 2EC component C of H is a bowtie (i.e., two triangles that share a common node) or a $K_{2,3}$ (i.e., a complete bipartite graph with one side of size 2 and the other of size 3). In both cases it is possible to reduce the number of edges or the number of connected components (using the fact that G is structured). In order to enforce (ii) we consider any block B with at most 5 edges and such that all the bridges incident to B are incident to a single node (this includes leaf blocks as a special case). The 3-matching Lemma 2.3 allows us to reduce the number of connected components or the number of bridges while not increasing the number of edges.

2.3 A Credit Assignment Scheme We will next assume that we are given a minimum-size 2-edge-cover H which is also canonical. The high-level structure of our algorithm is as follows. Starting from $S = H$, we gradually transform S by adding and removing edges, until S becomes a feasible solution. In order to keep the size of S under control, we define a credit assignment scheme that assigns credits $cr(S) \geq 0$ to S . Let us define the cost of S as $cost(S) = |S| + cr(S)$. Initially $cost(S) = cost(H) \leq \alpha|H|$ for some $\alpha < \frac{4}{3}$. During the process we iteratively replace S with a different solution S' satisfying $cost(S') \leq cost(S)$. Typically we make progress towards a feasible solution by guaranteeing that S' has fewer bridges and/or fewer connected components than S . This way at the end of the algorithm one obtains a feasible solution S of size $|S| \leq cost(S) \leq cost(H) \leq \alpha|H|$. This provides the desired $\alpha < \frac{4}{3}$ approximation since, as already observed, $|H| \leq \text{opt}$.

In a preliminary draft of this paper we initially assigned $\delta < \frac{1}{3}$ credits to each edge of H , so that $cost(H) < \frac{4}{3}\text{opt}$. The notation and case analysis however were extremely complex. So we decided to present a simpler (still unfortunately not very simple) approach which leads to a slightly worse approximation factor. This simplified approach uses two different algorithms and credit assignment schemes depending on the number of 2EC components of H which are triangles. In more detail, suppose that $t|H|$ edges of H belong to triangle 2EC components, and $b|H|$ edges of H are bridges. Obviously $0 \leq t + b \leq 1$. Then the following Lemma (proof in Section C) will give a better than 4/3 approximation when t is sufficiently close to 1 (namely, for any $t > 19/20$).

LEMMA 2.5. *Given a canonical minimum-size 2-edge-cover H of a structured graph G with $b|H|$ bridges and $t|H|$ edges belonging to triangle 2EC components. In polynomial time one can compute a $\frac{13}{8} - \frac{1}{3}t + \frac{1}{2}b$ approximate solution for 2-ECSS on G .*

In the complementary case (i.e. for t sufficiently close to 0, namely for any $t < 1$), we exploit the following Lemma (proof in Section D).

LEMMA 2.6. *Given a canonical minimum-size 2-edge-cover H of a structured graph G with $b|H|$ bridges and $t|H|$ edges belonging to triangle EC components. In polynomial time one can compute a $\frac{13}{10} + \frac{1}{30}t - \frac{1}{20}b$ approximate solution for 2-ECSS on G .*

Combining Lemmas 2.2, 2.4, 2.5, and 2.6, one obtains a $\min\{\frac{13}{8} - \frac{1}{3}t + \frac{1}{2}b, \frac{13}{10} + \frac{1}{30}t - \frac{1}{20}b\} + \varepsilon$ approximation for 2-ECSS. The reader might easily check that this is $\frac{117}{88} + \varepsilon$ in the worst case, hence already improving on $\frac{4}{3}$ for a small enough constant $\varepsilon > 0$.

The refined approximation factor claimed in Theorem 1.1 is based on a refinement of Lemma 2.5 to a $\frac{14}{9} - \frac{8}{27}t + \frac{4}{9}b$ approximation (Lemma E.1 in Section E) that exploits some ideas in [34]. The rest of this section is devoted to a high-level description of the proofs of Lemmas 2.5 and 2.6

2.4 Case of Many Triangles: Overview We sketch the proof of Lemma 2.5 (for the details see Section C). The first step of the construction is to remove all the bridges H_{brg} of H , hence obtaining a subgraph $S = H \setminus H_{brg}$ whose connected components are 2EC. We call the singleton nodes *lonely*, and we consider them as (degenerate) 2EC components.

Then there is a *gluing stage* where we gradually merge the 2EC components of S by adding and removing edges (keeping the invariant that each connected component of S is 2EC). In order to keep the size of S under control, we assign certain credits to the 2EC components of S as follows. We say that a 2EC component of S is *light* if it is a triangle connected component of H (hence of the initial S) and *heavy* otherwise. We remark that the gluing phase might create new triangle connected components, which are therefore heavy. Let $\mathcal{C}(S)$ be the 2EC components of S . We define a credit assignment scheme which assigns $cr(C)$ credits to each $C \in \mathcal{C}(S)$ according to the following simple rule:

- each light 2EC component C of S receives $cr(C) = \frac{1}{2}$ credits;
- each heavy 2EC component C of S receives $cr(C) = 2$ credits.

Let $cr(S) := \sum_{C \in \mathcal{C}(S)} cr(C)$ be the total amount of credits assigned to S , and define the cost of S as $cost(S) = |S| + cr(S)$. Recall that H contains $b|H|$ bridges and $t|H|$ edges in its triangle 2EC components. Let us upper bound the cost of the initial S . Recall that H is a canonical 2-edge-cover.

LEMMA 2.7. $cost(H \setminus H_{brg}) \leq (\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b)|H|$.

Proof. We initially assign $\frac{1}{6}$ credits to each edge of a triangle 2EC component of H , 2 credits to each bridge of H , and $\frac{1}{2}$ credits to the remaining edges in H . Hence the total number of credits is $(\frac{1}{6}t + 2b + \frac{1}{2}(1 - t - b))|H| = (\frac{1}{2} - \frac{1}{3}t + \frac{3}{2}b)|H|$. We next redistribute these credits to the 2EC components of $H \setminus H_{brg}$ as follows. Each light (triangle) component C of $H \setminus H_{brg}$ keeps the $\frac{1}{6} \cdot 3$ credits of its edges. Similarly each 2EC component C of $H \setminus H_{brg}$ deriving from a block of H or from a 2EC component of H which is not a triangle, keeps the credits of its edges (at least $\frac{1}{2} \cdot 4$ credits since $|E(C)| \geq 4$ being H canonical). Each bridge assigns its 2 credits to some lonely node: this is doable since the number of lonely nodes is upper bounded by $|H_{brg}|$. Altogether $cost(S) = |H| - |H_{brg}| + cr(S) \leq |H| - b|H| + (\frac{1}{2} - \frac{1}{3}t + \frac{3}{2}b)|H|$. \square

The gluing phase halts when we obtain an S which has the following special form (see Figure 4.b):

DEFINITION 6. (CORE-TRIANGLE 2-EDGE-COVER) *A 2-edge-cover of a graph G is core-triangle if it consists of $k+1$ 2EC components, namely one core component C plus k triangles T_1, \dots, T_k for $k \geq 0$. Furthermore, nodes of distinct triangles are not adjacent in G .*

We make progress towards a 2-edge-cover of the latter form by iteratively applying the following lemma.

LEMMA 2.8. *Suppose that S is not core-triangle. Then in polynomial time we can compute $S' \subseteq E(G)$ such that: (1) the connected components of S' are 2EC, (2) S' has fewer connected components than S , and (3) $cost(S') \leq cost(S)$.*

In order to prove the above lemma, we introduce the notion of merging cycle¹⁰ (see Figure 4.a).

¹⁰Here and in other parts of the paper we interpret a path or cycle as a subgraph, a subset of nodes or edges, or a sequence of nodes or edges, depending on which notation is more convenient.

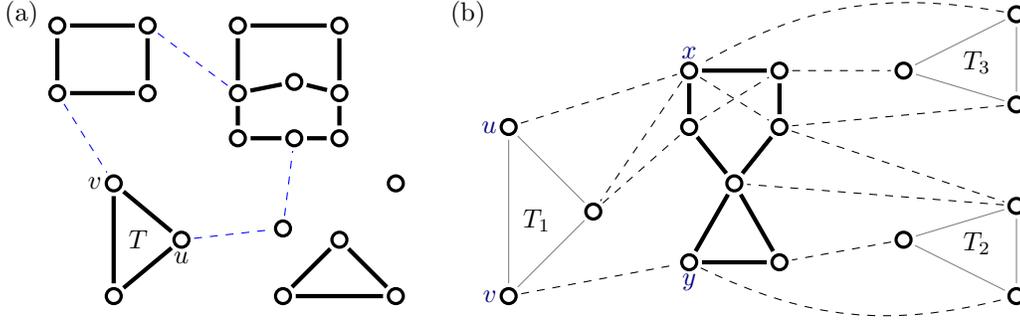


Figure 4: (a) The dashed blue edges induce a merging cycle M w.r.t. the current solution S given by black edges. Observe that all the connected components of S are 2EC. A new 2-edge-cover S' is obtained from S by adding M and removing uv . (b) The solid edges define a core-triangle 2-edge-cover S , where the core C is induced by the thick edges and the triangles T_1, T_2, T_3 by the gray edges. The remaining edges of the input graph G are dashed. Observe that C is 2EC and there are no edges in G between distinct triangles. A feasible solution can be obtained by adding the edges $\{ux, vy\}$ and removing the edge uv w.r.t. T_1 , and performing a similar update for the remaining 2 triangles.

DEFINITION 7. [*merging cycle*] A merging cycle M for S is a collection of edges $M \subseteq E(G) \setminus S$ such that: (1) M induces one cycle (excluding loops) in the graph obtained by collapsing the connected components $\mathcal{C}(S)$ of S ; (2) if two edges of M are incident to a light component $C \in \mathcal{C}(S)$, they are incident to two distinct nodes of C .

Given a merging cycle M of S , one naturally obtains a new S' as follows: we add the edges of M to S and, for each light (triangle) component C of S such that M is incident to the nodes u and v of C , we remove the edge uv . Notice that S' satisfies conditions (1) and (2) of Lemma 2.8. More specifically, all the 2EC components of S incident to M become a unique 2EC component C' in S' , and the remaining 2EC of S remain the same in S' . Let us analyze $cost(S')$. Suppose that M is incident to n_L light components of S and n_H heavy components of S . One has that $|S'| - |S| = n_H + n_L - n_L = n_H$. Furthermore $cr(S') - cr(S) \leq 2 - \frac{1}{2}n_L - 2n_H$, where the term 2 is due to $cr(C')$. Thus $cost(S) - cost(S') \geq n_H + \frac{1}{2}n_L - 2$. Notice that $n_H = 0$ and $n_L \leq 3$ is not possible since this would imply that the initial 2-edge-cover H is not canonical (in particular the 3-optimality property would not hold). A simple case analysis shows that $cost(S) - cost(S') \geq 0$ (hence (3) holds) unless $n_H = 1 = n_L$. In other words, the merging cycle M is incident to precisely one light and one heavy component: we call a merging cycle of the latter type *expensive*, and any other merging cycle *cheap*.

The proof of Lemma 2.8 proceeds constructively as follows. Given the current S (which is not core-triangle), we maintain an auxiliary forest $G' = (\mathcal{C}(S), F)$ whose nodes are the 2EC components of S and whose edges correspond to expensive merging cycles that we discovered along the way. Initially $F = \emptyset$. While G' consists of at least 2 trees we are able to compute efficiently a merging cycle M , exploiting the fact that G is structured (see Lemma C.1 in Section C). If the merging cycle M obtained this way is cheap, we use it to obtain an S' with the desired properties as described before. Otherwise, we add to F the edge between the two 2EC components of S (one heavy and one light) incident to M . Clearly the process ends in a polynomial number of steps by returning the desired cheap merging cycle or by obtaining a single tree G' .

In the latter case, we check whether there exists an edge e in G between two connected components C_1 and C_2 of S which are *not* adjacent in G' . If such an e exists, it is not hard to derive a cheap merging cycle involving all the (at least 3) components along the C_1 - C_2 path in F (see Lemma C.2 in Section C).

If there is no edge e of the above type, since by assumption S is not core-triangle, we can infer that there exists a light component C adjacent to two heavy components C_1 and C_2 in F . A simple case analysis shows that we can merge these 3 components together, hence obtaining the desired S' (see Lemma C.3 in Section C).

Once S is core-triangle, with core C and triangles T_1, \dots, T_k , we can transform it into a 2EC spanning subgraph APX via the following simple construction (see Figure 4.b). For each T_i , let $e_i = u_i u'_i$ and $f_i = v_i v'_i$ be two edges with $u_i, v_i \in V(T_i)$, $u_i \neq v_i$, and $u'_i, v'_i \in V(C)$. Notice that the pair e_i, f_i must exist otherwise some

node in $V(T_i)$ would be a 1-vertex-cut. Set $Q_i = E(T_i) \setminus \{u_i v_i\} \cup \{e_i, f_i\}$ and $\text{APX} := E(C) \cup \bigcup_{i=1}^k Q_i$. The proof of Lemma 2.5 follows easily.

Proof. [Proof of Lemma 2.5] Consider the feasible solution APX obtained with the above construction (which takes polynomial time). Observe that $|\text{APX}| = |S| + k$. By construction and due to Lemma 2.8 one has $\text{cost}(S) \leq \text{cost}(H \setminus H_{\text{brg}})$. Thus $\text{cost}(S) = |S| + cr(S) \leq \text{cost}(H \setminus H_{\text{brg}}) \leq (\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b)|H|$, where the last inequality comes from Lemma 2.7.

We also observe that $cr(S) \geq \frac{k}{2}$ where in the worst case each triangle T_i is light¹¹. Then $|\text{APX}| = |S| + k \leq (\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b)|H| + \frac{k}{2}$. Observe that any feasible solution must contain at least 4 distinct edges per triangle T_i and nodes in distinct triangles are not adjacent by Definition 6, hence $\text{opt} \geq 4k$. We already observed that $\text{opt} \geq |H|$. Altogether

$$|\text{APX}| \leq (\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b)\text{opt} + \frac{1}{8}\text{opt} = (\frac{13}{8} - \frac{1}{3}t + \frac{1}{2}b)\text{opt}.$$

□

In Section E we will present a more refined construction inspired by [34], and a corresponding refinement of Lemma 2.5 (namely, Lemma E.1). at a high level, instead of choosing the Q_i 's arbitrarily, we do that so as to minimize the number $\alpha(S)$ of connected components induced by $Q := \bigcup_i Q_i$ (this can be done in polynomial time analogously to [34] via a reduction to matroid intersection). Then, in a way similar to [34], we add a minimal set of edges F to Q to make $Q \cup F$ connected, and then add a minimum-size T -join J over the odd-degree nodes of $Q \cup F$ to obtain an alternative feasible solution $\text{APX}' := Q \cup F \cup J$. We show that APX' has size at most $4k + \alpha(S) + \frac{1}{2}|E(C)| - 1$, and that $\text{opt} \geq 4k + \alpha(S) - 1$. Taking the best solution among APX and APX' gives the refined bound.

2.5 Case of Few Triangles: Overview We next sketch the proof of Lemma 2.6. Let us initially set $S = H$. We gradually transform S by adding and deleting edges, until S becomes a 2EC spanning subgraph. During the process we will maintain the following invariant:

INVARIANT 1. *The 2EC components of S are an i -cycle for $3 \leq i \leq 6$ belonging to H or contain at least 7 edges (we call large the latter components).*

Clearly S initially satisfies this property since H is canonical. During our construction we always merge together different connected components of S so that the resulting connected component C contains at least 7 nodes (hence 7 edges if C is 2EC).

In order to keep the size of S under control, we use a credit assignment scheme analogously to Section 2.4, however more complex. In this scheme we assign certain credits to the connected components, bridges, and blocks of S . In more detail, we assign the following credits:

1. Each 2EC component C of S receives $cr(C) = 2$ credits if C is large (i.e., contains at least 7 edges), $cr(C) = 1$ credits if C is a triangle, and $cr(C) = \frac{3}{10}k$ credits if C is a k -cycle, $4 \leq k \leq 6$.
2. Each connected component C of S which is not 2EC receives $cr(C) = 1$ credits.
3. Each bridge e of S receives $cr(e) = \frac{1}{4}$ credits.
4. Each block B of S receives $cr(B) = 1$ credits.

Let $cr(S)$ be the total number of credits assigned to S according to the above rule, and define $\text{cost}(S) = |S| + cr(S)$ analogously to Section 2.4. The next lemma shows that $\text{cost}(S)$ is initially not too large. Its simple proof is based on initially assigning $\frac{1}{3}$, $\frac{1}{4}$ and $\frac{3}{10}$ credits to each edge in H which belongs to a triangle 2EC component, which is a bridge, and to the remaining edges, resp. We then redistribute these credits in a natural way, exploiting the fact that H is canonical.

LEMMA 2.9. *For $S = H$, $\text{cost}(S) \leq (\frac{13}{10} + \frac{1}{30}t - \frac{1}{20}b)|H|$.*

¹¹Indeed the construction can be slightly modified so that each heavy T_i is merged with the core C by exploiting the credits of T_i to pay for the extra needed edge.

The transformation of S into a feasible solution consists of two main phases. In an initial *bridge-covering* phase, we gradually reduce the number of bridges of S until all its connected components are 2EC. This is done by iteratively applying the following lemma.

LEMMA 2.10. [*Bridge-Covering*] Assume that S contains at least one bridge. Then in polynomial time one can compute a 2-edge-cover $S' \subseteq E(G)$ such that: (1) S' has fewer bridges than S ; (2) $\text{cost}(S') \leq \text{cost}(S)$.

The main building block in our construction is a *bridge-covering path* (see Figure 3). The idea is to consider any connected component C of S which is not 2EC, and collapse into a single node each block of C and any other connected component $C' \neq C$. Observe that C becomes a tree T_C . In the obtained graph G_C we consider any path P_C with edges not in $E(T_C)$, two distinct endpoints in $V(T_C)$, and all the internal nodes not in $V(T_C)$ (P_C might consist of a single edge). Consider $S' = S \cup E(P_C)$, which obviously has fewer bridges than S . W.r.t. S , we increase the number of edges by $|E(P_C)|$ and we create a new block B' in C (which requires 1 credit) or we transform C into a 2EC component C' (which requires 1 additional credit w.r.t. the credit owned by C). However we also decrease the credits by at least $|E(P_C)| - 1$ due to the connected components which are merged with C . Hence it is sufficient to collect a total of 2 credits among the bl blocks and br bridges of C that disappear in S' . In particular, to guarantee $\text{cost}(S) \geq \text{cost}(S')$, it is sufficient that $\frac{1}{4}br + bl - 2 \geq 0$. We say that P_C is cheap if the latter quantity is non-negative, and expensive otherwise. In particular, any P_C involving at least 2 block nodes or 1 block node and at least 4 bridges satisfies this condition. If there exists a cheap bridge-covering path P_C , we simply set $S' = S \cup E(P_C)$. Notice that this condition can be easily checked in polynomial time. If no cheap bridge-covering path exists, by exploiting the 3-matching Lemma 2.3 and a relatively simple case analysis, we show how to obtain the desired S' : to achieve that we will combine together up to two bridge-covering paths and/or we will remove some edges from S .

After the bridge-covering phase there is a *gluing* phase where we gradually merge together the 2EC components of S until we achieve a feasible solution. This is similar in spirit to the gluing phase of Section 2.4, with the difference that here we can directly reach a feasible solution without the need for a separate procedure to handle a core-triangle 2-edge-cover as done in Lemma 2.5. Intuitively this is due to the fact that each triangle T_i in a core-triangle 2-edge-cover now has 1 credit which is sufficient to glue it with the core C . In the gluing phase we iteratively apply the following lemma.

LEMMA 2.11. [*Gluing*] Assume that the connected components of S are 2EC, and there are at least 2 such components. Then in polynomial time one can compute a 2-edge-cover $S' \subseteq E(G)$ such that: (1) the connected components of S' are 2EC; (2) S' has fewer connected components than S ; (3) $\text{cost}(S') \leq \text{cost}(S)$.

In order to prove the above lemma we introduce an auxiliary graph \hat{G}_S (*component graph*) whose nodes are the 2EC components of S and where there is an edge C_1C_2 iff G contains at least one edge between the components C_1 and C_2 of S . A relatively simple case analysis allows us to rule out certain local configurations (or find the desired S' in polynomial time), see Lemmas D.3-D.9. Given that, it is relatively easy to find the desired S' in the case that \hat{G}_S is a tree (see Lemma D.11): here we exploit the 3-matching Lemma.

The most complicated part of our analysis is when \hat{G}_S is *not* a tree (see Lemma D.12). at a high level, the idea (see Figure 5) is to build a variant of the merging cycle M from Section 2.4 touching the components $C_0, \dots, C_{|M|}$ of S , where we guarantee that the two edges of M incident to C_i , $0 < i < |M|$, are incident to adjacent nodes in_i and out_i of C_i if C_i is a triangle or a 4-cycle. We obtain S' by adding M to S and removing the edges of type in_iout_i mentioned before. Notice that S' contains a 2EC component C' spanning the nodes of the components C_i . W.r.t. $\text{cost}(S)$, $\text{cost}(S')$ grows by $|M| + 2$ due to the edges of M and the 2 credits of C' . However, considering the removed credits and edges corresponding to each C_i , we obtain a decrease of $\text{cost}(S)$ by strictly more than 1 from each C_i , $0 < i < |M|$ (the smallest decrease is 1.5 when C_i is a 5-cycle). This allows one to easily rule out large values of $|M|$: for the remaining cases a more careful choice of M and a refined case analysis provides the desired S' .

The desired approximation factor follows easily from Lemmas 2.9, 2.10 and 2.11.

Proof. [Proof of Lemma 2.6] Consider the above process. Clearly it ends within a polynomial number of transformation steps of S . Let S' denote the final value of S . Observe that by construction $|S'| \leq |S| + cr(S') = \text{cost}(S') \leq \text{cost}(H) \leq (\frac{13}{10} + \frac{1}{30}t - \frac{1}{20}b)|H|$, where the first inequality comes from the fact that credits are non-negative and the last one from Lemma 2.9. The claim follows since $|H| \leq \text{opt}$. \square

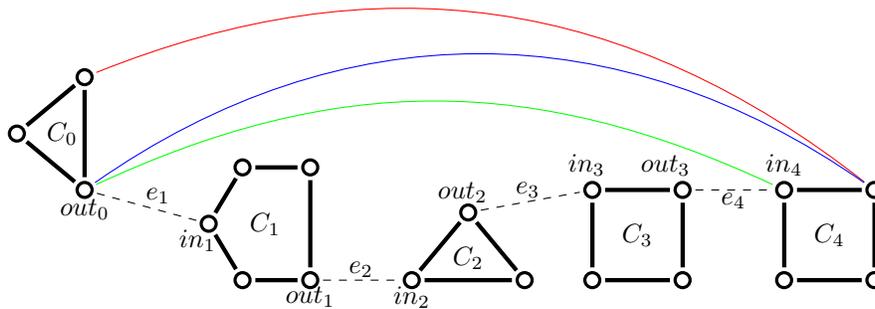


Figure 5: The dashed edges $\{e_1, \dots, e_4\}$ form a gluing path involving the component C_0, \dots, C_4 . Let e' be any edge among the green, blue, and red edges. By adding $\{e_1, e_2, e_3, e_4, e'\}$ to the components C_0, \dots, C_4 and removing in_2out_2 and in_3out_3 , one obtains a 2EC component C' spanning $V(C_0) \cup \dots \cup V(C_4)$. If e' is the blue (resp., red) edge, then we can remove from C' one extra edge (resp., two extra edges) while keeping it 2EC.

References

- [1] D. Adjiashvili. Beating approximation factor two for weighted tree augmentation with bounded costs. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2384–2399, 2017.
- [2] É. Bamas, M. Drygala, and O. Svensson. A simple LP-based approximation algorithm for the matching augmentation problem. In K. Aardal and L. Sanità, editors, *Integer Programming and Combinatorial Optimization - 23rd International Conference, IPCO 2022, Eindhoven, The Netherlands, June 27-29, 2022, Proceedings*, volume 13265 of *Lecture Notes in Computer Science*, pages 57–69. Springer, 2022.
- [3] J. Byrka, F. Grandoni, and A. Jabal Ameli. Breaching the 2-approximation barrier for connectivity augmentation: a reduction to steiner tree. In K. Makarychev, Y. Makarychev, M. Tulsiani, G. Kamath, and J. Chuzhoy, editors, *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 815–825. ACM, 2020.
- [4] F. Cecchetto, V. Traub, and R. Zenklusen. Bridging the gap between tree and connectivity augmentation: Unified and stronger approaches. *STOC 2021*, 2021.
- [5] J. Cheriyan, R. Cummings, J. Dippel, and J. Zhu. An improved approximation algorithm for the matching augmentation problem. *CoRR*, abs/2007.11559, 2020.
- [6] J. Cheriyan, J. Dippel, F. Grandoni, A. Khan, and V. V. Narayan. The matching augmentation problem: a $\frac{7}{4}$ -approximation algorithm. *Math. Program.*, 182(1):315–354, 2020.
- [7] J. Cheriyan and Z. Gao. Approximating (unweighted) tree augmentation via lift-and-project, part I: stemless TAP. *Algorithmica*, 80(2):530–559, 2018.
- [8] J. Cheriyan and Z. Gao. Approximating (unweighted) tree augmentation via lift-and-project, part II. *Algorithmica*, 80(2):608–651, 2018.
- [9] J. Cheriyan, A. Sebö, and Z. Szigeti. Improving on the 1.5-approximation of a smallest 2-edge connected spanning subgraph. *SIAM J. Discret. Math.*, 14(2):170–180, 2001.
- [10] J. Cheriyan and R. Thurimella. Approximating minimum-size k -connected spanning subgraphs via matching. *SIAM J. Comput.*, 30(2):528–560, 2000.
- [11] A. Civril. A new approximation algorithm for the minimum 2-edge-connected spanning subgraph problem. *CoRR*, 1911.07232v6, 2022.
- [12] N. Cohen and Z. Nutov. A $(1+\ln 2)$ -approximation algorithm for minimum-cost 2-edge-connectivity augmentation of trees with constant radius. *Theor. Comput. Sci.*, 489-490:67–74, 2013.
- [13] A. Czumaj and A. Lingas. On approximability of the minimum-cost k -connected spanning subgraph problem. In R. E. Tarjan and T. J. Warnow, editors, *Proceedings of the Tenth Annual ACM-SIAM Symposium on Discrete Algorithms, 17-19 January 1999, Baltimore, Maryland, USA*, pages 281–290. ACM/SIAM, 1999.
- [14] E. A. Dinits, A. V. Karzanov, and M. V. Lomonosov. On the structure of a family of minimal weighted cuts in a graph. *Studies in Discrete Optimization*, pages 290–306, 1976.
- [15] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In M. Jünger, G. Reinelt, and G. Rinaldi, editors, *Combinatorial Optimization - Eureka, You Shrink!, Papers Dedicated to Jack Edmonds, 5th International*

- Workshop, Aussois, France, March 5-9, 2001, Revised Papers*, volume 2570 of *Lecture Notes in Computer Science*, pages 11–26. Springer, 2001.
- [16] G. Even, J. Feldman, G. Kortsarz, and Z. Nutov. A 1.8 approximation algorithm for augmenting edge-connectivity of a graph from 1 to 2. *ACM Transactions on Algorithms*, 5(2):21:1–21:17, 2009.
- [17] C. G. Fernandes. A better approximation ratio for the minimum size k -edge-connected spanning subgraph problem. *J. Algorithms*, 28(1):105–124, 1998.
- [18] S. Fiorini, M. Groß, J. Könemann, and L. Sanità. Approximating weighted tree augmentation via Chvatal-Gomory cuts. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 817–831, 2018.
- [19] A. Frank. Conservative weightings and ear-decompositions of graphs. *Combinatorica*, 13:65–81, 1993.
- [20] H. N. Gabow and S. Gallagher. Iterated rounding algorithms for the smallest k -edge connected spanning subgraph. *SIAM J. Comput.*, 41(1):61–103, 2012.
- [21] W. Galvez, F. Grandoni, A. Jabal Ameli, and K. Sornat. On the cycle augmentation problem: Hardness and approximation algorithms. In *Workshop on Approximation and Online Algorithms (WAOA)*, 2019.
- [22] F. Grandoni, A. Jabal Ameli, and V. Traub. Breaching the 2-approximation barrier for the forest augmentation problem. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2022, Roma, Italy, June 20-24, 2022*. ACM, 2022.
- [23] F. Grandoni, C. Kalaitzis, and R. Zenklusen. Improved approximation for tree augmentation: saving by rewiring. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 632–645, 2018.
- [24] C. Hunkenschröder, S. S. Vempala, and A. Vetta. A $4/3$ -approximation algorithm for the minimum 2-edge connected subgraph problem. *ACM Trans. Algorithms*, 15(4):55:1–55:28, 2019.
- [25] K. Jain. A factor 2 approximation algorithm for the generalized steiner network problem. *Combinatorica*, 21(1):39–60, 2001.
- [26] R. Jothi, B. Raghavachari, and S. Varadarajan. A $5/4$ -approximation algorithm for minimum 2-edge-connectivity. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA*, pages 725–734. ACM/SIAM, 2003.
- [27] S. Khuller and U. Vishkin. Biconnectivity approximations and graph carvings. In S. R. Kosaraju, M. Fellows, A. Wigderson, and J. A. Ellis, editors, *Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada*, pages 759–770. ACM, 1992.
- [28] G. Kortsarz and Z. Nutov. Lp-relaxations for tree augmentation. In *Proceedings of 19th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, pages 13:1–13:16, 2016.
- [29] G. Kortsarz and Z. Nutov. A simplified 1.5-approximation algorithm for augmenting edge-connectivity of a graph from 1 to 2. *ACM Transactions on Algorithms (TALG)*, 12(2):23:1–23:20, 2016.
- [30] P. Krysta and V. S. A. Kumar. Approximation algorithms for minimum size 2-connectivity problems. In A. Ferreira and H. Reichel, editors, *STACS 2001, 18th Annual Symposium on Theoretical Aspects of Computer Science, Dresden, Germany, February 15-17, 2001, Proceedings*, volume 2010 of *Lecture Notes in Computer Science*, pages 431–442. Springer, 2001.
- [31] H. Nagamochi. An approximation for finding a smallest 2-edge-connected subgraph containing a specified spanning tree. *Discrete Applied Mathematics*, 126(1):83–113, 2003.
- [32] Z. Nutov. On the tree augmentation problem. In *25th Annual European Symposium on Algorithms, ESA 2017, September 4-6, 2017, Vienna, Austria*, pages 61:1–61:14, 2017.
- [33] A. Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.
- [34] A. Sebö and J. Vygen. Shorter tours by nicer ears: $7/5$ -approximation for the graph-tsp, $3/2$ for the path version, and $4/3$ for two-edge-connected subgraphs. *Comb.*, 34(5):597–629, 2014.
- [35] V. Traub and R. Zenklusen. A better-than-2 approximation for weighted tree augmentation. *CoRR*, abs/2104.07114, 2021.
- [36] V. Traub and R. Zenklusen. Local search for weighted tree augmentation and steiner tree. In J. S. Naor and N. Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 3253–3272. SIAM, 2022.
- [37] S. S. Vempala and A. Vetta. Factor $4/3$ approximations for minimum 2-connected subgraphs. In K. Jansen and S. Khuller, editors, *Approximation Algorithms for Combinatorial Optimization, Third International Workshop, APPROX 2000, Saarbrücken, Germany, September 5-8, 2000, Proceedings*, volume 1913 of *Lecture Notes in Computer Science*, pages 262–273. Springer, 2000.

A Reduction to Structured Graphs

In this section we show how to reduce the 2-ECSS problem to (α, ε) -structured graphs. An optimal solution to a 2-ECSS instance G will be denoted by $\text{OPT}(G)$ or simply by OPT when G is clear from the context. Likewise we will use $\text{opt}(G) = |\text{OPT}(G)|$ and $\text{opt} = |\text{OPT}|$.

Given $W \subseteq V$ we let $G|W$ denote the multi-graph obtained by contracting W into one node w . For a subgraph H of G , we use $G|H$ as a shortcut for $G|V(H)$. We will use the following facts and lemmas.

FACT A.1. *Let G be a 2EC graph and $W \subseteq V(G)$. Then $G|W$ is 2EC.*

FACT A.2. *Let H be a 2EC subgraph of a 2EC graph G , and S and S' be 2EC spanning subgraph of H and $G|H$, resp. Then $S \cup S'$ is a 2EC spanning subgraph of G , where we interpret S' as edges of G after decontracting $V(H)$.*

We will also use the following standard results (see, e.g., [24]) which will allow us to focus on 2VC simple graphs¹².

LEMMA A.1. *Let G be a 2EC graph, v be a 1-vertex-cut of G , and (V_1, V_2) , $V_1 \neq \emptyset \neq V_2$, be a partition of $V \setminus \{v\}$ so that there are no edges between V_1 and V_2 . Then $\text{opt}(G) = \text{opt}(G_1) + \text{opt}(G_2)$, where $G_i := G[V_i \cup \{v\}]$.*

LEMMA A.2. *Let G be a 2VC multigraph with $|V(G)| \geq 3$. Then, there exists a minimum size 2EC spanning subgraph of G that is simple (i.e. contains no self-loops nor parallel edges).*

We already showed how to handle irrelevant edges (Lemma 2.1). In order to handle the case of a non-isolating 2-vertex-cut $\{u, v\}$ we need to introduce some notation and technical lemmas.

LEMMA A.3. *Given a non-isolating 2 vertex-cut $\{u, v\}$ of G and assuming $|V(G)| \geq 6$, in polynomial time one can find a partition (V_1, V_2) of $V \setminus \{u, v\}$ with $2 \leq |V_1| \leq |V_2|$ such that there are no edges between V_1 and V_2 .*

Proof. Let C_1, \dots, C_k be the connected components of $G' := G \setminus \{u, v\}$ in non-decreasing order of size. If $k = 2$ simply set $V_i = V(C_i)$. Observe that $|V_1| \geq 2$ as required since otherwise $\{u, v\}$ would be isolating. If $k \geq 4$, set $V_1 = V(C_1) \cup V(C_2)$ and $V_2 = V(C_2) \cup \dots \cup V(C_k)$. Otherwise ($k = 3$) notice that $|V(C_3)| \geq 2$ since otherwise $|V(G)| = 5$. In this case set $V_1 = V(C_1) \cup V(C_2)$ and $V_2 = V(C_3)$. \square

For the following definition and lemma, see Figure 2.

DEFINITION 8. *Consider any graph H and two nodes $u, v \in V(C)$. Let H' be obtained by contracting each 2EC component C of H into a single super-node C , and let $C(u)$ and $C(v)$ be the super-nodes corresponding to u and v , resp. Then H is:*

1. of type A w.r.t. $\{u, v\}$ if H' consists of a single super-node $C(u) = C(v)$;
2. of type B w.r.t. $\{u, v\}$ if H' is a $C(u)$ - $C(v)$ path (of length at least one);
3. of type C w.r.t. $\{u, v\}$ if H' consists of two isolated super-nodes $C(u)$ and $C(v)$.

LEMMA A.4. *Let G be a simple 2VC graph without irrelevant edges, $\{u, v\}$ be a 2-vertex-cut of G , (V_1, V_2) , $V_1 \neq \emptyset \neq V_2$, be a partition of $V \setminus \{u, v\}$ such that there are no edges between V_1 and V_2 , and H be a 2EC spanning subgraph of G . Define $G_i = G[V_i \cup \{u, v\}]$ and $H_i = E(G_i) \cap H$ for $i \in \{1, 2\}$. The following claims hold:*

1. Both H_1 and H_2 are of type A, B or C (w.r.t. $\{u, v\}$). Furthermore, if one of the two is of type C, then the other must be of type A.
2. If H_i is of type C, then there exists one edge $f \in E(G_i)$ such that $H'_i := H_i \cup \{f\}$ is of type B. As a consequence, there exists a 2EC spanning subgraph H' where $H'_i := H' \cap E(G_i)$ is of type A or B.

¹²Part of our reductions temporarily create multi-graphs.

Proof. Notice that $uv \notin E(G)$ (hence $uv \notin E(H)$) since there are no irrelevant edges. (1) We prove the claim for H_1 , the other case being symmetric. First notice that if H_1 contains a connected component not containing u nor v , H would be disconnected. Hence H_1 consists of one connected component or two connected components $H_1(u)$ and $H_1(v)$ containing u and v , resp.

Suppose first that H_1 consists of one connected component. Let us contract the 2EC components of H_1 , hence obtaining a tree T . If T consists of a single node, H_1 is of type A. Otherwise, consider the path P (possibly of length 0) between the super-nodes resulting from the contraction of the 2EC components $C(u)$ and $C(v)$ containing u and v , resp.. Assume by contradiction that T contains an edge e not in P . Then e does not belong to any cycle of H , contradicting the fact that H is 2EC. Thus H_1 is of type B (in particular P has length at least 1 since H_1 is not 2EC).

Assume now that H_1 consists of 2 connected components $H_1(u)$ and $H_1(v)$. Let T_u and T_v be the two trees obtained by contracting the 2EC components of $H_1(u)$ and $H_2(v)$, resp. By the same argument as before, the 2-edge-connectivity of H implies that these two trees contain no edge. Hence H_1 is of type C.

The second part of the claim follows easily since if H_1 is of type C and H_2 is not of type A, then H would either be disconnected or it would contain at least one bridge edge (here we exploit the fact that $uv \notin E(H)$).

(2) There must exist an edge f in G_i between the two connected components of H_i since otherwise at least one of u or v would be a cut node. Clearly $H'_i := H_i \cup \{f\}$ satisfies the claim. For the second claim, consider any 2EC spanning subgraph H' and let $H'_i := H' \cap E(G_i)$. The claim holds if there exists one such H' where H'_i is of type A or B, hence assume this is not the case. Hence H'_i is of type C by property (1). The same argument as above implies the existence of an edge f in G_i such that $H''_i := H'_i \cup \{f\}$ is of type B, implying the claim for $H'' := H' \cup \{f\}$. \square

We are now ready to describe the reduction behind Lemma 2.2. We assume that we are given a polynomial time algorithm ALG that, given a (α, ε) -structured graph $G = (V, E)$, returns an α -approximate solution $\text{ALG}(G) \subseteq E$ to 2-ECSS on G . We consider the recursive procedure RED described in Algorithm 1: given a 2EC multi-graph $G = (V, E)$, RED returns a solution $\text{RED}(G) \subseteq E$ to 2-ECSS on G . RED uses ALG as a black-box subroutine. In Step 8 we interpret the edges of $\text{RED}(G|H)$ as the corresponding edges after decontracting H . We interpret similarly the edge sets H'_i from Step 15. We will show that the graphs G'_1 and G'_2 from Line 14, the graph G''_2 from Line 20, and the graph G'''_2 from Line 24 are all 2EC: this justifies the recursive use of RED in the following lines. We will also show that a valid F as in Steps 16 and 22 exists. The computation from Line 18 is performed as follows. First notice that G_1 contains a constant number of nodes, hence by brute force we can compute minimum-size subgraphs OPT_{1A} , OPT'_{1B} and OPT'_{1C} of the desired types, when any such subgraph exists. Obviously OPT_{1A} , if defined, forms a 2EC spanning subgraph with G_2 . If OPT'_{1C} is defined, it belongs to a 2EC spanning subgraph iff G_2 is 2EC (i.e., G_2 is a type A subgraph) by Lemma A.4. If the latter condition holds, we set $\text{OPT}_{1C} := \text{OPT}'_{1C}$. Otherwise, no type C subgraph of G_1 belongs to a feasible solution and we consider OPT_{1C} as undefined. Similarly, if OPT'_{1B} is defined, it belongs to a feasible solution iff G_2 is a type A or B subgraph. We set OPT_{1B} accordingly. Finally notice that if Step 26 is reached, then the graph is (α, ε) -structured.

LEMMA A.5. *RED runs in polynomial time in $|V(G)|$ if ALG does so.*

Proof. Let $n = |V(G)|$ and $m = |E(G)|$. Define $S = n^2 + m^2$ as the *size* of the problem. Clearly every step of the algorithm can be performed in polynomial time $\text{poly}(S)$ (which is polynomial in n) excluding the recursive calls. In particular, it is possible to check in polynomial time whether G contains an α -contractible subgraph with at most $1/\varepsilon$ nodes: simply enumerate over all the subsets W of nodes of the desired size; for each such W one can compute in polynomial time a minimum-size 2EC spanning subgraph OPT' of $G[W]$, if any such subgraph exists. Furthermore, one can compute in polynomial time a maximum cardinality subset of edges F with endpoints in W such that $G \setminus F$ is 2EC. Then it is sufficient to compare $|\text{OPT}'|$, in case OPT' is defined, with $|E(G[W]) \setminus F|$.

Let $T(S)$ be the running time of the algorithm as a function of S . Observe that in each step from a problem of size S we generate either a single subproblem of size $S' < S$ or two subproblems of size S' and S'' with $S', S'' < S$ and $S' + S'' \leq S$. Thus $T(S)$ satisfies $T(S) \leq \max\{T(S'), T(S') + T(S'')\} + \text{poly}(S)$ with $S', S'' < S$ and $S' + S'' \leq S$. In particular $S \cdot \text{poly}(S)$ is a valid upper bound for $T(S)$. \square

It is also not hard to prove that RED returns a feasible solution.

LEMMA A.6. *If G is 2EC, RED returns a 2EC spanning subgraph for G .*

Algorithm 1: Reduction from non-structured to structured instances of 2-ECSS. Here ALG is an algorithm for structured instances. We assume w.l.o.g. that $\varepsilon > 0$ is a small enough constant, so that $2/\varepsilon$ is lower bounded by a sufficiently large constant. G is assumed to be 2EC.

```

1 if  $|V(G)| \leq \frac{2}{\varepsilon}$  then
2    $\lfloor$  compute  $\text{OPT}(G)$  by brute force (in constant time) and return  $\text{OPT}(G)$ 
3 if  $G$  has a 1-vertex-cut  $v$  then
4    $\lfloor$  let  $(V_1, V_2)$ ,  $V_1 \neq \emptyset \neq V_2$ , be a partition of  $V \setminus \{v\}$  such that there are no edges between  $V_1$  and  $V_2$ .
5     return  $\text{RED}(G[V_1 \cup \{v\}]) \cup \text{RED}(G[V_2 \cup \{v\}])$ 
6 if  $G$  contains a self loop or a parallel edge  $e$  then
7    $\lfloor$  return  $\text{RED}(G \setminus \{e\})$ 
8 if  $G$  contains an  $\alpha$ -contractible subgraph  $H$  with  $|V(H)| \leq 1/\varepsilon$  then
9    $\lfloor$  return  $H \cup \text{RED}(G|H)$ .
10 if  $G$  contains an irrelevant edge  $e$  then
11    $\lfloor$  return  $\text{RED}(G \setminus \{e\})$ 
12 if  $G$  contains a non-isolating 2-vertex-cut  $\{u, v\}$  then
13   let  $(V_1, V_2)$ ,  $2 \leq |V_1| \leq |V_2|$  be a partition of  $V \setminus \{u, v\}$  as in Lemma A.3
14   let  $G_i := G[V_i \cup \{u, v\}]$  for  $i \in \{1, 2\}$  if  $|V_1| > \frac{1}{\varepsilon} - 2$  then
15     let  $G'_i$  for  $i \in \{1, 2\}$  be obtained from  $G_i$  by contracting  $\{u, v\}$  into one node
16     let  $H'_i = \text{RED}(G'_i)$  for  $i \in \{1, 2\}$ 
17     let  $F' \subseteq E$ ,  $|F'| \leq 2$ , such that  $H' := H'_1 \cup H'_2 \cup F'$  is 2EC. return  $H'$ 
18 else
19   let  $\text{OPT}_{1A}$ ,  $\text{OPT}_{1B}$ , and  $\text{OPT}_{1C}$  be minimum-size subgraphs of  $G_1$  of type  $A$ ,  $B$  and  $C$  belonging
20     to some 2EC spanning subgraph, resp., if any such subgraph exists. Let  $\text{OPT}_{1min}$  be the
21     subgraph of type  $\text{OPT}_{1X}$  which is defined and has minimum-size, where ties are broken by
22     preferring  $\text{OPT}_{1A}$ ,  $\text{OPT}_{1B}$  and  $\text{OPT}_{1C}$  in this order
23   if  $\text{OPT}_{1C}$  is defined and  $|\text{OPT}_{1C}| \leq |\text{OPT}_{1B}| - 1$  then
24     let  $G''_2 := (V(G_2), E(G_2) \cup \{uv\})$  where  $uv$  is a dummy edge
25     let  $H''_2 := \text{RED}(G''_2) \setminus \{uv\}$ 
26     let  $F'' \subseteq E$ ,  $|F''| \leq 1$ , such that  $H'' := \text{OPT}_{1min} \cup H''_2 \cup F''$  is 2EC. return  $H''$ 
27   else
28     let  $G'''_2 := (V(G_2) \cup \{w\}, E(G_2) \cup \{wu, wv\})$  where  $w$  is a dummy node and  $wu$  and  $wv$  are
29     dummy edges
30     let  $H'''_2 := \text{RED}(G'''_2) \setminus \{wu, wv\}$ . return  $H''' := \text{OPT}_{1min} \cup H'''_2$ 
31 return  $\text{ALG}(G)$ 

```

Proof. Let us prove the claim inductively for increasing values of the vector $(|V(G)|, |E(G)|)$ (sorted lexicographically). The base cases are given by Lines 2 and 26, where the claim holds trivially and by the correctness of ALG, resp. If Line 4 is executed, both G_1 and G_2 (which must be 2EC since G is so) contain at most $|V(G)| - 1$ nodes, hence by inductive hypothesis $\text{RED}(G_1)$ and $\text{RED}(G_2)$ are 2EC spanning subgraphs of $V_1 \cup \{v\}$ and $V_2 \cup \{v\}$, resp. Thus their union is a 2EC spanning subgraph for G . If Line 6 is executed, $G' := G \setminus \{e\}$ must be 2EC by Lemma A.2, and $(|V(G')|, |E(G')|) = (|V(G)|, |E(G)| - 1)$. Hence by inductive hypothesis $\text{RED}(G')$ is a 2EC spanning subgraph for G' , hence for G . The case when Line 10 is executed can be handled similarly as in the previous case, using Lemma 2.1. When Line 8 is executed, the claim follows from Fact A.2.

It remains to consider the case when the condition of Line 11 holds. We remark that in this case G is 2VC since the condition of Line 3 does not hold. Hence Lemma A.4 applies. We also notice that $G'_1, G'_2, G''_2,$ and G'''_2 all have strictly fewer nodes than G (since $|V_1| \geq 2$), which will be needed to apply the inductive hypothesis.

Suppose first that the condition of Line 13 holds. Notice that G'_1 and G'_2 are 2EC by Fact A.1 and, implicitly, by Lemma A.1 (since the contracted node becomes a 1-vertex-cut). Hence the inductive hypothesis holds for these graphs. Let us argue that a proper F' exists. Notice that each H'_i must be of type A, B or C (otherwise $\text{RED}(G'_i)$ would not be 2EC). If one of them is of type A or they are both of type B, then $H'_1 \cup H'_2$ is 2EC. If one of them is of type B, say H'_1 , and the other of type C, say H'_2 , then let $H'_2(u)$ and $H'_2(v)$ be the two 2EC components of H'_2 containing u and v resp. By Lemma A.4.2 there must exist an edge $f \in G_2$ between $H'_2(u)$ and $H'_2(v)$. Hence $H'_1 \cup H'_2 \cup \{f\}$ is 2EC. The remaining case is that H'_1 and H'_2 are both of type C. In this case $H'_1 \cup H'_2$ consists of precisely two 2EC components $H'(u)$ and $H'(v)$ containing u and v , resp. Since G is 2EC, there must exist two edges f and g between $H'(u)$ and $H'(v)$. Thus $H'_1 \cup H'_2 \cup \{f, g\}$ is 2EC. In all the cases the desired F' exists.

Suppose next that the condition of Line 19 holds. Observe that G_2 must contain a subgraph of type A or B by Lemma A.4.2. As a consequence G''_2 is 2EC as required. Let us argue that a proper F'' exists. Notice that H''_2 must be of type A or B. By the tie-breaking rule of Line 18, $\text{OPT}_{1\min}$ must be of type A or C. In the first case $\text{OPT}_{1\min} \cup H''_2$ is 2EC. In the second case by Lemma A.4.2 there exists one edge f such that $\text{OPT}_{1\min} \cup \{f\}$ is of type B for G_1 , implying that $\text{OPT}_{1\min} \cup H''_2 \cup \{f\}$ is 2EC.

It remains to consider the case when the condition of Line 23 holds. Again by Lemma A.4.2, G_2 must contain a subgraph of type A or B, hence G'''_2 is 2EC as required. Notice that $\text{RED}(G'''_2)$ must contain the two dummy edges wv and wu , which induce a type B graph. Hence by Lemma A.4.1, H'''_2 must be of type A or B. Notice also that by the tie-breaking rule of Line 18 and given that the condition of Line 19 does not hold, $\text{OPT}_{1\min}$ must be of type A or B. Hence $\text{OPT}_{1\min} \cup H'''_2$ is 2EC as desired. \square

It remains to upper bound the approximation factor of the algorithm. Let $\text{red}(G) = |\text{RED}(G)|$. The following claim implies the approximation factor of RED since trivially $|V(G)| \leq \text{opt}(G)$.

LEMMA A.7. *One has $\text{red}(G) \leq \begin{cases} \text{opt}(G) & \text{if } |V(G)| \leq \frac{2}{\epsilon}; \\ \alpha \cdot \text{opt}(G) + 2\epsilon \cdot |V(G)| - 2 & \text{if } |V(G)| > \frac{2}{\epsilon}. \end{cases}$*

Proof. Let us prove the claim by induction on $(|V(G)|, |E(G)|)$ similarly to Lemma A.6. The base case is when Line 2 is executed: here the claim trivially holds. For the inductive hypothesis, we distinguish a few cases depending on which line of RED is executed. If Line 6 is executed we consider a graph $G' := G \setminus \{e\}$ with the same number of nodes as $V(G)$ and the same optimal size by Lemma A.2. The claim follows by inductive hypothesis. If Line 10 is executed, the claim follows similarly using Lemma 2.1. In Line 26 is executed, one has $\text{red}(G) \leq \alpha \text{opt}(G) \leq \alpha \text{opt}(G) + 2\epsilon |V(G)| - 2$ since ALG is α approximate and $|V(G)| > \frac{2}{\epsilon}$.

If Line 4 is executed, recall that by Lemma A.1 one has $\text{opt}(G) = \text{opt}(G_1) + \text{opt}(G_2)$. We distinguish a few cases depending on the sizes of V_1 and V_2 . Assume w.l.o.g. $|V_1| \leq |V_2|$. If $|V_2| \leq \frac{2}{\epsilon} - 1$ then $\text{red}(G) = \text{opt}(G_1) + \text{opt}(G_2) = \text{opt}(G) \leq \alpha \text{opt}(G) + 2\epsilon |V(G)| - 2$, where the last inequality follows from $\alpha \geq 1$ and $|V(G)| \geq \frac{2}{\epsilon}$. Otherwise, if $|V_1| \leq \frac{2}{\epsilon} - 1$, by inductive hypothesis one has $\text{red}(G) \leq \text{opt}(G_1) + \alpha \text{opt}(G_2) + 2\epsilon |V(G_2)| - 2 \leq \alpha \text{opt}(G) + 2\epsilon |V(G)| - 2$, where we used $\alpha \geq 1$ and $|V(G_2)| \leq |V(G)|$. The remaining case is $|V_1| \geq \frac{2}{\epsilon}$. In this case $\text{red}(G) \leq \alpha \text{opt}(G_1) + \alpha \text{opt}(G_2) + 2\epsilon (|V(G_1)| + |V(G_2)|) - 4 = \alpha \text{opt}(G) + 2\epsilon |V(G)| + 2\epsilon - 4 \leq \alpha \text{opt}(G) + 2\epsilon |V(G)| - 2$, where we used $2\epsilon \leq 2$.

If Line 8 is executed, one has $\text{opt}(G) = |\text{OPT}(G) \cap \binom{V(H)}{2}| + |\text{OPT}(G) \setminus \binom{V(H)}{2}| \geq \frac{1}{\alpha} |H| + \text{opt}(G|H)$ by the definition of H and observing that $\text{OPT}(G) \setminus \binom{V(H)}{2}$ induces a feasible 2EC spanning subgraph of $G|H$. If

$|V(G|H)| \leq \frac{2}{\varepsilon}$ one has $\text{red}(G) = |H| + \text{opt}(G|H) \leq \alpha \text{opt}(G) \leq \alpha \text{opt}(G) + 2\varepsilon|V(G)| - 2$ since $|V(G)| \geq 2/\varepsilon$. Otherwise $\text{red}(G) \leq |H| + \alpha \text{opt}(G|H) + 2\varepsilon|V(G|H)| - 2 \leq \alpha \text{opt}(G) + 2\varepsilon|V(G)| - 2$ since $|V(G)| \geq |V(G|H)|$.

We next assume that the condition of Line 11 holds. Let $\text{OPT}_i = \text{OPT}(G) \cap E(G_i)$ and $\text{opt}_i = |\text{OPT}_i|$. Assume first that the condition of Line 13 holds. Observe that OPT_i induces a feasible solution for G'_i . Thus by inductive hypothesis

$$|H'| = |H'_1| + |H'_2| + |F| \leq \alpha(\text{opt}_1 + \text{opt}_2) + 2\varepsilon(|V_1| + |V_2| + 2) - 4 + 2 = \alpha \text{opt}(G) + 2\varepsilon|V(G)| - 2.$$

We next assume that one of the conditions of Lines 19 or 23 holds. Let OPT_{iX} be a minimum-size subgraph of type $X \in \{A, B, C\}$ in G_i and set $\text{opt}_{iX} = |\text{OPT}_{iX}|$ (is there is no such subgraph, $\text{opt}_{iX} = +\infty$). We will next assume w.l.o.g. that OPT_1 is of type B or C (hence OPT_2 of type A or B by Lemma A.4.1) in view of the following claim.

CLAIM 1. *There exists an optimal solution $\text{OPT}(G)$ such that OPT_1 is of type B or C .*

Proof. Let us assume that OPT_1 is of type A , otherwise the claim holds. Notice that every feasible solution must use at least $\min\{\text{opt}_{1A}, \text{opt}_{1B}, \text{opt}_{1C}\}$ edges from G_1 by Lemma A.4.1. Hence $\text{opt}_{1A} > \alpha \cdot \min\{\text{opt}_{1B}, \text{opt}_{1C}\}$ since otherwise OPT_{1A} would be an α -contractible subgraph on at most $1/\varepsilon$ nodes (which is excluded by the check of Line 7). If $\text{opt}_{1A} \geq \text{opt}_{1B} + 1$, then we obtain an alternative optimum solution with the desired property by taking $\text{OPT}_{1B} \cup \text{OPT}_2$, and, in case OPT_2 is of type C , by adding one edge f between the connected components of OPT_2 whose existence is guaranteed by Lemma A.4.2. Otherwise ($\text{opt}_{1A} \leq \text{opt}_{1B}$), we must have $\text{opt}_{1A} > \alpha \text{opt}_{1C}$, hence in particular $\text{opt}_{1A} \geq \text{opt}_{1C} + 1 \geq 4$ since $\alpha \geq 1$ and $\text{opt}_{1C} \geq 3$. Observe however that Lemma A.4.2 implies $\text{opt}_{1B} \leq \text{opt}_{1C} + 1$. Altogether $\text{opt}_{1A} = \text{opt}_{1B} = \text{opt}_{1C} + 1$. Then $\text{opt}_{1A} > \alpha \text{opt}_{1C} = \alpha(\text{opt}_{1A} - 1)$ implies $\text{opt}_{1A} \leq 5$ since $\alpha \geq 6/5$. Altogether $\text{opt}_{1A} \in \{4, 5\}$.

If $\text{opt}_{1A} = 4$, OPT_{1A} has to be a 4-cycle C of type u, a, v, b, u . Notice that the edge ab cannot exist since otherwise $\text{opt}_{1B} = 3$ due to the path u, a, b, v . Then every feasible solution must include the 4 edges of C (to guarantee degree at least 2 on a and b), which makes C an α -contractible subgraph on $4 \leq 1/\varepsilon$ nodes: this is excluded by the check of Line 7.

The other possible case is $\text{opt}_{1A} = 5$. The minimality of OPT_{1A} implies that OPT_{1A} is a 5-cycle C : w.l.o.g. let $C = u, a, v, b, c, u$. Observe that the edges ab cannot exist, since otherwise $\{va, ab, bc, cu\}$ would be a type B graph in G_1 of size 4, contradicting $\text{opt}_{1B} = \text{opt}_{1A} = 5$. A symmetric construction shows that edge ac cannot exist. Hence every feasible solution restricted to G_1 must include the edges $\{au, av\}$ and furthermore at least 3 more edges incident to b and c , so at least 5 edges altogether. This implies that C is an α -contractible subgraph of size $5 \leq 1/\varepsilon$: this is excluded by the check of Line 7. \square

Claim 1 implies $\text{opt}_2 \geq \min\{\text{opt}_{2A}, \text{opt}_{2B}\}$ and $\text{opt}_1 \geq \min\{\text{opt}_{1B}, \text{opt}_{1C}\}$. By the check of Line 7 we also have that $\text{opt}_{1A} > \frac{1}{\alpha} \text{opt}_1 \geq \frac{1}{\alpha} \min\{\text{opt}_{1B}, \text{opt}_{1C}\}$, hence in particular $\text{OPT}_{1\min}$ is of type B or C in this case. More precisely, by the tie-breaking rule, $\text{OPT}_{1\min}$ is of type C if the condition of Line 19 holds, and of type B otherwise (i.e. if the condition of Line 23 holds).

Suppose that the condition of Line 19 holds. In this case $\text{OPT}_{1\min}$ is of type B by the tie-breaking rule and the previous arguments, hence $|H_{1\min}| = \text{opt}_{1B} \leq \text{opt}_1$. Notice also that OPT_{2A} and OPT_{2B} plus the two dummy edges are feasible solutions for G'_2 , thus $\text{opt}'_2 \leq \min\{\text{opt}_{2A}, \text{opt}_{2B}\} + 2 \leq \text{opt}_2 + 2$. Thus

$$\begin{aligned} |H'''| &= |\text{OPT}_{1\min}| + |H_2'''| - 2 \leq \text{opt}_1 + \alpha(\text{opt}_2 + 2) + 2\varepsilon(|V_2| + 3) - 2 - 2 \\ &\leq \alpha \text{opt} + 2\varepsilon|V(G)| - 2 + (2\alpha - 2 - (\alpha - 1)\text{opt}_1) \leq \alpha \text{opt} + 2\varepsilon|V(G)| - 2, \end{aligned}$$

where in the last inequality we used the obvious fact that $\text{opt}_1 \geq 2$.

It remains to consider the case that the condition of Line 23 holds. Recall that in this case $\text{OPT}_{1\min}$ is of type C . Remember also that OPT_2 must be of type A or B . Suppose first that OPT_2 is of type A . Then OPT_1 has to be of type C (otherwise $\text{OPT}(G)$ would not be optimal), implying $\text{opt} = \text{opt}_{1C} + \text{opt}_{2A}$. Notice that OPT_2 is a feasible solution for G'_2 , hence $|\text{OPT}(G'_2)| \leq \text{opt}_{2A}$. Then

$$|H''| \leq \text{opt}_{1\min} + \alpha \text{opt}_{2A} + 2\varepsilon(|V_2| + 2) - 2 - 1 + 1 \leq \alpha \text{opt} + 2\varepsilon|V(G)| - 2.$$

Otherwise, i.e. OPT_2 is of type B , OPT_1 must be of type A or B , hence of type B by Claim 1. Hence $\text{opt} = \text{opt}_{1B} + \text{opt}_{2B}$. Notice that $\text{OPT}_2 \cup \{uv\}$ is a feasible solution for G'_2 , hence $|\text{OPT}(G'_2)| \leq \text{opt}_{2B} + 1$.

Since $\text{opt}_{1\min} = \text{opt}_{1C} \leq \text{opt}_{1B} - 1$, one has

$$\begin{aligned} |H''| &\leq \text{opt}_{1B} - 1 + \alpha(\text{opt}_{2B} + 1) + 2\varepsilon(|V_2| + 2) - 2 - 1 + 1 \\ &\leq \alpha \text{opt} + 2\varepsilon|V(G)| - 2 - (\alpha - 1)(\text{opt}_{1B} - 1) \leq \alpha \text{opt} + 2\varepsilon|V(G)| - 2. \end{aligned}$$

□

B Finding a Canonical 2-Edge-Cover

In this section we prove Lemma 2.4, which we restate here together with the definition of canonical 2-edge-cover.

DEFINITION 5. [*canonical 2-edge-cover*] A 2-edge-cover H of G is canonical if the following properties hold:

- (i) The 2EC components of H are i -cycles for $3 \leq i \leq 6$ or contain at least 7 edges.
- (ii) The leaf blocks contain at least 6 edges and the inner ones at least 4 edges.
- (iii) (3-optimality) It is not possible to obtain from H a 2-edge-cover of the same size with fewer connected components by adding up to 3 edges and removing the same number of edges.

LEMMA 2.4. Given a 2-edge-cover H of a $(5/4, \varepsilon)$ -structured graph G , $\varepsilon \leq 1/5$, in polynomial time one can compute a 2-edge-cover H' of no larger size which is canonical.

Proof. Starting with $H' = H$, while H' is not canonical we perform one of the following updates, in the specified order of priority:

- (1) If there exists $e \in H'$ such that $H' \setminus \{e\}$ is a 2-edge-cover, replace H' with $H' \setminus \{e\}$.
- (2) If there exist $D \subseteq H'$ and $A \subseteq E(G) \setminus H'$, with $|D| = |A| \leq 3$, such that $H' \setminus D \cup A$ is a 2-edge-cover with fewer components than H' , replace H' with $H' \setminus D \cup A$.
- (3) Suppose that there exists a 2EC component C of H' with fewer than 7 edges that is not an $|V(C)|$ -cycle. Notice that $|E(C)| \geq |V(C)|$ since otherwise C would not be 2EC. This implies $|V(C)| \leq 5$. However $|V(C)| \geq 5$ since case (1) does not apply. Thus C contains precisely 5 nodes, it is 2EC, minimal, and not a 5-cycle. The only possibilities are that C is a bowtie (i.e., two triangles that share a common node) or a $K_{2,3}$ (i.e a complete bipartite graph with a side of size 2 and the other of size 3). Depending on which case applies, we perform one of the following steps:
 - (3.a) C is a bowtie (see Figure 6.a). Let $\{v_1, v_2, u\}$ and $\{v_3, v_4, u\}$ be the two triangles in the bowtie. Assume that there exists some edge in $E(G) \setminus E(H')$ between the nodes of the bowtie, say v_1v_3 w.l.o.g. In this case we replace C with the cycle $v_1, v_3, v_4, u, v_2, v_1$. In other words, we replace H' with $H' \setminus E(C) \cup \{v_1v_3, v_3v_4, v_4u, uv_2, v_2v_1\}$. Notice that the number of edges of H' decreases. Otherwise, since u is not a 1-vertex-cut of G (as G is structured), there must exist an edge $zw \in E(G) \setminus E(H')$ with $z \in V(C) \setminus \{u\}$ and $w \in V(G) \setminus V(C)$. This is however excluded by case (2) since $H' \setminus \{uz\} \cup \{zw\}$ would be a 2-edge-cover with the same size as H' but with fewer components.
 - (3.b) C is a $K_{2,3}$ (see Figure 6.b). Let $\{v_1, v_2\}$ and $\{w_1, w_2, w_3\}$ be the two sides of C . Suppose that there exists an edge between the nodes $\{w_1, w_2, w_3\}$, say w_1w_2 . In this case we replace C with the 5-cycle $w_1, w_2, v_1, w_3, v_2, v_1$, namely we replace H' with $H' \setminus E(C) \cup \{w_1w_2, w_2v_1, v_1w_3, w_3v_2, v_2v_1\}$. Notice that the number of edges of H' decreases.

Suppose next that the nodes $\{w_1, w_2, w_3\}$ have degree precisely 2 in G . Then every feasible 2EC spanning subgraph contains all the edges of C , making C a $5/4$ -contractible subgraph on 5 nodes, which contradicts the assumption that G is structured.

The only remaining case is that there exists an edge $f \in E(G) \setminus E(H')$ between $\{w_1, w_2, w_3\}$ and $V(G) \setminus V(C)$, w.l.o.g. $f = w_1u$. This is however excluded by case (2) since $H' \setminus \{w_1v_1\} \cup \{w_1u\}$ would be a 2-edge-cover with fewer components.

- (4) There exists a block B of H' which contains at most 5 edges and such that all the edges of H' with exactly one endpoint in $V(B)$ are incident to the same node v_1 of B . We remark that this includes leaf blocks as a special case. Notice that by Case (1) B must be an ℓ -cycle v_1, \dots, v_ℓ, v_1 for $3 \leq \ell \leq 5$. Since G is structured, by the

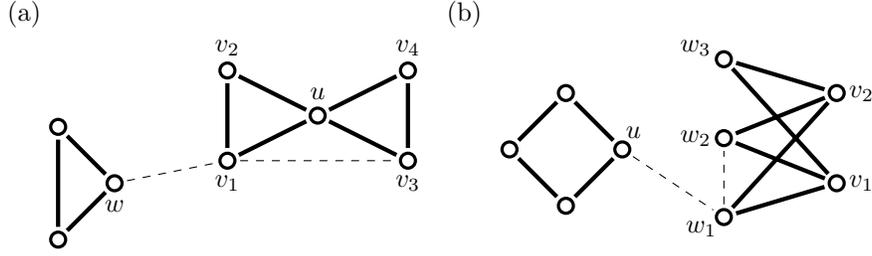


Figure 6: (a) Illustration of case (3.a) of Lemma 2.4. By adding v_1v_3 and removing $\{v_1u, v_3u\}$ we obtain a smaller 2-edge-cover. Alternatively, by adding v_1w and removing v_1u we obtain a 2-edge-cover of the same size with fewer components. (b) Illustration of case (3.b) of Lemma 2.4. By adding w_1w_2 and removing $\{v_2w_1, v_1w_2\}$ we obtain a smaller 2-edge-cover. Alternatively, by adding w_1u and removing w_1v_2 we obtain a 2-edge-cover of the same size with fewer components.

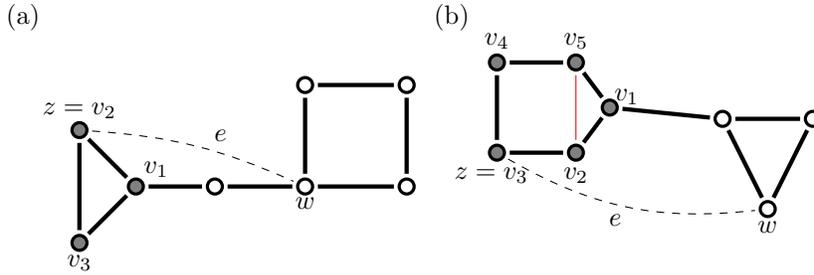


Figure 7: (a) Illustration of case (4.a) of Lemma 2.4. By adding v_2w and removing v_2v_1 we obtain a 2-edge-cover with fewer bridges. (b) Illustration of case (4.b) of Lemma 2.4. Nodes v_2 and v_5 are adjacent only to nodes in the 5-cycle $B = v_1, v_2, v_3, v_4, v_5, v_1$. Furthermore edge v_2v_5 must exist since otherwise B would be a 5/4-contractible subgraph. By adding $\{v_2v_5, v_3w\}$ and removing $\{v_2v_3, v_1v_5\}$ we obtain a 2-edge-cover with fewer bridges.

3-matching Lemma 2.3 there is a matching M of size 3 between $V(B)$ and $V \setminus V(B)$ in G . We distinguish two subcases:

(4.a) There exists an edge $zw \in E(G)$ with $z \in V(B)$ adjacent to v_1 in B , say $z = v_2$, and $w \notin V(B)$ (see Figure 7.a). In this case we replace H' with $H' \setminus \{v_1v_2\} \cup \{v_2w\}$. Notice that w must be in the same connected component of B by case (2). Furthermore, the number of bridges of H' decreases since some original bridge incident to v_1 is not a bridge any longer and no new bridge is created.

(4.b) v_2 and v_ℓ are adjacent only to nodes in $V(B)$ in G (see Figure 7.b). Since we have the matching M , then $\ell = 5$ in this case. Furthermore, we can assume w.l.o.g. that an edge $zw \in E(G)$, with $z \in \{v_3, v_4\}$ and $w \notin V(B)$, exists (otherwise v_1 would be a 1-vertex-cut). Assume by contradiction that $v_2v_5 \notin E(G)$. In this case any feasible solution would contain 4 distinct edges (incident to v_2 and v_5) with both endpoints in $V(B)$: this makes B a 5/4-contractible subgraph on 5 nodes, contradicting the assumption that G is structured. Thus we can assume $v_2v_5 \in E(G)$. In this case replace H' with $H' \setminus \{v_1v_5, v_2v_3\} \cup \{zw, v_2v_5\}$. As in Case (4.a), the number of connected components of H' decreases or the number of bridges of H' decreases and its number of connected components does not change.

In Steps (1) and (3) the number of edges decreases strictly, while in Steps (2) and (4) it remains unchanged. Furthermore in Steps (2) and (4) either the number of connected components decreases or the number of bridges decreases and the number of connected components does not change. It is then easy to check that the process ends within a polynomial number of steps (where each step can be executed in polynomial time).

At the end of the process clearly H' satisfies properties (i) (by case (3)) and (iii) (by case (2)). By case (4), leaf blocks contain at least 6 edges. The same holds for inner blocks B such that all the bridge edges incident to

B are incident to the same node $v \in V(B)$. Consider any remaining inner block B with 3 edges, in particular B is a triangle v_1, v_2, v_3, v_1 . By assumption there are at least two bridge edges incident to distinct nodes of B , say wv_1 and zv_2 . However in this case we would apply Step (1) and remove the edge v_1v_2 , a contradiction. \square

C Case of Many Triangles

In this section we prove Lemma 2.8, which we next restate together with the definition of merging cycle for the reader's convenience. We recall that S is the current infeasible solution, whose connected components are 2EC.

LEMMA 2.8. *Suppose that S is not core-triangle. Then in polynomial time we can compute $S' \subseteq E(G)$ such that: (1) the connected components of S' are 2EC, (2) S' has fewer connected components than S , and (3) $\text{cost}(S') \leq \text{cost}(S)$.*

DEFINITION 7. [*merging cycle*] *A merging cycle M for S is a collection of edges $M \subseteq E(G) \setminus S$ such that: (1) M induces one cycle (excluding loops) in the graph obtained by collapsing the connected components $\mathcal{C}(S)$ of S ; (2) if two edges of M are incident to a light component $C \in \mathcal{C}(S)$, they are incident to two distinct nodes of C .*

Recall that we build an auxiliary forest $G' = (\mathcal{C}(S), F)$ on the components $\mathcal{C}(S)$ of S , where initially $F = \emptyset$. The main tool to add edges to F is the following lemma, whose proof is postponed.

LEMMA C.1. *If G' consists of at least 2 trees, in polynomial time one can compute a merging cycle M incident to the components of at least 2 distinct trees of G' .*

If the merging cycle M obtained via Lemma C.1 is cheap, we derive the desired S' from S and M as described earlier.. Otherwise, we add to F the edge between the two 2EC components of S incident to M . We recall that M is incident to one light and one heavy component. Clearly the process ends in a polynomial number of steps by returning the desired cheap merging cycle or by obtaining a single tree G' .

In the latter case, we check whether there exists an edge e in G between two connected components of S which are *not* adjacent in G' . If such an e exists, we can find a cheap merging cycle via the following lemma.

LEMMA C.2. *Suppose that G' consists of a single tree and there exists an edge e_0 in G with endpoints in two distinct 2EC components of S which are not adjacent in G' . Then in polynomial time one can compute a cheap merging cycle of S .*

Proof. Assume that e_0 is incident to the components C_1 and C_k , and let $C_1, C_2, \dots, C_k, k \geq 3$, be the components along the (unique) path in G' between C_1 and C_k . We will define edges e_1, \dots, e_{k-1} such that each $e_i, i \geq 1$, has one endpoint in C_i and the other one in C_{i+1} , and such that $M := \{e_0, \dots, e_{k-1}\}$ is a merging cycle. The claim follows since $k \geq 3$ (hence M has to be cheap).

Remember that for each pair $C_i C_{i+1}$ there is an expensive merging cycle incident to exactly C_i and C_{i+1} (by the construction of G'). In particular there will be two edges a_i and b_i with one endpoint in C_i and the other in C_{i+1} . Furthermore, if C_i is light, the endpoints of a_i and b_i in C_i are distinct. By construction we also have that C_i is light and C_{i+1} is heavy or vice versa.

Given the edges $e_0, \dots, e_i, i \in \{0, \dots, k-2\}$, we choose e_{i+1} as follows. If C_{i+1} is light, we choose any edge in $\{a_{i+1}, b_{i+1}\}$ which is not incident to e_i (this edge must exist since a_{i+1} and b_{i+1} are incident to distinct nodes of C_{i+1}). Otherwise, i.e. if C_{i+1} is heavy, we set e_{i+1} to any edge in $\{a_{i+1}, b_{i+1}\}$ which is not incident to e_0 (the latter condition is always satisfied unless $i+1 = k-1$). \square

If there is no edge e_0 as in the previous lemma, then we apply the following lemma to compute a desired S' .

LEMMA C.3. *Suppose that G' consists of a single tree, all the edges e of G between distinct 2EC components C_1 and C_2 of S involve two adjacent components in G' , and S is not core-triangle. Then in polynomial time one can compute an $S' \subseteq E$ such that: (1) the connected components of S' are 2EC, (2) S' has fewer connected components than S , and (3) $\text{cost}(S') \leq \text{cost}(S)$.*

Proof. By assumption all the edges of G between different components of S involve adjacent components in G' . Furthermore, as remarked before, by construction if C_1 and C_2 are adjacent in G' , then exactly one of them is

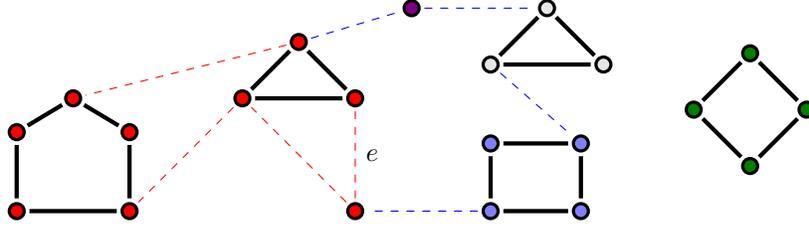


Figure 8: The black edges are the current solution S . The forest G' has 5 trees, which are identified by nodes of distinct colors. The red edges correspond to expensive merging cycles identified in previous steps of the construction. Let Π the partition of the node set induced by distinct colors. The blue dashed edges form a nice cycle N of Π . In particular if two edges in N touch a given set V_i in the partition, this happens at distinct nodes of V_i unless $|V_i| = 1$. A merging cycle M can be easily derived from N using the red edges. In particular, $M = N \cup \{e\}$ in this case, which is cheap since it involves at least 3 components of S .

light and the other is heavy. Notice that there must exist a light component $C = \{v_1, v_2, v_3\}$ of degree at least 2 in G' : indeed otherwise G' would be a star graph whose center is a heavy component and whose leaves are light components (hence in particular triangles). Since by assumption there are no edges between the leaf components of G' , this implies that S is core-triangle, contradicting the assumptions.

Suppose by contradiction that there exists a component C' adjacent to C in G' such that C' is adjacent only to two nodes in $V(C)$, say v_1 and v_2 . Then $\{v_1, v_2\}$ is a 2-vertex-cut and the edge v_1v_2 is irrelevant, contradicting the assumption that G is structured. Therefore we can choose any two distinct components C_1 and C_2 adjacent to C in G' , and two edges $\{a_1v_1, b_1v_2\}$ (resp., $\{a_2v_2, b_2v_3\}$) between C_1 (resp., C_2) and C . Consider $S' := (S \setminus \{v_1v_2, v_2v_3\}) \cup \{a_1v_1, b_1v_2, a_2v_2, b_2v_3\}$. Notice that S' satisfies (1) and (2), in particular it contains a 2EC component C' spanning the nodes $V(C) \cup V(C_1) \cup V(C_2)$. Furthermore S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + \text{cr}(S) - \text{cr}(S') \\ &= -2 + \text{cr}(C_1) + \text{cr}(C_2) + \text{cr}(C) - \text{cr}(C') = -2 + 2 + 2 + \frac{1}{2} - 2 > 0. \end{aligned}$$

□

Proof. [Proof of Lemma 2.8] The claim follows directly from the above construction and Lemmas C.1, C.2, and C.3. □

It remains to prove Lemma C.1. It is convenient to introduce the following definition (see Figure 8).

DEFINITION 9. (NICE CYCLE AND PATH) Let $\Pi = (V_1, \dots, V_k)$, $k \geq 2$, be a partition of the node-set of a graph G . A nice cycle (resp. nice path) N of Π is a subset of edges with endpoints in distinct subsets of Π such that: (1) N induces one cycle of length at least 2 (resp., one path of length at least 1) in the graph obtained from G by collapsing each V_i into a single node; (2) given any two edges of N incident to some V_i , these edges are incident to distinct nodes of V_i unless $|V_i| = 1$.

Notice that, considering the partition Π induced by the node sets of the 2EC components of S , a nice cycle for Π is a merging cycle for S . More important for our goals, if we consider the partition $\Pi(G')$ induced by the nodes of the components in each tree of G' (assuming that it contains at least 2 trees), then we can derive a merging cycle M for S from a nice cycle N for $\Pi(G')$ via the following lemma (see Figure 8).

LEMMA C.4. Given a nice cycle N of $\Pi(G') = \{V_1, \dots, V_k\}$, $k \geq 2$, in polynomial time one can compute a merging cycle M of S incident to the components of at least two distinct trees of G' .

Proof. Let U_1, \dots, U_q be the subsets in the partition incident to N . The set M consists of N plus a subset of edges M_j for each set U_j which is constructed as follows. Let e' and e'' be the two edges of N incident to U_j . Let C_1 and C_k be the components of S incident to e' and e'' , resp., (possibly $k = 1$) and let C_1, C_2, \dots, C_k the

set of components along the unique path between C_1 and C_k in G' . If $k = 1$, set $M_i = \emptyset$ (Notice that, if C_1 is light, the edges e' and e'' are incident to distinct nodes of C_1 by the definition of nice cycle). Otherwise M_i is constructed analogously to the construction of the edge set $\{e_1, \dots, e_{k-1}\}$ in the proof of Lemma C.2 with the following difference: In the definition of e_1 the role of e_0 is played by e' and in the definition of e_{k-1} the role of e_0 is replaced by e'' (notice that e_0 did not play any role in the definition of the remaining e_i 's). \square

A nice cycle can be computed efficiently via the following lemma.

LEMMA C.5. *Let $\Pi = (V_1, \dots, V_k)$, $k \geq 2$, be a partition of the node-set of a 2VC simple graph G . In polynomial time one can compute a nice cycle N of Π .*

Given the above claims, it is easy to conclude the proof of Lemma C.1.

Proof. [Proof of Lemma C.1] It is sufficient to compute a nice cycle N of $\Pi(G')$ via Lemma C.5, and then convert it into a merging cycle M via Lemma C.4. \square

It remains to prove Lemma C.5 (see Figure 9). We remark that the claim of this lemma sounds rather basic and natural, hence it is reasonable that it is already known in the literature (possibly with a simpler/more elegant proof).

Proof. [Proof of Lemma C.5] We use induction on $|\Pi|$ to obtain the desired cycle. The base case is $|\Pi| = 2$. Since G is 2VC, if $|V_1|, |V_2| > 1$, there is a matching N of size 2 between V_1 and V_2 , which induces a nice cycle. If one of V_1 and V_2 , say V_1 , has cardinality 1, then there are at least two edges e_1 and e_2 between V_1 and V_2 (with distinct endpoints in V_2 since G is simple). These edges form a nice cycle N .

Now assume $|\Pi| > 2$. We need the following (rather involved) definition of almost-nice cycle of a partition Π' :

DEFINITION 10. *Let Π' be a partition of V , $U_0, U_1, \dots, U_r \in \Pi'$ be distinct sets, $|U_0| \geq 2$, and $u_0 \in U_0$. Furthermore for every $1 \leq i \leq r$ there exist two nice-paths (not necessarily distinct) $P_{i,1}$ and $P_{i,2}$ from U_0 to U_i such that:*

- $P_{i,1}$ and $P_{i,2}$ are incident to u_0 and all their nodes belong to $U_0 \cup \dots \cup U_r$.
- If $|U_i| > 1$ then $P_{i,1}$ and $P_{i,2}$ are incident to different nodes of U_i .

In this case we say that $A = (U_0, u_0, (U_1, P_{1,1}, P_{1,2}), \dots, (U_r, P_{r,1}, P_{r,2}))$ is an almost-nice cycle of Π' .

We start by constructing an almost-nice cycle for Π involving at least two subsets of Π (or finding a nice cycle, in which case the claim holds). To this aim, we build a sequence V_0, V_1, \dots, V_k of distinct sets in Π , and a sequence of edges e_1, \dots, e_k such that e_i has one endpoint in V_{i-1} and the other in V_i . Furthermore, the two edges e_i and e_{i+1} , $1 \leq i < k$, are incident to distinct nodes of V_i if $|V_i| > 1$. We set V_0 to be any set in Π . Then we take as e_1 any edge with exactly one endpoint in V_0 and we let V_1 be the set containing the other endpoint of e_1 . Given V_0, \dots, V_k and e_1, \dots, e_k , we proceed as follows. We take any edge $e_{k+1} = vu$ with $v \in V_k$ and $u \notin V_k$, with the extra condition that v is not an endpoint of e_k if $|V_k| > 1$. Notice that such edge e_{k+1} must exist since the graph is 2VC. Now if $u \notin V_0 \cup \dots \cup V_{k-1}$, we expand the sequence by adding the edge e_{k+1} and the set $V_{k+1} \in \Pi$ containing u . Notice that this can happen at most $|\Pi| - 1$ times since $k + 1 \leq |\Pi|$. Otherwise, let $u \in V_j$ for some $0 \leq j < k$. If $|V_j| = 1$, e_{j+1}, \dots, e_{k+1} is a nice cycle and the claim holds. The same holds if $|V_j| > 1$ and u is not incident to e_{j+1} . The only remaining case is that u is incident to e_{j+1} . In this case we set $u_0 = u$ and $(U_0, \dots, U_r) = (V_j, \dots, V_k)$. For each U_{i-j} , the edges $\{e_{j+1}, \dots, e_k\}$ can be partitioned into two nice paths $P_{i-j,1}$ and $P_{i-j,2}$. These paths involve only nodes in $U_0 \cup \dots \cup U_r$, they have one endpoint in u_0 and the other in U_{i-j} . Furthermore these two paths have distinct endpoints in U_{i-j} if $|U_{i-j}| > 1$. Thus $A = (U_0, u_0, (U_1, P_{1,1}, P_{1,2}), \dots, (U_r, P_{r,1}, P_{r,2}))$ is an almost-nice cycle involving at least 2 sets in Π .

Given the current almost-nice cycle $A = (U_0, u_0, (U_1, P_{1,1}, P_{1,2}), \dots, (U_r, P_{r,1}, P_{r,2}))$, $r \geq 1$, we either find a nice cycle N , hence the claim holds, or we find a larger almost-nice cycle A' (incident to more subsets of Π). Clearly this process ends within a polynomial number of steps.

Given A , we proceed as follows. From Π we obtain a new partition Π' of V by replacing U_0, \dots, U_r with their union $U' = U_0 \cup \dots \cup U_r$. We observe that $|\Pi'| = k - r < k$. We have the following cases:

Proof. Let us initially assign $\frac{1}{3}$, $\frac{1}{4}$ and $\frac{3}{10}$ credits to each edge in H which belongs to a triangle 2EC component, which is a bridge, and to the remaining edges, resp. We show how to use these credits to assign the credits according to the assignment scheme from Section 2.5. The claim then follows since

$$\text{cost}(H) = |H| + \text{cr}(H) \leq |H| + \left(\frac{1}{3}t + \frac{1}{4}b + \frac{3}{10}(1-t-b)\right)|H|.$$

Each 2EC component C of H retains the credits of its edges. The number of such credits is $\frac{1}{3} \cdot 3 = 1$ if C is a triangle, and $\frac{3}{10}|E(C)|$ credits otherwise. This is sufficient to satisfy case (1) of the credit assignment scheme. Each bridge retains its credits: this satisfies case (3) of the credit assignment scheme.

Each 2EC block B retains 1 credit of the total amount of credits assigned to its edges, and the remaining credits of the mentioned edges are assigned to the connected component C containing B . Notice that $|E(B)| \geq 4$ since H is canonical, hence the edges of B have at least $\frac{3}{10} \cdot 4 > 1$ credits, which is sufficient to assign one credit to B : this satisfies case (4) of the credit assignment scheme.

Next consider any connected component C of H which is not 2EC. Notice that C contains at least 2 leaf blocks, namely blocks with exactly one edge of C (which is a bridge) incident to them. Let B_1 and B_2 be any two such blocks. Since H is canonical, each such B_i contains at least 6 edges. Therefore C collects from B_1 and B_2 at least $2 \cdot 6 \cdot \frac{3}{10} - 2 > 1$ credits: hence also case (2) of the credit assignment scheme is satisfied. \square

The proofs of the Bridge-Covering Lemma 2.10 and the Gluing Lemma 2.11 are given in the next two subsections.

D.1 Bridge-Covering In this section we prove Lemma 2.10, which we restate next.

LEMMA 2.10. [*Bridge-Covering*] Assume that S contains at least one bridge. Then in polynomial time one can compute a 2-edge-cover $S' \subseteq E(G)$ such that: (1) S' has fewer bridges than S ; (2) $\text{cost}(S') \leq \text{cost}(S)$.

Let C be any connected component of S containing at least one bridge (i.e., C is not 2EC). Let G_C be the multi-graph obtained from G by contracting into a single node each block B of C and each connected component C' of S other than C . Let T_C be the tree in G_C induced by the bridges of C : we call the nodes of T_C corresponding to blocks *block nodes*, and the remaining nodes of T_C *lonely nodes*. Observe that the leaves of T_C are necessarily block nodes (otherwise S would not be a 2-edge-cover).

at a high level, in this stage of our construction we will transform S into a new solution S' containing a component C' spanning the nodes of C (and possibly the nodes of some other components of S). Furthermore, no new bridge is created and at least one bridge e of C is not a bridge of C' (intuitively, the bridge e is covered). We remark that each component of the initial (canonical) 2-edge-cover H which is not 2EC contains at least 12 nodes, and we only possibly merge together components in this stage of the construction. As a consequence Invariant 1 is preserved (in particular, if C' happens to be 2EC, it contains at least 12 edges).

A *bridge-covering path* P_C is any path in $G_C \setminus E(T_C)$ with its (distinct) endpoints u and v in T_C , and the remaining (internal) nodes outside T_C . Notice that P_C might consist of a single edge, possibly parallel to some edge in $E(T_C)$. Augmenting S along P_C means adding the edges of P_C to S , hence obtaining a new 2-edge-cover S' . Notice that S' obviously has fewer bridges than S : in particular all the bridges of S along the u - v path in T_C are not bridges in S' (we also informally say that such bridges are *removed*), and the bridges of S' are a subset of the bridges of S . Let us analyze $\text{cost}(S')$. Suppose that the distance between u and v in T_C is br and such path contains bl blocks. Then the number of edges w.r.t. S grows by $|E(P_C)|$. The number of credits w.r.t. S decreases by at least $\frac{1}{4}br + bl + |E(P_C)| - 1$ since we remove br bridges, bl blocks and $|E(P_C)| - 1$ components (each one bringing at least one credit). However the number of credits also grows by 1 since we create a new block B' (which needs 1 credit) or a new 2EC component C' (which needs 1 additional credit w.r.t. the credit of C). Altogether $\text{cost}(S) - \text{cost}(S') \geq \frac{1}{4}br + bl - 2$. We say that P_C is cheap if the latter quantity is non-negative, and expensive otherwise. In particular P_C is cheap if it involves at least 2 block nodes or 1 block node and at least 4 bridges. Notice that a bridge-covering path, if at least one such path exists, can be computed in polynomial time.

Before proving Lemma 2.10 we need the following two technical lemmas (see Figures 10 and 11). We say that a node $v \in V(T_C) \setminus \{u\}$ is reachable from $u \in V(T_C)$ if there exists a bridge-covering path between v and u . Let $R(W)$ be the nodes in $V(T_C) \setminus W$ reachable from some node in $W \subseteq V(T_C)$, and let us use the shortcut $R(u) = R(\{u\})$ for $u \in V(T_C)$. Notice that $v \in R(u)$ iff $u \in R(v)$.

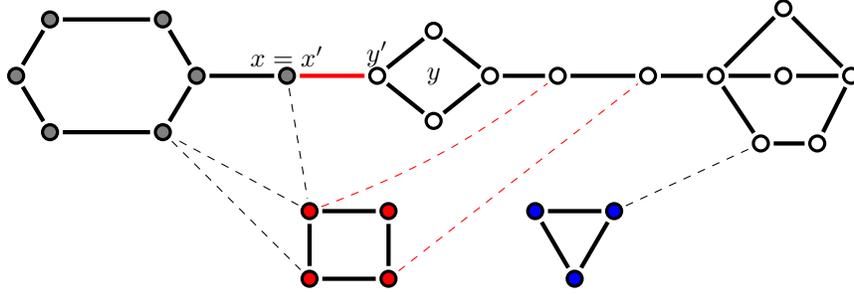


Figure 10: Illustration of Lemma D.1. The solid edges define S and the dashed ones are part of the remaining edges. The gray and white nodes induce the sets X_C and Y_C , resp. The set X is induced by the gray and red nodes. Notice that y corresponds to the block with 4 nodes (belonging to C). The red edges form the 3-matching used in the proof of the mentioned lemma.

LEMMA D.1. *Let $e = xy \in E(T_C)$ and let X_C and Y_C be the nodes of the two trees obtained from T_C after removing the edge e , where $x \in X_C$ and $y \in Y_C$. Then $R(X_C)$ contains a block node or $R(X_C) \setminus \{y\}$ contains at least 2 lonely nodes.*

Proof. Let us assume that $R(X_C)$ contains only lonely nodes, otherwise the claim holds. Let X be the nodes in G_C which are connected to X_C after removing Y_C , and let Y be the remaining nodes in G_C . Let X' and Y' be the nodes in G corresponding to X and Y , resp. Let e' be the edge in G that corresponds to e and let y' be the node of Y' that is incident to e' . In particular if y is a lonely node then $y' = y$ and otherwise y' belongs to the block of C corresponding to y .

Observe that both X_C and Y_C (hence X and Y) contain at least one (leaf) block node, hence both X' and Y' contain at least 6 nodes¹³. Therefore we can apply the 3-matching Lemma 2.3 to X' and obtain a matching $M = \{u'_1v'_1, u'_2v'_2, u'_3v'_3\}$ between X' and Y' , where $u'_i \in X'$ and $v'_i \in Y'$. Let u_i and v_i be the nodes in G_C corresponding to u'_i and v'_i , resp. We remark that $v_i \in Y_C$: indeed otherwise v_i would be connected to X_C in $G_C \setminus Y_C$, contradicting the definition of X . Let us show that the v_i 's are all distinct. Assume by contradiction that $v_i = v_j$ for $i \neq j$. Since $v'_i \neq v'_j$ (being M a matching) and v'_i, v'_j are both associated with v_i , this means that v_i is a block node in $R(X_C)$, a contradiction.

For each u_i there exists a path $P_{w_i u_i}$ in $G_C \setminus E(T_C)$ between u_i and some $w_i \in X_C$ (possibly $w_i = u_i$). Observe that $P_{w_i u_i} \circ u_i v_i$ is a bridge-covering path between $w_i \in X_C$ and $v_i \in Y_C$ unless $u_i v_i = xy$ (recall that the edges $E(T_C)$ cannot be used in a bridge-covering path). In particular, since the v_i 's are all distinct, at least two such paths are bridge-covering paths from X_C to distinct (lonely) nodes of $Y_C \setminus \{y\}$, implying $|R(X_C) \setminus \{y\}| \geq 2$. \square

Let $T_C(u, v)$ denote the path in T_C between nodes u and v .

LEMMA D.2. *Let b and b' be two leaf (block) nodes of T_C . Let $u \in R(b)$ and $u' \in R(b')$ be nodes of $V(T_C) \setminus \{b, b'\}$. Suppose that $T_C(b, u)$ and $T_C(b', u')$ both contain some node w (possibly $w = u = u'$) and $|E(T_C(b, u)) \cup E(T_C(b', u'))| \geq 4$. Then in polynomial time one can find a 2-edge-cover S' satisfying the conditions (1) and (2) of Lemma 2.10.*

Proof. Let P_{bu} (resp., $P_{b'u'}$) be a bridge-covering path between b and u (resp., b' and u'). Suppose that P_{bu} and $P_{b'u'}$ share an internal node. Then there exists a (cheap) bridge-covering path between b and b' and we can find the desired S' by the previous discussion. So next assume that P_{bu} and $P_{b'u'}$ are internally node disjoint. Let $S' = S \cup E(P_{bu}) \cup E(P_{b'u'})$. All the nodes and bridges induced by $E(P_{bu}) \cup E(P_{b'u'}) \cup E(T_C(b, u)) \cup E(T_C(b', u'))$ become part of the same block or 2EC component C' of S' . Furthermore C' contains at least 4 bridges of C and the two blocks B and B' corresponding to b and b' , resp. One has $|S'| = |S| + |E(P_{bu})| + |E(P_{b'u'})|$. Moreover $cr(S') \leq cr(S) + 1 - (|E(P_{bu})| - 1) - (|E(P_{b'u'})| - 1) - 2 - \frac{1}{4}$. In the latter inequality the +1 is due to the extra

¹³Leaf blocks initially contain at least 6 nodes being H canonical, and the bridge-covering stage does not create smaller leaf blocks.

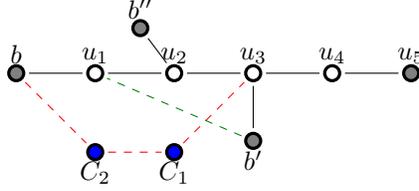


Figure 11: The solid edges induce T_C . The blue nodes correspond to contracted connected components other than C . The black and white nodes are block and lonely nodes of T_C , resp. The red edges induce an (expensive) bridge-covering path covering the bridges u_1u_2 , u_2u_3 and a bridge corresponding to bu_1 . The green edge induces another (expensive) bridge-covering path. An edge bb'' or bu_4 would induce a cheap bridge-covering path. Lemma D.2 can be applied with $u = u_3 \in R(b)$ and $u' = u_1 \in R(b')$ (both $T_C(b, u_3)$ and $T_C(b', u_1)$ contain $w = u_2$). In particular, one obtains S' by adding to S the red and green edges. Notice that the new component C' of S' which contains $V(C)$ has 4 fewer bridges and 1 fewer block than C .

credit required by C' , the -2 to the removed blocks B and B' (i.e., blocks of S which are not blocks of S'), and the $-\frac{1}{4}4$ to the removed bridges. Altogether $cost(S) - cost(S') \geq \frac{1}{4} \cdot 4 + 2 - 3 = 0$, hence (2) holds. Notice that S' satisfies (1) too. \square

We are now ready to prove Lemma 2.10.

Proof. [Proof of Lemma 2.10] If there exists a cheap bridge-covering path P_C (condition that we can check in polynomial time), we simply augment S along P_C hence obtaining the desired S' . Thus we next assume that no such path exists.

Let $P' = b, u_1, \dots, u_\ell$ in T_C be a longest path in T_C (interpreted as a sequence of nodes). Notice that b must be a leaf of T_C , hence a block node (corresponding to some leaf block B of C). Let us consider $R(b)$. Since by assumption there is no cheap bridge-covering path, $R(b)$ does not contain any block node. Hence $|R(b) \setminus \{u_1\}| \geq 2$ and $R(b)$ contains only lonely nodes by Lemma D.1 (applied to $xy = bu_1$).

We next distinguish a few subcases depending on $R(b)$. Let V_i , $i \geq 1$, be the nodes in $V(T_C) \setminus V(P')$ such that their path to b in T_C passes through u_i . Notice that $\{V(P'), V_1, \dots, V_\ell\}$ is a partition of $V(T_C)$. We observe that any node in V_i is at distance at most i from u_i in T_C (otherwise P' would not be a longest path in T_C). We also observe that as usual the leaves of T_C in V_i are block nodes.

REMARK 1. *This implies that all the nodes in V_1 are block nodes, and all the nodes in V_2 are block nodes or are lonely non-leaf nodes at distance 1 from u_2 in T_C .*

(1) There exists $u \in R(b)$ with $u \notin \{u_1, u_2, u_3\} \cup V_1 \cup V_2$. By definition there exists a bridge-covering path between b and u containing at least 4 bridges, hence cheap. This is excluded by the previous steps of the construction.

(2) There exists $u \in R(b)$ with $u \in V_1 \cup V_2$. Since u is not a block node, by Remark 1 u must be a lonely non-leaf node in V_2 at distance 1 from u_2 . Furthermore V_2 must contain at least one leaf block node b' adjacent to u . Consider $R(b')$. By the assumption that there are no cheap bridge-covering paths and Lemma D.1 (applied to $xy = b'u$), $|R(b') \setminus \{u\}| \geq 2$. In particular $R(b')$ contains at least one lonely node $u' \notin \{b', b, u\}$. The tuple (b, b', u, u') satisfies the conditions of Lemma D.2 (in particular both $T_C(b, u)$ and $T_C(b', u')$ contain $w = u_2$), hence we can obtain the desired S' .

(3) $R(b) \setminus \{u_1\} = \{u_2, u_3\}$. Recall that u_2 and u_3 are lonely nodes. We distinguish 2 subcases:

(3.a) $V_1 \cup V_2 \neq \emptyset$. Take any leaf (block) node $b' \in V_1 \cup V_2$, say $b' \in V_i$. Let ℓ' be the node adjacent to b' . By the assumption that there are no cheap bridge-covering paths and Lemma D.1 (applied to $xy = b'\ell'$), $R(b') \setminus \{\ell'\}$ has cardinality at least 2 and contains only lonely nodes. Choose any $u' \in R(b') \setminus \{\ell'\}$. Notice that $u' \notin V_1$ by Remark 1 (but it could be a lonely node in V_2 other than ℓ'). The tuple (b, b', u_3, u') satisfies the conditions of Lemma D.2 (in particular both $T_C(b, u_3)$ and $T_C(b', u')$ contain $w = u_2$), hence we can compute the desired S' .

(3.b) $V_1 \cup V_2 = \emptyset$. By Lemma D.1 (applied to $xy = u_1u_2$) the set $R(\{b, u_1\})$ contains a block node or $R(\{b, u_1\}) \setminus \{u_2\}$ contains at least 2 lonely nodes. Suppose first that $R(\{b, u_1\})$ contains a block node b' . Notice that $b' \notin R(b)$ by the assumption that there are no cheap bridge-covering paths, hence $u_1 \in R(b')$. Notice also that $b' \notin \{u_2, u_3\}$ since those are lonely nodes. Thus the tuple (b, b', u_2, u_1) satisfies the conditions of Lemma D.2 (in particular both $T_C(b, u_2)$ and $T_C(b', u_1)$ contain $w = u_2$), hence we can obtain the desired S' .

The remaining case is that $R(\{b, u_1\}) \setminus \{u_2\}$ contains at least 2 lonely nodes. Let us choose $u' \in R(\{b, u_1\}) \setminus \{u_2\}$ with $u' \neq u_3$. Let P_{bu_2} (resp., $P_{u'u_1}$) be a bridge-covering path between b and u_2 (resp., u' and u_1). Notice that P_{bu_2} and $P_{u'u_1}$ must be internally node disjoint otherwise $u' \in R(b)$, which is excluded since $R(b) \subseteq \{u_1, u_2, u_3\}$ by assumption. Consider $S' = S \cup E(P_{bu_2}) \cup E(P_{u'u_1}) \setminus \{u_1u_2\}$. Notice that S' satisfies (1). One has $|S'| = |S| + |E(P_{bu_2})| + |E(P_{u'u_1})| - 1$, where the -1 comes from the edge u_1u_2 . Furthermore $cr(S') \leq cr(S) + 1 - (|E(P_{bu_2})| - 1) - (|E(P_{u'u_1})| - 1) - \frac{1}{4}4 - 1$, where the $+1$ comes from the extra credit needed for the block or 2EC component C' containing $V(B)$, the final -1 from the removed block B , and the $-\frac{1}{4}4$ from the at least 4 bridges removed from S (namely a bridge corresponding to bu_1 , the edges u_1u_2 and u_2u_3 , and one more bridge incident to u'). Altogether $cost(S) - cost(S') \geq 0$, hence S' satisfies (2). \square

D.2 Gluing In this section we complete the description of the Gluing stage by proving Lemma 2.11, which we next restate.

LEMMA 2.11. [*Gluing*] Assume that the connected components of S are 2EC, and there are at least 2 such components. Then in polynomial time one can compute a 2-edge-cover $S' \subseteq E(G)$ such that: (1) the connected components of S' are 2EC; (2) S' has fewer connected components than S ; (3) $cost(S') \leq cost(S)$.

Recall that the component graph \hat{G}_S has one node per 2EC component of S and one edge C_1C_2 iff G contains at least one edge between the components C_1 and C_2 of S . In the gluing step associated with S we will check if \hat{G}_S satisfies certain properties, and based on that we derive a convenient S' . We start by dealing with certain *local configurations* involving 2 or 3 2EC components that can be merged together in Section D.2.1. Then we consider the case that \hat{G}_S is a tree in Section D.2.2. Finally we consider the case that \hat{G}_S is *not* a tree in Section D.2.3.

D.2.1 Local Merges If one of the following cases applies (which we can check in polynomial time), then we obtain the desired S' . We remark that in each case S' has the same 2EC components of S a part for a new 2EC component C' that spans the nodes of two or three 2EC components of S . In particular conditions (1) and (2) of Lemma 2.11 immediately hold. We also remark that C' contains at least 7 edges, hence Invariant 1 is preserved by a simple induction.

The next 3 lemmas can be used to merge two 2EC components of S into one.

LEMMA D.3. Suppose that there exists two components C_1 and C_2 of S where C_1 is a 4-cycle, C_2 is not a 5-cycle, and there exists a 3-matching M between C_1 and C_2 . Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.

Proof. Let $M = \{u_1v_1, u_2v_2, u_3v_3\}$ with $u_i \in C_1$ and $v_j \in C_2$.

Assume by contradiction that C_2 is a triangle or a 4-cycle. By a simple case analysis we can always choose 2 edges in M , say u_1v_1 and u_2v_2 , such that u_1 and u_2 (resp., v_1 and v_2) are adjacent in C_1 (resp., in C_2). Let $S' = S \cup \{u_1v_1, u_2v_2\} \setminus \{u_1u_2, v_1v_2\}$. Notice that S' is a 2-edge-cover of the same size as S but with fewer connected components than S . By Invariant 1, both C_1 and C_2 are 2EC components of the initial canonical 2-edge-cover H . Hence the existence of S' contradicts the 3-optimality of H .

Therefore C_2 is a 6-cycle or large. We remark that $cr(C_2) \geq \frac{3}{10}6$. In this case we choose any two edges in M , say u_1v_1 and u_2v_2 , such that u_1 and u_2 are adjacent in C_1 . Let $S' = S \setminus \{u_1u_2\} \cup \{u_1v_1, u_2v_2\}$. Notice that S' satisfies (1) and (2). In particular the nodes of C_1 and C_2 belong to a unique (large) 2EC component C' of S' . One has $|S'| = |S| + 1$ and

$$cr(S) - cr(S') = cr(C_1) + cr(C_2) - cr(C') \geq \frac{3}{10}4 + \frac{3}{10}6 - 2 = 1.$$

Hence $cost(S') \leq cost(S)$, i.e. S' satisfies (3). \square

LEMMA D.4. *Suppose that there exists two components C_1 and C_2 of S where C_1 is a triangle, C_2 is not a 6-cycle, and there exists a 3-matching M between C_1 and C_2 . Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let $M = \{u_1v_1, u_2v_2, u_3v_3\}$, where $u_i \in C_1$ and $v_j \in C_2$. We can exclude the case where C_2 is a 4-cycle since it is captured by Lemma D.3. If C_2 is a triangle or a 5-cycle, we can choose two edges in M , say u_1v_1 and u_2v_2 , such that v_1 and v_2 are adjacent in C_2 (u_1 and u_2 are obviously adjacent in C_1). Consider $S' = S \cup \{u_1v_1, u_2v_2\} \setminus \{u_1u_2, v_1v_2\}$. The same argument as in the proof of Lemma D.3 leads to a contradiction (of the 3-optimality of the initial H).

The remaining case is that C_2 is large. In this case consider $S' = S \setminus \{u_1u_2\} \cup \{u_1v_1, u_2v_2\}$. Notice that S' satisfies (1) and (2). Let C' be the (large) 2EC component of S' containing the nodes of C_1 and C_2 . One has $|S'| = |S| + 1$ and

$$cr(S) - cr(S') = cr(C_1) + cr(C_2) - cr(C') \geq 1 + 2 - 2 = 1.$$

Hence $cost(S') \leq cost(S)$, i.e. S' satisfies (3). \square

LEMMA D.5. *Suppose that there exists two components C_1 and C_2 of S where C_1 is a 5-cycle, C_2 is not a 4-cycle, and that there exists a 3-matching M between C_1 and C_2 . Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let $M = \{u_1v_1, u_2v_2, u_3v_3\}$, with $u_i \in C_1$ and $v_j \in C_2$. If C_2 is a triangle then we can apply Lemma D.4 and obtain the desired S' . Therefore we assume C_2 is 5-cycle, 6-cycle or large. Select any two edges in M , say u_1v_1 and u_2v_2 , such that u_1 and u_2 are adjacent in C_1 . Consider $S' = S \cup \{u_1v_1, u_2v_2\} \setminus \{u_1u_2\}$ and define C' analogously to the previous case. Also in this case S' satisfies (1) and (2). One has $|S'| = |S| + 1$ and

$$cr(S) - cr(S') = cr(C_1) + cr(C_2) - cr(C') \geq \frac{3}{10}5 + \frac{3}{10}5 - 2 = 1.$$

Hence $cost(S') \leq cost(S)$, i.e. S' satisfies (3). \square

LEMMA D.6. *Suppose that there exists two large components C_1 and C_2 and there exists a 3-matching M between C_1 and C_2 . Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. In this case take any two edges $e, f \in M$, and consider $S' = S \cup \{e, f\}$. Notice that S' satisfies (1) and (2). Let C' be the 2EC component of S' containing the nodes of C_1 and C_2 . One has $|S'| = |S| + 2$ and

$$cr(S) - cr(S') = cr(C_1) + cr(C_2) - cr(C') = 2 + 2 - 2 = 2.$$

Hence $cost(S') \leq cost(S)$, i.e. S' satisfies (3). \square

We next describe two cases when we can merge 3 2EC components of S into one.

LEMMA D.7. *Let C_1, C_2 and C_3 be 2EC components of S such that C_1 is a 4-cycle, C_2 and C_3 are 5-cycles, and there is a 3-matching M_{12} (resp., M_{13}) between C_1 and C_2 (resp., C_3). Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let us choose two edges $u_1v_1, u_2v_2 \in M_{12}$ with $u_i \in C_2$ and $v_j \in C_1$, such that u_1 and u_2 are adjacent in C_2 . Similarly, let us choose two edges $w_1v'_1, w_2v'_2 \in M_{13}$ with $w_i \in C_3$ and $v'_j \in C_1$, such that w_1 and w_2 are adjacent in C_3 . Consider $S' = S \setminus \{u_1u_2, w_1w_2\} \cup \{u_1v_1, u_2v_2, w_1v'_1, w_2v'_2\}$. Observe that S' satisfies (1) and (2). In particular it contains a 2EC component C' spanning the nodes $V(C_1) \cup V(C_2) \cup V(C_3)$. S' satisfies (3) since

$$\begin{aligned} cost(S) - cost(S') &= |S| - |S'| + cr(C_1) + cr(C_2) + cr(C_3) - cr(C') \\ &= -2 + \frac{3}{10}4 + \frac{3}{10}5 + \frac{3}{10}5 - 2 > 0 \end{aligned}$$

\square

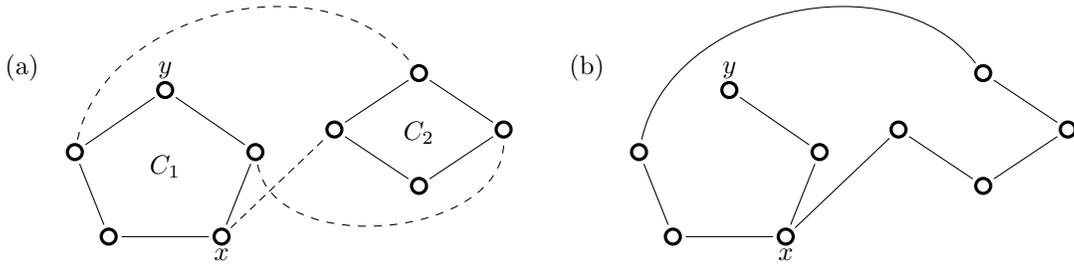


Figure 12: (a) Illustration of Lemma D.8, where the dashed edges correspond to the edges of M . (b) A valid subset F of 9 edges. Note that $F \cup \{xy\}$ forms a 2EC graph on $V(C_1) \cup V(C_2)$.

We need the following technical lemma (see Figure 12).

LEMMA D.8. *Let C_1 and C_2 be two 2EC components of S such that C_1 is a 5-cycle, C_2 is a 4-cycle, and there is a 3-matching M between C_1 and C_2 . Then for any two nodes $x, y \in V(C_1)$ there exists a set F of 9 edges in G such that $F \cup \{xy\}$ induces a 2EC spanning subgraph of $V(C_1) \cup V(C_2)$.*

Proof. Let $C_1 = v_1, v_2, v_3, v_4, v_5, v_1$ and $C_2 = u_1, u_2, u_3, u_4, u_1$. We consider the following cases, each time assuming that the previous cases do not apply:

(a) There exist $e_1, e_2 \in M$ which are incident to adjacent nodes of C_1 , say v', v'' , and adjacent nodes of C_2 , say u', u'' , the cycle $F = (E(C_1) \setminus \{v'v''\}) \cup (E(C_2) \setminus \{u'u''\}) \cup \{e_1, e_2\}$ satisfies the claim.

In the following we can assume w.l.o.g. that $M = \{u_1v_1, u_2v_3, u_4v_4\}$.

(b) x and y are adjacent nodes of C_1 . Consider $F = (E(C_1) \setminus \{xy\}) \cup (E(C_2) \setminus \{u_1u_2\}) \cup \{u_1v_1, u_2v_3\}$. Clearly $F \cup \{xy\}$ is 2EC and spans all the considered nodes.

In the following we can assume $\{x, y\} \neq \{v_1, v_3\}$, since $\{x, y\} = \{v_1, v_4\}$ is a symmetric case.

(c) $\{x, y\} \cap \{v_1, v_3\} = \emptyset$. The case $\{x, y\} = \{v_4, v_5\}$ is excluded by case (b). Thus, $\{x, y\} = \{v_2, v_4\}$ or $\{x, y\} = \{v_2, v_5\}$ may occur. In both cases we will not use the edge v_4u_4 in F . Under this restriction, the two mentioned cases are symmetric hence assume w.l.o.g. $\{x, y\} = \{v_2, v_5\}$. In this case $F = (E(C_1) \setminus \{v_1v_5\}) \cup (E(C_2) \setminus \{u_1u_2\}) \cup \{u_1v_1, u_2v_3\}$ satisfies the claim.

(d) $|\{x, y\} \cap \{v_1, v_3\}| = 1$. W.l.o.g. $x \in \{v_1, v_3\}$, therefore either $(x, y) = (v_1, v_4)$ or $(x, y) = (v_3, v_5)$ by case (b). In both cases we will not use the edge u_4v_4 in F . Under this restriction, the two mentioned cases are symmetric hence assume w.l.o.g. $(x, y) = (v_1, v_4)$. In this case $F = (E(C_1) \setminus \{v_3v_4\}) \cup (E(C_2) \setminus \{u_1u_2\}) \cup \{u_1v_1, u_2, v_3\}$ satisfies the claim. \square

LEMMA D.9. *Let C_1, C_2 and C_3 be 2EC components of S such that C_1 is a 5-cycle, C_2 and C_3 are 4-cycles, and there is a 3-matching M_{12} (resp., M_{13}) between C_1 and C_2 (resp., C_3). Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let $C_1 = v_1, v_2, v_3, v_4, v_5, v_1$ and $C_2 = z_1, z_2, z_3, z_4, z_1$. Notice that M_{12} must contain two edges with adjacent nodes in C_2 , say $v_i z_1, v_j z_2 \in M_{12}$. We apply Lemma D.8 to the tuple (C_1, C_3, v_i, v_j) , hence obtaining a set F of 9 edges such that $F \cup \{v_i v_j\}$ is a 2EC spanning subgraph of $V(C_1) \cup V(C_3)$. As a consequence $F' := F \cup E(C_2) \cup \{v_i z_1\} \cup \{v_j z_2\} \setminus \{z_1 z_2\}$ induces a 2EC component C' spanning the nodes $V(C_1) \cup V(C_2) \cup V(C_3)$. Define $S' = S \setminus (E(C_1) \cup E(C_2) \cup E(C_3)) \cup F'$. Notice that S' satisfies (1) and (2) trivially, and satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + \text{cr}(C_1) + \text{cr}(C_2) + \text{cr}(C_3) - \text{cr}(C') \\ &= 13 - 14 + \frac{3}{10}5 + \frac{3}{10}4 + \frac{3}{10}4 - 2 > 0 \end{aligned}$$

\square

D.2.2 When the Component Graph is a Tree Assuming that no local merge of the above kind can be performed, in this section we show how to obtain the desired S' when \hat{G}_S (which is a connected graph) is a tree. A critical observation here is that each edge $C_1C_2 \in E(\hat{G}_S)$ naturally induces a partition (V_1, V_2) of the node set (with both sides containing at least 3 nodes). The 3-matching Lemma 2.3 then guarantees that there is a 3-matching between V_1 and V_2 , hence between C_1 and C_2 by the definition of \hat{G}_S . The next lemma follows immediately:

LEMMA D.10. *Assume that the conditions of Lemmas D.3-D.9 do not hold and \hat{G}_S is a tree. Then, for any edge C_1C_2 and any two edges $C'_1C'_2, C'_1C'_3 \in E(\hat{G}_S)$:*

1. *If C_1 is a 4-cycle, then C_2 is a 5-cycle.*
2. *If C_1 is a 5-cycle then C_2 is a 4-cycle.*
3. *If C_1 is a triangle, then C_2 is a 6-cycle.*
4. *If C_1 is a large then C_2 is not large.*
5. *If C'_1 is a 4-cycle, then C'_2 and C'_3 cannot be both 5-cycles.*
6. *If C'_1 is a 5-cycle, then C'_2 and C'_3 cannot be both 4-cycles.*

LEMMA D.11. *Assume that the conditions of Lemmas D.3-D.9 do not hold and \hat{G}_S is a tree. Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Assume by contradiction that S contains a component C which is a 4-cycle or a 5-cycle. By Lemma D.10.1-2 all the neighbors of a 4-cycle (resp., 5-cycle) C' are 5-cycles (resp., 4-cycles). In particular, S must consist only of 4-cycles and 5-cycles. Notice that S must contain at least 3 components (otherwise G would contain at most 9 nodes), hence there exists a node in \hat{G}_S of degree at least 2. However a 5-cycle C_1 cannot have degree at least 2 in \hat{G}_S , otherwise C_1 would have 2 neighbouring 4-cycles C_2 and C_3 which is excluded by Lemma D.10.6. Similarly, a 4-cycle C_1 cannot have degree at least 2 by Lemma D.10.5. This leads to the desired contradiction.

Thus the components of S are triangles, 6-cycles or large. Assume by contradiction that there is *not* a 6-cycle in S . Then the existence of a triangle would contradict Lemma D.10.2. However in the complementary case all the components of S would be large, contradicting D.10.3. Thus we can assume that S contains a 6-cycle C . By the 3-matching Lemma 2.3 and the assumptions, for each neighbor C' of C in \hat{G}_S , there is a 3-matching between C and C' . We distinguish three subcases:

(a) For some triangle component C_1 of S adjacent to C in \hat{G}_S , there exists a 2-matching $M = \{u_1v_1, u_2v_2\}$ with $u_i \in C_1$ and $v_j \in C$ such that v_1 and v_2 are adjacent nodes of C . Consider $S' = S \setminus \{v_1v_2, u_1u_2\} \cup \{u_1v_1, u_2v_2\}$. The same argument as in the proof of Lemma D.3 leads to a contradiction, hence this case cannot happen.

(b) For some 6-cycle or large component C_1 of S adjacent to C in \hat{G}_S , there exists two edges u_1v_1, u_2v_2 with $u_i \in C_1$ and $v_j \in C$ such that v_1 and v_2 are adjacent nodes of C . Consider $S' = S \setminus \{v_1v_2\} \cup \{u_1v_1, u_2v_2\}$, which trivially satisfies (1) and (2). Let C' be the 2EC component in S' with node set $V(C) \cup V(C_1)$. Notice that S' satisfies (3) since

$$\text{cost}(S) - \text{cost}(S') = |S| - |S'| + cr(C) + cr(C_1) - cr(C') \geq -1 + \frac{3}{10}6 + \frac{3}{10}6 - 2 > 0.$$

(c) None of the above cases applies. Let $C = v_1, v_2, v_3, v_4, v_5, v_6, v_1$. Notice that for each neighbour C' of C in \hat{G}_S there is a 3-matching between C and C' . In particular C' is adjacent (in G) to at least 3 nodes in C . Furthermore the subset of nodes in $V(C)$ adjacent to C' are not adjacent in C , namely they are either $V_{\text{odd}} := \{v_1, v_3, v_5\}$ or $V_{\text{even}} := \{v_2, v_4, v_6\}$. This allows us to partition the neighbours of C in \hat{G}_S into the components \mathcal{C}_{odd} and $\mathcal{C}_{\text{even}}$ adjacent to V_{odd} and V_{even} , resp. We distinguish two subcases:

(c.1) One of the two sets \mathcal{C}_{odd} and $\mathcal{C}_{\text{even}}$ is empty, say $\mathcal{C}_{\text{odd}} = \emptyset$. Notice that any neighbour of $\{v_1, v_3, v_5\}$ in G must belong to C in this case. Assume that $v_1v_3 \in E(G)$. Consider any $C_1 \in \mathcal{C}_{\text{even}}$ (which cannot be empty). Let $u_2v_2, u_4v_4 \in E(G)$, with $u_2, u_4 \in V(C_1)$ and $u_2 \neq u_4$. If C_1 is *not* a triangle (hence it is a 6-cycle or large by

the previous cases), let $S' := S \cup \{u_2v_2, u_4v_4, v_1v_3\} \setminus \{v_1v_2, v_3v_4\}$. Clearly S' satisfies (1) and (2), and it satisfies (3) since

$$\text{cost}(S) - \text{cost}(S') \geq -1 + \frac{3}{10}6 + \frac{3}{10}6 - 2 > 0.$$

If C_1 is a triangle, consider $S' := S \cup \{u_2v_2, u_4v_4, v_1v_3\} \setminus \{v_1v_2, v_3v_4, u_2u_4\}$ (i.e., the same S' as before, but without the edge u_2u_4). The same argument as in the proof of Lemma D.3 leads to a contradiction, hence this case cannot happen.

A symmetric argument holds if $v_1v_5 \in E(G)$ or $v_3v_5 \in E(G)$, hence we next assume $v_1v_3, v_1v_5, v_3v_5 \notin E(G)$. This however implies that any feasible solution must include at least 6 edges with endpoints in $V(C)$, making C a contractible subgraph. This contradicts the assumption that G is structured.

(c.2) Let $C_{\text{odd}} \in \mathcal{C}_{\text{odd}}$ and $C_{\text{even}} \in \mathcal{C}_{\text{even}}$. Let $M_1 = \{u_1v_1, u_2v_3, u_3v_5\}$ be a 3-matching between C and C_{odd} , with $u_i \in C_{\text{odd}}$. Similarly, let $M_2 = \{w_1v_2, w_2v_4, w_3v_6\}$ be a 3-matching between C and C_{even} , with $w_i \in C_{\text{even}}$. We initially set $S' = S \setminus \{v_1v_2, v_3v_4\} \cup \{u_1v_1, u_3v_3, w_1v_2, w_2v_4\}$. Then, if C_{odd} (resp., C_{even}) is a triangle, we also remove u_1u_3 (resp., w_1w_2). In any case, S' satisfies (1) and (2). In particular it contains a 2EC component C' spanning the nodes set $V(C) \cup V(C_{\text{odd}}) \cup V(C_{\text{even}})$. If C_{odd} and C_{even} are both 6-cycles or large (in any combination), one has

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C) + cr(C_{\text{odd}}) + cr(C_{\text{even}}) - cr(C') \\ &\geq -2 + \frac{3}{10}6 + \frac{3}{10}6 + \frac{3}{10}6 - 2 > 0. \end{aligned}$$

If exactly one of the two components, say C_{even} , is a triangle, then

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C) + cr(C_{\text{odd}}) + cr(C_{\text{even}}) - cr(C') \\ &\geq -1 + \frac{3}{10}6 + \frac{3}{10}6 + 1 - 2 > 0. \end{aligned}$$

If both the components are triangles, one has

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C) + cr(C_{\text{odd}}) + cr(C_{\text{even}}) - cr(C') \\ &\geq \frac{3}{10}6 + 1 + 1 - 2 > 0. \end{aligned}$$

In all the cases S' satisfies (3). \square

D.2.3 When the Component Graph is Not a Tree It remains to consider the case where \hat{G}_S is not a tree. In the rest of this section we will prove the following lemma.

LEMMA D.12. *Assume that the conditions of Lemmas D.3-D.9 do not hold and \hat{G}_S is not a tree. Then in polynomial time one can compute an S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

We first need some more notation. Let $\mathcal{C}(S^*)$ be a subset of $\mathcal{C}(S)$ obtained as follows: iteratively remove from $\mathcal{C}(S)$ any component which has degree 1 in \hat{G}_S , and update \hat{G}_S consequently. Let S^* be the set of edges in $\mathcal{C}(S^*)$, and \hat{G}_{S^*} the subgraph of \hat{G}_S induced by S^* . Notice that by assumption \hat{G}_{S^*} is not empty and by construction each node in \hat{G}_{S^*} has degree at least 2. We also denote by $V^* \subseteq V$ the nodes set of $\mathcal{C}(S^*)$. Notice that each removed component C belongs to a subtree of \hat{G}_S rooted at some component C^* of S^* : we say that C is rooted at C^* .

We start by showing a variant of the 3-matching Lemma restricted to S^* .

LEMMA D.13. *Let $\mathcal{C}_1, \mathcal{C}_2 \neq \emptyset$ be a partition of $\mathcal{C}(S^*)$, and let V_i be the node set of \mathcal{C}_i . Then there is a 3-matching between V_1 and V_2 .*

Proof. Let \mathcal{C}'_i be the (removed) components in $\mathcal{C}(S) \setminus \mathcal{C}(S^*)$ which are rooted at some component in \mathcal{C}_i , and V'_i be the corresponding node set. The 3-matching Lemma 2.3 guarantees that there is a 3-matching M between $V_1 \cup V'_1$ and $V_2 \cup V'_2$. However by construction nodes in V'_1 (resp., V'_2) are not adjacent to nodes in $V_2 \cup V'_2$ (resp., $V_1 \cup V'_1$). Consequently the endpoints of M must belong to $V_1 \cup V_2$. The claim follows. \square

In order to prove Lemma D.12, we introduce the notion of *gluing path* of S^* which has some similarities with the notion of nice path (Definition 9 in Section C), however with some critical differences (see Figure 5).

DEFINITION 11. A *gluing path* P^* of S^* is a sequence of edges with endpoints in distinct components of S^* such that: (1) P^* induces a simple path in the graph obtained by contracting each connected component of S^* into a single node; (2) If e_1 and e_2 are two edges of P^* incident to a component $C^* \in \mathcal{C}(S^*)$ which is a triangle or a 4-cycle, then the two endpoints of these edges in $V(C^*)$ are adjacent in C^* .

The intuition behind this definition is as follows (see Figure 5). Suppose that we are given a gluing path $P_\ell = e_1, \dots, e_\ell$, with $e_i = out_{i-1}in_i$. Let C_{i-1} and C_i be the two components incident to e_i . Notice that C_i contains in_i for $2 \leq i \leq \ell$ and out_i for $1 \leq i \leq \ell - 1$. Suppose now that we are also given an edge $e' = v'u'$ with $v' \in C_\ell$ and $u' \in C_{\ell'}$ for some $\ell' < \ell$ (in Figure 5 one has $\ell' = 0$). Consider $S' = S \cup \{e_{\ell'+1}, \dots, e_\ell, e'\}$. Notice that S' contains a 2EC component C' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell)$. This construction can be refined as follows (which motivated the definition of gluing path). For each component C_j , $\ell' < j < \ell$, which is a triangle or a 4-cycle, remove from S' (and C') the edge in_jout_j . Notice that C' remains 2EC. Observe also that S' satisfying the conditions (1) and (2) of Lemma 2.11.

It remains to check condition (3) of Lemma 2.11, namely $cost(S') \leq cost(S)$. Our goal is to show that $cost(S) - cost(S') = |S| - |S'| + \sum_{j=\ell'}^{\ell} cr(C_j) - cr(C') \geq 0$. Excluding the -2 coming from e_ℓ and e' , and the -2 coming from $cr(C') = 2$, we can decompose the above difference as follows. We associate to each pair (C_j, e_{j+1}) , $\ell' < j < \ell$, the following contribution: -1 due to e_{j+1} , $+1$ due to in_jout_j if C_j is a triangle or 4-cycle, and $+cr(C_j)$. Hence the neat contribution of the pair is $+1, +\frac{12}{10}, +\frac{5}{10}, +\frac{8}{10}$ and $+1$ if C_j is a triangle, 4-cycle, 5-cycle, 6-cycle, or large, resp. In particular these contributions are always positive, which implies that $cost(S) \geq cost(S')$ trivially if ℓ' is sufficiently smaller than ℓ . In the remaining cases we need to be more careful in the analysis. In particular, we will need to carefully choose e' . Furthermore, sometimes we will need to remove one edge from $C_{\ell'}$ and/or C_ℓ . In one case we will need also to involve some component not in S^* . We remark that in the above construction the new 2EC component C' that we create always contains at least 7 nodes, hence Invariant 1 is maintained by a simple induction.

In more details, the construction proceeds as follows: we start with a gluing path P_ℓ of length $|P_\ell| = \ell \geq 1$. The base case $\ell = 1$ is obtained by taking any component $C_0 \in \mathcal{C}(S^*)$, and selecting any edge e_1 with one endpoint in C_0 and the other endpoint in some other component $C_1 \in \mathcal{C}(S^*)$ (notice that e_1 must exist since C_0 is not an isolated node in \hat{G}_S). Given such a P_ℓ , we either build a new gluing path $P_{\ell+1}$ of length $\ell + 1$ (which replaced P_ℓ in the next steps), or we build an S' satisfying the properties of Lemma D.12. Clearly we obtained a desired S' within a polynomial number of steps (since $\ell \leq |\mathcal{C}(S^*)| - 1$).

Consider the given gluing path P_ℓ with the above notation. We say that P_ℓ ends at C_ℓ . We sometimes will need to modify P_ℓ as in the next lemma.

LEMMA D.14. Given a gluing path $P_\ell = e_1, \dots, e_\ell$ ending at C_ℓ , in polynomial time one can compute a gluing path $P'_\ell = e_1, \dots, e'_\ell$ ending at C_ℓ and a 3-matching M_ℓ between $V(C_\ell)$ and $V^* \setminus V(C_\ell)$ with $e'_\ell \in M_\ell$.

Proof. By the 3-matching Lemma D.13, there exists a 3-matching M_ℓ between $V(C_\ell)$ and $V^* \setminus V(C_\ell)$. Suppose that M_ℓ is not incident to $out_{\ell-1}$. In this case we add e_ℓ to M_ℓ and remove some other edge from M_ℓ giving priority to the edge containing in_ℓ if any. Then $P'_\ell = P_\ell$ and (the modified) M_ℓ satisfy the claim. Otherwise, let e'_ℓ be the edge in M_ℓ incident to $out_{\ell-1}$. In this case $P'_\ell = e_1, \dots, e'_\ell$ and M_ℓ satisfy the claim. \square

In the next few lemmas we perform a case analysis based on the type of C_ℓ , namely if it is an i -cycle for some $3 \leq i \leq 6$ or it is large (i.e., if it contains at least 7 edges). We start with the case that C_ℓ is a 5-cycle, the most complicated case in our analysis. We need to distinguish between the subcases when C_ℓ has an adjacent component in $S \setminus S^*$ or not w.r.t. \hat{G}_S .

LEMMA D.15. Suppose that C_ℓ is a 5-cycle and it is adjacent to some component C'_ℓ in $S \setminus S^*$. Then in polynomial time one can either compute a gluing path $P_{\ell+1}$ of length $\ell + 1$ or a solution S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.

Proof. By construction $C'_\ell C_\ell$ is a bridge edge in \hat{G}_S . Hence by the 3-matching Lemma 2.3 there exists a 3-matching M'_ℓ between $V(C'_\ell)$ and $V(C_\ell)$. Since by assumption the conditions of Lemma D.5 do not hold, it must be the case that C'_ℓ is a 4-cycle.

Let us apply Lemma D.14 to P_ℓ , in order to obtain a pair (P'_ℓ, M_ℓ) as in the claim, and set P_ℓ to P'_ℓ . Suppose that there exists an edge $e_{\ell+1}$ with one endpoint in $V(C_\ell)$ and the other endpoint in $V^* \setminus (V(C_0) \cup \dots \cup V(C_\ell))$. Then $P_{\ell+1} = P_\ell, e_{\ell+1}$ satisfies the claim.

Otherwise, all the edges in M_ℓ must have one endpoint in $V(C_\ell)$ and the other one in $V(C_0) \cup \dots \cup V(C_{\ell-1})$. We distinguish two cases:

(a) All the edges in M_ℓ are between $V(C_\ell)$ and $V(C_{\ell-1})$. Since by assumption the conditions of Lemma D.5 do not hold, $C_{\ell-1}$ must be a 4-cycle. However in this case the triple $(C_\ell, C_{\ell-1}, C'_\ell)$ satisfies the conditions of Lemma D.9, a contradiction. Hence this case cannot happen.

(b) There exists an edge $e' = u'v' \in M_\ell$ with $u' \in V(C_\ell) \setminus \{in_\ell\}$ and $v' \in V(C_{\ell'})$, $\ell' \leq \ell - 2$. In this case we invoke Lemma D.8 on the pair (C_ℓ, C'_ℓ) to obtain a subset F of 9 edges such that $F \cup \{in_\ell u'\}$ is a 2EC spanning subgraph of $V(C_\ell) \cup V(C'_\ell)$. Let us initially set $S' = S \setminus (E(C_\ell) \cup E(C'_\ell)) \cup F \cup \{e_{\ell'}, \dots, e_\ell, e'\}$. We then remove from S' any edge $in_j out_j \in E(C_j)$, $\ell' + 1 \leq j \leq \ell - 1$, whenever C_j is a triangle or a 4-cycle (recall that in_j and out_j must be adjacent in C_j since P_ℓ is a gluing path). Let n' be the number of such removed edges and $n'' = \ell - \ell' - 1 - n'$.

Observe that S' satisfies (1) and (2). In particular, it contains a 2EC component C' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell) \cup V(C'_\ell)$. One has $|S'| = |S| + n'' + 2$. Furthermore

$$\begin{aligned} cr(S) - cr(S') &= cr(C'_\ell) + cr(C_\ell) + cr(C_{\ell'}) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq \frac{3}{10}4 + \frac{3}{10}5 + 1 + n' + \frac{3}{10}5n'' - 2. \end{aligned}$$

Altogether $cost(S) - cost(S') \geq \frac{1}{2}n'' + n' - \frac{3}{10} > 0$, hence S' satisfies (3). \square

LEMMA D.16. *Suppose that C_ℓ is a 5-cycle and it is not adjacent to any component in $S \setminus S^*$. Then in polynomial time one can either compute a gluing path $P_{\ell+1}$ of length $\ell + 1$ or a solution S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let us apply Lemma D.14 to P_ℓ , in order to obtain a pair (P'_ℓ, M_ℓ) as in the claim, and set P_ℓ to P'_ℓ . Suppose that there exists an edge $e' = u'v'$ with $u' \in V(C_\ell)$ and $v' \in V^* \setminus (V(C_0) \cup \dots \cup V(C_\ell))$. In this case $P_{\ell+1} := P_\ell, e'$ satisfies the claim.

Otherwise, let us prove the following technical claim, which exploits a simple case analysis and the fact that G is structured.

CLAIM 2. *There exists a 2-matching $M = \{wout_{\ell-1}, v'u'\}$, with $v', w \in V(C_\ell)$, between $V(C_\ell)$ and $V^* \setminus V(C_\ell)$ such that there exists a v' - w Hamiltonian path H_ℓ restricted to $V(C_\ell)$.*

Proof. Let $C_\ell = v_1, v_2, v_3, v_4, v_5, v_1$. Consider the 3-matching $M_\ell = \{e_\ell, e_3, e_4\}$. W.l.o.g. assume $e_\ell = v_1 out_{\ell-1}$. If one of the nodes v_2 or v_5 , say v_2 , is adjacent in G to some node $u_2 \in V^* \setminus (V(C_\ell) \cup \{out_{\ell-1}\})$, then $M = \{e_\ell, v_2 u_2\}$ satisfies the claim. We next assume that v_2 and v_5 are adjacent only to nodes in $V(C_\ell) \cup \{out_{\ell-1}\}$. Notice that by the previous case the endpoints of e_3 and e_4 are distinct from v_2 and v_5 . W.l.o.g. assume $e_3 = v_3 u_3$ and $e_4 = v_4 u_4$. If one of the nodes v_2 and v_5 , say v_2 , is adjacent to $out_{\ell-1}$, then $M = \{v_2 out_{\ell-1}, e_3\}$ satisfies the claim with $H_\ell = v_2, v_1, v_5, v_4, v_3$. If $v_2 v_5 \in E(G)$, then $M = \{e_\ell, e_3\}$ satisfies the claim with $H_\ell = v_1, v_2, v_5, v_4, v_3$. Otherwise every feasible solution must contain at least 4 distinct edges incident to v_2 or v_5 with the other endpoint in C_ℓ . Hence C_ℓ is a contractible subgraph of size 5, a contradiction. \square

We next update P_ℓ such that its final edge e_ℓ is the first edge in the 2-matching M (in particular, $w = in_\ell$) guaranteed by the above lemma. Notice that $|E(H_\ell)| = 4 = |E(C_\ell)| - 1$. We let $e' = v'u'$ be the remaining edge in M . W.l.o.g. assume $in_\ell = v_1$. By the previous case analysis we know that $u' \in V(C_{\ell'})$ for some $0 \leq \ell' < \ell$. We distinguish a few cases:

(a) $\ell' \leq \ell - 2$. Define $S' = S \setminus E(C_\ell) \cup H_\ell \cup \{e_{\ell'}, \dots, e_\ell, e'\}$. Remove from S' any edge $in_j out_j$ for $\ell' + 1 \leq j \leq \ell - 1$ such that C_j is a triangle or a 4-cycle. Observe that S' trivially satisfies (1) and (2). With the usual definition of

C' , n' , and n'' , S' also satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell'}) + cr(C_{\ell}) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -(n'' + 2 - 1) + 1 + \frac{3}{10}5 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - \frac{1}{2} \geq 0. \end{aligned}$$

(b) $\ell' = \ell - 1$ and $C_{\ell-1}$ is a 5-cycle, 6-cycle, or large. Define S' and C' as in case (a) (in particular S' satisfies (1) and (2)). S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell-1}) + cr(C_{\ell}) - cr(C') \\ &\geq -(2 - 1) + \frac{3}{10}5 + \frac{3}{10}5 - 2 = 0. \end{aligned}$$

(c) $\ell' = \ell - 1$ and $C_{\ell-1}$ is a triangle. In this case consider $S' = S \setminus E(C_{\ell}) \cup H_{\ell} \setminus \{u'out_{\ell-1}\} \cup \{e_{\ell}, e'\}$, which satisfies (1) and (2). With the usual definition of C' , one has that S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell-1}) + cr(C_{\ell}) - cr(C') \\ &\geq -(2 - 2) + 1 + \frac{3}{10}5 - 2 > 0. \end{aligned}$$

(d) $\ell' = \ell - 1$ and $C_{\ell-1}$ is a 4-cycle. By the definition of S^* , there must exist an edge $e'' = v''u''$ with $v'' \in V(C_{\ell})$ and $u'' \in V(C_{\ell'})$ for some $\ell'' \leq \ell - 2$.

(d.1) $out_{\ell-1}$ and u' are adjacent in $C_{\ell-1}$. Then essentially the same construction and analysis as in case (c) gives the desired S' . In particular one has

$$\text{cost}(S) - \text{cost}(S') \geq -(2 - 2) + \frac{3}{10}4 + \frac{3}{10}5 - 2 > 0.$$

We next assume w.l.o.g. that $C_{\ell-1} = u_1, u_2, u_3, u_4, u_1$ where $u_1 = out_{\ell-1}$, $u_3 = u'$, and $\ell - 1 = 0$ or $in_{\ell-1} = u_2$.

(d.2) $v' = v_3$ (the case $v' = v_4$ is symmetric). If $v'' \in \{v_2, v_3, v_5\}$, we can exploit the analysis of case (a) where the role of e' is replaced by e'' . This is because there exists a Hamiltonian $v''-v_1$ path over $V(C_{\ell})$. Otherwise, consider the alternative gluing path $P'_{\ell} = e_1, \dots, e_{\ell-1}, e'$ of length ℓ . If $v'' \in \{v_1, v_4\}$, thanks to the existence of edge e'' , we can apply to P'_{ℓ} the analysis of case (a).

(d.3) $v' = v_2$ (the case $v' = v_5$ is symmetric). If $v'' \in \{v_3, v_5\}$, we can exploit the analysis of case (a) where the role of e' is replaced by e'' . If $v'' \in \{v_1, v_3\}$, consider the alternative gluing path $P'_{\ell} = e_1, \dots, e_{\ell-1}, e'$ of length ℓ . Thanks to the existence of edge e'' , we can apply to P'_{ℓ} the analysis of case (a).

The remaining case is $v'' = v_4$. If $v_3v_5 \in E(G)$, then there exists a v_1-v_4 Hamiltonian path over $V(C_{\ell})$, namely v_1, v_2, v_3, v_5, v_4 . Hence again case (a) applies with e' replaced by e'' . Otherwise, assume by contradiction that v_3 and v_5 are not adjacent to any node outside $V(C_{\ell})$. This would make $V(C_{\ell})$ a contractible subgraph on 5 nodes (at least 4 edges with endpoints in $V(C_{\ell})$ belong to every feasible solution), which is a contradiction since G is structured. Thus there must exist an edge $e''' = v'''u'''$ with $v''' \in \{v_3, v_5\}$ and $u''' \notin V(C_{\ell})$. If $u''' \notin V(C_{\ell-1})$, again we can reduce to case (a) applied to P_{ℓ} or P'_{ℓ} with edge e''' playing the role of e' . If $u''' \in \{u_2, u_4\}$, we can exploit the analysis of case (d.1) applied to P_{ℓ} (if $v''' = v_5$) or to P'_{ℓ} (if $v''' = v_3$). The remaining case is $u''' \in \{u_1, u_3\}$. Consider $P''_{\ell} = e_1, \dots, e_{\ell-1}, e'''$. We can apply to P''_{ℓ} the analysis of case (a) where the role of e' is played by e'' (there is a $v'''-v_4$ Hamiltonian path over $V(C_{\ell})$). \square

In the next three lemmas we address the case when C_{ℓ} is a 4-cycle, a 6-cycle or large, or a triangle, resp.

LEMMA D.17. *Suppose that C_{ℓ} is a 4-cycle. Then in polynomial time one can either compute a gluing path $P_{\ell+1}$ of length $\ell + 1$ or a solution S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. We distinguish a few subcases (each time excluding the previous cases):

(a) There exists an edge $e' = v'u'$ with $v' \in V(C_\ell)$ adjacent to in_ℓ in C_ℓ and $u' \in V^* \setminus (V(C_0) \cup \dots \cup V(C_\ell))$. Then $P_{\ell+1} := P_\ell, e_{\ell+1}$ satisfies the claim.

(b) There exists an edge $e' = u'v' \in E(G)$ with $u' \in V(C_\ell) \setminus \{in_\ell\}$ adjacent to in_ℓ and $v' \in V(C_{\ell'})$, $\ell' \leq \ell - 2$. Let $S' = S \cup \{e_{\ell'}, \dots, e_\ell, e'\} \setminus \{in_{\ell'}u'\}$. Next remove from S' any edge $in_j out_j$ with $\ell' + 1 \leq j \leq \ell - 1$ where C_j is a triangle or a 4-cycle. Let n' be the number of removed edges and $n'' = \ell - \ell' - 1 - n'$. Observe that S' satisfies conditions (1) and (2). In particular, S' has a 2EC component C' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell)$. Notice also that $|S'| = |S| + n'' + 1$. We further distinguish a few subcases (each time excluding the previous subcases):

(b.1) $C_{\ell-1}$ is not a 5-cycle. If $C_{\ell-1}$ is a triangle or 4-cycle (which implies $n' \geq 1$), one has

$$\begin{aligned} cost(S) - cost(S') &= -n'' - 1 + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} ct(C_j) - cr(C') \\ &\geq -n'' - 1 + 1 + \frac{3}{10}4 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - \frac{8}{10} > 0. \end{aligned}$$

Otherwise $C_{\ell-1}$ is a 6-cycle or large, hence

$$\begin{aligned} cost(S) - cost(S') &= -n'' - 1 + cr(C_{\ell-1}) + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-2} cr(C_j) - cr(C') \\ &\geq -n'' - 1 + \frac{3}{10}6 + 1 + \frac{3}{10}4 + n' + \frac{3}{10}5(n'' - 1) - 2 = \frac{1}{2}n'' + n' - \frac{1}{2} \geq 0. \end{aligned}$$

In both cases (3) holds.

(b.2) $\ell' \leq \ell - 3$. S' satisfies (3) since

$$\begin{aligned} cost(S) - cost(S') &= -n'' - 1 + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -n'' - 1 + 1 + \frac{3}{10}4 + n' + \frac{3}{10}5n'' - 2 \geq \frac{1}{2}n'' + n' - \frac{8}{10} > 0, \end{aligned}$$

where in the last inequality we used $n' + n'' \geq 2$. We next assume $\ell' = \ell - 2$.

(b.3) $C_{\ell-1}$ is a 5-cycle and v' is adjacent to $out_{\ell'}$. In this case modify S' (and C') by removing $v'out_{\ell'}$. Notice that S' still satisfies (1) and (2). Furthermore it satisfies (3) since

$$\begin{aligned} cost(S) - cost(S') &= -n'' + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} ct(C_j) - cr(C') \\ &\geq -n'' + 1 + \frac{3}{10}4 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' + \frac{1}{5} > 0. \end{aligned}$$

(b.4) $C_{\ell-1}$ is a 5-cycle and v' is not adjacent to $out_{\ell'}$. Consider $P'_\ell = e_1, \dots, e_{\ell'}, e', e_\ell$. Notice that this is a gluing path of length ℓ . This is trivial if $C_{\ell'}$ is not a triangle nor a 4-cycle. Otherwise, by Case (b.3), $v' = out_{\ell'}$ if $C_{\ell'}$ is a triangle and $v' = out_{\ell'}$ or v' is the node opposite to $out_{\ell'}$ in $C_{\ell'}$ if $C_{\ell'}$ is a 4-cycle. In both cases v' is adjacent to $in_{\ell'}$ unless the latter node is not defined (i.e., if $\ell' = 0$). Hence P'_ℓ satisfies all the desired properties. Since P'_ℓ ends with a 5-cycle by construction, we can apply Lemma D.15 or D.16 to infer the claim.

In the following cases we assume $\ell' = \ell - 1$.

(c) There exists an edge $e' = u'v'$ with $u' \in V(C_\ell)$ adjacent to in_ℓ in C_ℓ and $v' \in V(C_{\ell-1})$ adjacent to $out_{\ell-1}$ in $C_{\ell-1}$. Consider $S'' = S \setminus \{in_\ell u', out_{\ell-1} v'\} \cup \{e_\ell, e'\}$. Notice that S'' satisfies (1) and (2). In particular S'' has a 2EC component C' spanning the nodes $V(C_\ell) \cup V(C_{\ell-1})$. S'' also satisfies (3) since

$$cost(S) - cost(S'') = cr(C_\ell) + cr(C_{\ell-1}) - cr(C') \geq \frac{3}{10}4 + 1 - 2 > 0$$

(d) All the edges $e' = u'v'$, with $u' \in V(C_\ell) \setminus \{in_\ell\}$ adjacent to in_ℓ in C_ℓ and $v' \notin V(C_\ell)$, have $v' \in V(C_{\ell-1})$ not adjacent to $out_{\ell-1}$ in $C_{\ell-1}$. Let $e' = u'v'$ be one such edge. Notice that such an edge must exist since there is a 3-matching between $V(C_\ell)$ and $V^* \setminus V(C_\ell)$. Let $C_\ell = v_1, v_2, v_3, v_4, v_1$ with $v_1 = in_\ell$. By the definition of S^* , one among v_1 and v_3 must be adjacent to some node in $V^* \setminus (V(C_\ell) \cup V(C_{\ell-1}))$. Consider $P'_\ell = e_1, \dots, e_{\ell-1}, e'$. This happens to be a gluing path of length ℓ . In particular, if $C_{\ell-1}$ is a triangle or a 4-cycle, then v' is adjacent to $in_{\ell-1}$ in $C_{\ell-1}$ or $\ell - 1 = 0$. Observe that u' is adjacent to both v_1 and v_3 , and at least one of them is incident to an edge with the other endpoint in $V^* \setminus (V(C_\ell) \cup V(C_{\ell-1}))$. The claim then follows by applying to P'_ℓ the construction of cases (a) or (b). \square

LEMMA D.18. *Suppose that C_ℓ is a 6-cycle or large. Then in polynomial time one can either compute a gluing path $P_{\ell+1}$ of length $\ell + 1$ or a solution S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Suppose there exists an edge $e' = u'v'$ with $u' \in V(C_\ell)$ and $v' \in V^* \setminus (V(C_0) \cup \dots \cup V(C_\ell))$. Then $P_{\ell+1} := P_\ell, e_{\ell+1}$ satisfies the claim. We next assume that such e' does not exist.

By the definition of S^* there exists an edge $e' = u'v' \in E(G)$ with $u' \in V(C_\ell)$ and $v' \in V(C_{\ell'})$ with $\ell' \leq \ell - 2$. In case of ties, we choose e' so that ℓ' is as small as possible, and as a second choice so that v' is adjacent to $out_{\ell'}$ in $C_{\ell'}$ if possible.

We define S' as follows. Set initially $S' = S \cup \{e_{\ell'+1}, \dots, e_\ell, e'\}$. Then remove from S' any edge $in_j out_j$ with $\ell' + 1 \leq j \leq \ell - 1$ such that C_j is a triangle or 4-cycle. Let n' be the number of such removed edges and $n'' = \ell - \ell' - 1 + 1 - n'$. Observe that S' satisfies (1) and (2). In particular, there is a component C' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell)$. Notice also that $|S'| = |S| + n'' + 2$ and $cr(C') = 2$.

We distinguish a few subcases, each time assuming that the previous cases do not apply:

(a) $\ell' \leq \ell - 3$. One has

$$\begin{aligned} cost(S) - cost(S') &= |S| - |S'| + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -(n'' + 2) + cr(C_{\ell'}) + \frac{3}{10}6 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' + cr(C_{\ell'}) - \frac{22}{10}. \end{aligned}$$

Notice that $n' + n'' \geq 2$. If $C_{\ell'}$ is not a triangle, then $\frac{1}{2}n'' + n' + cr(C_{\ell'}) \geq 1 + \frac{3}{10}4 = \frac{22}{10}$ and (3) holds. Similarly, if $n' \geq 1$ or $n'' \geq 3$ since in that case $\frac{1}{2}n'' + n' + cr(C_{\ell'}) \geq \frac{3}{2} + 1 > \frac{22}{10}$. Therefore we can assume that $C_{\ell'}$ is a triangle, $n' = 0$ and $n'' = 2$. Notice that $\ell' = \ell - 3$ in this case. Observe that $C_{\ell-1}$ and $C_{\ell-2}$ are not triangles nor 4-cycles. If at least one of them is a 6-cycle or large, we can refine the above inequality to

$$cost(S) - cost(S') \geq -(n'' + 2) + 1 + \frac{3}{10}6 + n' + \frac{3}{10}5(n'' - 1) + \frac{3}{10}6 - 2 = \frac{1}{2}n'' + n' - \frac{9}{10} > 0,$$

where we used $n'' = 2$. Thus we can assume that both $C_{\ell-1}$ and $C_{\ell-2}$ are 5-cycles.

Suppose now that $v' \neq out_{\ell'}$. In this case consider the same S' as before, but removing the edge $v'out_{\ell'}$ (this leads to a modified C' which is still 2EC, in particular (1) and (2) still hold). With a similar notation we obtain

$$cost(S) - cost(S') \geq -(n'' + 2 - 1) + 1 + \frac{3}{10}6 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - \frac{2}{10} > 0.$$

In the remaining case $v' = out_{\ell'}$, consider the alternative gluing path $P'_\ell = e_1, \dots, e_{\ell-3}, e', e_\ell, e_{\ell-1}$ of length ℓ . This path ends at $C_{\ell-2}$, which is a 5-cycle. The claim follows by applying to P'_ℓ one of the Lemmas D.15 and D.16. In the following cases we will assume $\ell' = \ell - 2$, hence in particular $v' \in C_{\ell-2}$.

(b) v' is adjacent to $out_{\ell-2}$. Modify S' (and C') by removing the edge $v'out_{\ell-2}$. Notice that C' remains 2EC, in particular (1) and (2) still hold. If $C_{\ell-dit1}$ is a triangle or a 4-cycle, one has

$$\begin{aligned} cost(S) - cost(S') &= |S| - |S'| + cr(C_{\ell-2}) + cr(C_{\ell-1}) + cr(C_\ell) - cr(C') \\ &\geq -1 + 1 + 1 + \frac{3}{10}6 - 2 > 0. \end{aligned}$$

Otherwise

$$\text{cost}(S) - \text{cost}(S') \geq -2 + 1 + \frac{3}{10}5 + \frac{3}{10}6 - 2 > 0.$$

In both cases (3) holds.

(c) $C_{\ell-2}$ is a 5-cycle, a 6-cycle or large.

(c.1) $C_{\ell-1}$ is a 4-cycle or a 5-cycle. Consider the alternative gluing path $P'_\ell = e_1, \dots, e_{\ell-2}, e', e_\ell$ of length ℓ . Notice that the last component of P'_ℓ is $C_{\ell-1}$, which is a 4-cycle or a 5-cycle. The claim follows by applying to P'_ℓ one of the Lemmas D.15, D.16, and D.17.

(c.2) $C_{\ell-1}$ is a triangle. S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell-2}) + cr(C_{\ell-1}) + cr(C_\ell) - cr(C') \\ &\geq -2 + \frac{3}{10}5 + 1 + \frac{3}{10}6 - 2 > 0. \end{aligned}$$

(c.3) $C_{\ell-1}$ is a 6-cycle, or large. S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell-2}) + cr(C_{\ell-1}) + cr(C_\ell) - cr(C') \\ &\geq -3 + \frac{3}{10}5 + \frac{3}{10}6 + \frac{3}{10}6 - 2 > 0. \end{aligned}$$

(d) $C_{\ell-2}$ is a triangle or a 4-cycle. Notice that by case (2), v' is not adjacent to $out_{\ell-2}$, hence it must be adjacent to $in_{\ell-2}$ or $\ell - 2 = 0$.

(d.1) $C_{\ell-1}$ is a 4-cycle or a 5-cycle. The claim follows analogously to case (c.1) by considering the same P'_ℓ and applying Lemma D.15, D.16, or D.17.

(d.2) $C_{\ell-1}$ is a triangle. Suppose first that $C_{\ell-2}$ is a 4-cycle. Then S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell-2}) + cr(C_{\ell-1}) + cr(C_\ell) - cr(C') \\ &\geq -2 + \frac{3}{10}4 + 1 + \frac{3}{10}6 - 2 = 0. \end{aligned}$$

Otherwise (i.e., $C_{\ell-2}$ is a triangle), by Lemma D.13, there exists a 3-matching M between $V(C_\ell) \cup V(C_{\ell-1})$ and $V^* \setminus (V(C_\ell) \cup V(C_{\ell-1}))$. By the previous cases M does not contain an edge $e'' = u''v''$ with $u'' \in V(C_\ell)$ and $v'' \in V(C_{\ell-2})$ adjacent to $out_{\ell-2}$. Hence M must contain an edge $e'' = u''v''$ with $u'' \in V(C_{\ell-1}) \setminus \{out_{\ell-1}\}$ and $v'' \neq out_{\ell-2}$. If $v'' \notin V(C_0) \cup \dots \cup V(C_{\ell-2})$, then $P_{\ell+1} = (e_1, \dots, e_{\ell-2}, e', e_\ell, e'')$ satisfies the claim. If $v'' \in V(C_{\ell-2})$, then consider $S'' = S \cup \{e_\ell, e', e''\} \setminus \{v''out_{\ell-2}, u''out_{\ell-1}\}$ and let C'' be the 2EC component spanning the nodes $V(C_{\ell-2}) \cup V(C_{\ell-1}) \cup V(C_\ell)$. Clearly S'' satisfies (1) and (2), and it satisfies (3) since:

$$\begin{aligned} \text{cost}(S) - \text{cost}(S'') &= |S| - |S''| + cr(C_{\ell-2}) + cr(C_{\ell-1}) + cr(C_\ell) - cr(C'') \\ &= -1 + 1 + 1 + \frac{3}{10}6 - 2 > 0. \end{aligned}$$

Otherwise $v'' \in V(C_{\ell'})$ for some $\ell' < \ell - 2$. In this case let $S''' = S \cup \{e_{\ell'+1}, \dots, e_{\ell-1}, e', e_\ell, e''\} \setminus \{u''out_{\ell-1}\}$. Remove from S''' any edge $in_j out_j$ with $\ell' + 1 \leq j \leq \ell - 2$ where C_j is a triangle or a 4-cycle, and let n' be the number of such removed edges. Notice that S''' satisfies (1) and (2). Observe that $n' \geq 1$ due to $C_{\ell-2}$. Let $n'' = \ell - 2 - \ell' - n'$. Then one has $|S'''| = |S| + n'' + 2$. Observe that S''' contains a 2EC component C''' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell)$. Then S' satisfies (3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S''') &= |S| - |S'''| + cr(C_{\ell'}) + cr(C_{\ell-1}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-2} cr(C_j) - cr(C''') \\ &\geq -(n'' + 2) + 1 + 1 + \frac{3}{10}6 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - \frac{2}{10} > 0, \end{aligned}$$

where in the last inequality we used $n' \geq 1$.

(d.3) $C_{\ell-1}$ is a 6-cycle or large. By Lemma D.13, there exists a 3-matching M between $V(C_\ell) \cup V(C_{\ell-1})$ and $V^* \setminus (V(C_\ell) \cup V(C_{\ell-1}))$. Since by the previous cases $V(C_\ell)$ can be adjacent only to $V(C_{\ell-1}) \cup \{out_{\ell-2}\}$, there must exist one edge $e'' = u''v'' \in E(G)$ with $u'' \in V(C_{\ell-1})$ and $v'' \neq out_{\ell-2}$. If there exists one such e'' with $v'' \notin V(C_0) \cup \dots \cup V(C_{\ell-2})$, then $P_{\ell+1} := (e_1, \dots, e_{\ell-2}, e', e_\ell, e'')$ satisfies the claim.

Otherwise, consider $P'_\ell := (e_1, \dots, e_{\ell-2}, e', e_\ell)$. Notice that P'_ℓ is a gluing path of length ℓ ending at $C_{\ell-1}$. In particular, the endpoint v' of e' in $V(C_{\ell-2})$ is not adjacent to $out_{\ell-2}$, hence it must be adjacent to $in_{\ell-2}$. If there exists one such e'' with $v'' \in V(C_{\ell-2})$ and $\ell'' \leq \ell - 3$, then we can apply case (a) to P'_ℓ . Otherwise the 3 edges of M must be all incident to $V(C_{\ell-2})$, and the ones incident to $V(C_\ell)$ (if any) cannot be adjacent to the two neighbours of $out_{\ell-2}$. Thus there exists one edge $e'' = u''v'' \in M$ with $u'' \in V(C_{\ell-1})$ and $v'' \in V(C_{\ell-2})$ adjacent to $out_{\ell-2}$. Then we can apply case (b) to P'_ℓ . \square

LEMMA D.19. *Suppose that C_ℓ is a triangle. Then in polynomial time one can either compute a gluing path $P_{\ell+1}$ of length $\ell + 1$ or a solution S' satisfying the conditions (1), (2) and (3) of Lemma 2.11.*

Proof. Let us apply Lemma D.14 to P_ℓ , in order to obtain a pair (P'_ℓ, M_ℓ) as in the claim, and set P_ℓ to P'_ℓ . Suppose there exists an edge $e_{\ell+1}$ with one endpoint in $V(C_\ell) \setminus \{in_\ell\}$ (in particular not incident to e_ℓ) and the other in $V^* \setminus (V(C_0) \cup \dots \cup V(C_\ell))$. Then $P_{\ell+1} := P_\ell, e_{\ell+1}$ satisfies the claim.

Otherwise all the edges in M_ℓ have one endpoint in $V(C_\ell)$ and the other one in $V(C_0) \cup \dots \cup V(C_{\ell-1})$. We distinguish two cases:

(a) There exists an edge $e' = u'v' \in E(G)$ with $u' \in V(C_\ell) \setminus \{in_\ell\}$ and $v' \in V(C_{\ell'})$, $\ell' \leq \ell - 2$. We further distinguish some subcases:

(a.1) $C_{\ell'}$ is not a triangle nor a 4-cycle. In this case consider $S' = S \setminus \{in_{\ell'}u'\} \cup \{e_{\ell'}, \dots, e_\ell, e'\}$. Let us remove from S' any edge of type in_jout_j where $\ell' + 1 \leq j \leq \ell - 1$ and C_j is a triangle or a 4-cycle. Let n' be the number of such removed edges, and $n'' = \ell - \ell' - 1 - n'$. Observe that S' satisfies (1) and (2). In particular it contains a 2EC component C' spanning the nodes $V(C_{\ell'}) \cup \dots \cup V(C_\ell)$. S' satisfies (3) since

$$\begin{aligned} cost(S) - cost(S') &= |S| - |S'| + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -n'' - 1 + \frac{3}{10}5 + 1 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - \frac{1}{2} \geq 0. \end{aligned}$$

(a.2) $C_{\ell'}$ is a triangle or a 4-cycle and $C_{\ell'+1}$ is a triangle or a 4-cycle. In this case we can take the same S' as in case (2.a), which satisfies (1) and (2). With the same notation, we observe that in this case the difference is that $n' \geq 1$ (due to $C_{\ell'+1}$) and $cr(C_{\ell'}) \geq 1$. Thus S' satisfies (3) since

$$\begin{aligned} cost(S) - cost(S') &= |S| - |S'| + cr(C_{\ell'}) + cr(C_\ell) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -n'' - 1 + 1 + 1 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' - 1 \geq 0. \end{aligned}$$

(a.3) $C_{\ell'}$ is a triangle or a 4-cycle and $C_{\ell'+1}$ is not a triangle nor a 4-cycle. If v' is a neighbour of $in_{\ell'}$ in $C_{\ell'}$ or $\ell' = 0$ (in which case $in_{\ell'}$ is not defined), consider $P'_\ell := e_1, \dots, e_{\ell'}, e', e_\ell, e_{\ell-1}, \dots, e_{\ell+2}$. Observe that P'_ℓ is an alternative gluing path of length ℓ whose ending component $C_{\ell'+1}$ is not a triangle nor a 4-cycle. The claim then follows by applying one of the Lemmas D.15, D.16, and D.18.

The remaining case is when v' is not a neighbour of $in_{\ell'}$ in $C_{\ell'}$ and $\ell' \neq 0$. Observe that, since P_ℓ is a gluing path and $C_{\ell'}$ is a triangle or a 4-cycle, then v' must be adjacent to $out_{\ell'}$. Consider $S' = S \setminus \{in_{\ell'}u', v'out_{\ell'}\} \cup \{e_{\ell'+1}, \dots, e_\ell, e'\}$. Remove as usual from S' any edge in_jout_j where $\ell' + 1 \leq j \leq \ell - 1$ and C_j is a triangle or a 4-cycle. S' satisfies (1) and (2). With the usual notation for n', n'' and C', S' satisfies

(3) since

$$\begin{aligned} \text{cost}(S) - \text{cost}(S') &= |S| - |S'| + cr(C_{\ell'}) + cr(C_{\ell}) + \sum_{j=\ell'+1}^{\ell-1} cr(C_j) - cr(C') \\ &\geq -n'' + 1 + 1 + n' + \frac{3}{10}5n'' - 2 = \frac{1}{2}n'' + n' \geq 0. \end{aligned}$$

(b) All the edges in M_{ℓ} are between $V(C_{\ell})$ and $V(C_{\ell-1})$ (notice that this must be the case when case (1) does not hold). Since by assumption the conditions of Lemma D.4 do not hold, $C_{\ell-1}$ must be a 6-cycle. Notice that by the definition of S^* , C_{ℓ} is adjacent to at least two components in S^* , hence to at least one component \tilde{C} other than $C_{\ell-1}$. This implies that there exists an edge $e' = vz$ with $v \in V(C_{\ell})$ and $z \in V(\tilde{C})$. More precisely, $v = in_{\ell}$ by the previous cases. In this case we choose any edge $e'' \in M_{\ell} \setminus \{e_{\ell}\}$. If $\tilde{C} \notin \{C_1, \dots, C_{\ell-2}\}$, then $P_{\ell+1} = e_1, \dots, e_{\ell-1}, e'', e'$ satisfies the claim. Otherwise $P'_{\ell} = e_1, \dots, e_{\ell-1}, e''$ is another gluing path of length ℓ whose ending component is C_{ℓ} . We can now apply case (a) to P'_{ℓ} .

□

We now have all the ingredients to prove Lemma D.12.

Proof. [Proof of Lemma D.12] It follows easily from the previous discussion and Lemmas D.15-D.19. □

E A Refined Approximation

In this section we present the following refined version of Lemma 2.5 for the case of many triangles.

LEMMA E.1. *Given a canonical minimum-size 2-edge-cover H of a structured graph G with $b|H|$ bridges and $t|H|$ edges belonging to triangle 2EC components. In polynomial time one can compute a $\frac{14}{9} - \frac{8}{27}t + \frac{4}{9}b$ approximate solution for 2-ECSS on G .*

The proof of Theorem 1.1 follows easily.

Proof. [Proof of Theorem 1.1] It is sufficient to describe a $\frac{118}{89}$ -approximation algorithm for the case of a structured graph G . The claim then follows from Lemma 2.2. Let H be obtained from a minimum-size 2-edge-cover of G via Lemma 2.4. Taking the best solution among the ones guaranteed by Lemmas E.1 and 2.6 w.r.t H , with the usual notation for t and b one obtains an approximation factor

$$\begin{aligned} \text{apx}(t, b) &:= \min\{\text{apx}_{mny}(t, b), \text{apx}_{few}(t, b)\}, \\ \text{apx}_{mny}(t, b) &:= \frac{14}{9} - \frac{8}{27}t + \frac{4}{9}b, \quad \text{apx}_{few} := \frac{13}{10} + \frac{1}{30}t - \frac{1}{20}b. \end{aligned}$$

For a fixed value of b , $\text{apx}_{few}(t, b)$ is an increasing function of t , while $\text{apx}_{mny}(t, b)$ is a decreasing function of t . The two factors are equal for $t = t(b) := \frac{69}{89} + \frac{3}{2}b$. Notice that $t(b) \geq 0$. For the values of b such that $t(b) \leq 1 - b$, namely for $b \leq \frac{8}{89}$, the worst case is achieved for $t = t(b)$, leading to the approximation factor

$$\text{apx}(b) := \frac{13}{10} + \frac{1}{30}\left(\frac{69}{89} + \frac{3}{2}b\right) - \frac{1}{20}b = \frac{118}{89}.$$

For $b \geq \frac{8}{89}$, the worst case is achieved for $t = 1 - b$ and it is given by $\text{apx}_{few}(t, b)$, leading to the approximation factor

$$\text{apx}(b) := \text{apx}_{few}(1 - b, b) = \frac{13}{10} + \frac{1}{30}(1 - b) - \frac{1}{20}b = \frac{4}{3} - \frac{1}{12}b \leq \frac{118}{89}.$$

□

It remains to prove Lemma E.1. The initial part of the construction, leading to a core-triangle 2-edge-cover, is the same as in Lemma 2.5. In particular, we will obtain a 2-edge-cover S consisting of a core C (which is 2EC) and k triangles T_1, \dots, T_k , such that nodes of distinct triangles are not adjacent in G . Recall that at this point in Lemma 2.5 we consider a solution APX obtained from S by adding any two edges between two distinct nodes

$v_i, u_i \in V(T_i)$ and C , and removing the edge $v_i u_i$: let Q_i be the (four) edges with at least one endpoint in $V(T_i)$. Our final solution is $APX = E(C) \cup \bigcup_{i=1}^k Q_i$. We recall that

$$|APX| = |E(C)| + 4k = |S| + k \leq \left(\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b\right)|H| + \frac{k}{2}.$$

The main idea here is to choose the Q_i 's more carefully among the possible options. To do that, we use an approach very similar in spirit to the one in [34].

In more detail, let $Q = \bigcup_{i=1}^k Q_i$. For a subset of edges F , let $\alpha(F)$ be the number of connected components of the graph (V, F) . Define $\alpha(S)$ to be the minimum possible value of $\alpha(Q)$ for any valid choice of Q for a given core-triangle 2-edge-cover S . The following fact follows similarly to the reduction to matroid intersection in [34] for the computation of a maximum earmuff:

FACT E.1. *Given a core-triangle 2-edge-cover S of G , there exists a polynomial time algorithm that computes Q^* minimizing $\alpha(Q)$ over the set of edges Q of the above type, i.e. $\alpha(Q^*) = \alpha(S)$.*

Proof. Consider the following algorithm. For each Q_i that contains a 2-matching $u_1 v_1, u_2 v_2$ between T_i and C (with $v_i \in V(C)$) we define a pseudo-edge $v_1 v_2$ of color i . We next compute a maximum size set F of pseudo-edges such that $(V(C), F)$ is a forest and no two edges of F have the same color. This can be done in polynomial time as follows. Let M_1 be the graphic matroid induced by the pseudo-edges E' , and M_2 be the partition matroid induced by the colors of the pseudo-edges E' . Then F is a maximum cardinality independent set in the intersection of M_1 and M_2 , which can be computed in polynomial time (see, e.g., [15, 33]).

Given F , we build a feasible Q as follows. For each pseudo-edge $f \in F$ of color i , we add the corresponding Q_i to Q . Notice that at this point $\alpha(Q)$ is precisely the number of connected components induced by F , namely $|V(C)| - |F|$. Then, for each missing color j in F , we add an arbitrary Q_j to Q . Notice that this does not reduce $\alpha(Q)$ by the optimality of F . In particular, $\alpha(Q) = |V(C)| - |F|$.

It remains to argue that the above Q minimizes $\alpha(Q)$. To see that, consider the optimal $Q^* = \bigcup_{i=1}^k Q_i^*$. We next iteratively remove from Q^* any Q_j^* whose removal does not increase $\alpha(Q^*)$ (in particular this always happens when Q_j^* is incident to a unique node of $V(C)$). Let F^* be the pseudo-edges corresponding to the remaining Q_i^* 's in Q^* . The number of connected components $|V(C)| - |F^*|$ induced by F^* is precisely $\alpha(Q^*)$. Furthermore, F^* is in the intersection of M_1 and M_2 , hence $|F^*| \leq |F|$ (implying $\alpha(Q) \leq \alpha(Q^*)$). \square

We next assume that Q is obtained as in the above fact in the computation of APX, a define the Q_i 's accordingly. Analogously to [34], we next define an alternative solution APX' of size $\text{apx}' = |APX'|$ starting from S as follows. We start from Q , and add a minimal number of edges $F \subseteq G[V(C)]$ so that $Q \cup F$ becomes a connected spanning subgraph. Obviously $|F| = \alpha(S) - 1$. Then we compute a minimum-size T -join¹⁴ J over the odd degree nodes T of $Q \cup F$. Notice that $T \subseteq V(C)$ since the nodes of the triangles have even degree in Q and this degree is not modified by F . We initially set $APX' := Q \cup F \cup J$. We observe that APX' is a connected Eulerian graph, hence in particular it is 2EC. Notice however that APX' might contain parallel edges: for any two parallel edges f and g (which are effectively copies of the same edge e), we remove g and add some other edge g' connecting the two sides of the cut induced by f in APX' . Notice that such g' must exist since G is 2EC. Furthermore, this process keeps APX' 2EC and does not change its size. By repeating this process we turn APX' into a simple 2EC spanning subgraph of size $\text{apx}' = 4k + \alpha(S) - 1 + |J|$.

We will next show that $|J| \leq \frac{1}{2}|E(C)|$. To this aim we will exploit a classical result by Frank [19]. We recall that an ear decomposition of a graph is obtained incrementally as follows. The first ear E_1 is a cycle. Given $F_i = E_1 \cup \dots \cup E_i$, the ear E_{i+1} is a path with its endpoints in $V(F_i)$ and the other nodes not in $V(F_i)$, or alternatively a cycle with exactly one node in $V(F_i)$. An ear is even (resp., odd) if its number of edges is so. A well-known fact is that every 2EC graph G admits an ear decomposition spanning all its edges. By $\phi(G)$ we denote the minimum possible number of even ears in any ear decomposition of G .

LEMMA E.2. [19] *Let G be a 2EC, and $T \subseteq V(G)$, $|T|$ even. Then there exists a T -join of G of size at most $\frac{L_{\phi(G)}}{2}$, where $L_{\phi(G)} = |V(G)| + \phi(G) - 1$. Furthermore $\text{opt}(G) \geq L_{\phi(G)}$.*

¹⁴Recall that, for a subset of nodes T of even cardinality, a T -join is a subgraph such that the nodes in T have odd degree and the other nodes even degree.

Recall that the core C is 2EC and that, for the T -join J , $T \subseteq V(C)$. Therefore we can apply Lemma E.2 to infer that $|E(C)| \geq \text{opt}(G[V(C)]) \geq L_{\phi(G[V(C)])} \geq 2|J|$ as desired. As a consequence

$$\text{apx}' = |APX'| = |Q| + |F| + |J| = 4k + \alpha(S) - 1 + |J| \leq 4k + \alpha(S) - 1 + \frac{1}{2}|E(C)|.$$

The next step is to define a refined lower bound on opt which takes $\alpha(S)$ into account.

LEMMA E.3. *Let F^* be a 2EC spanning subgraph of G including $E(C)$, and Q^* be the edges of F^* with at least one endpoint not in C . Then $\alpha(Q^*) \geq 4k - |Q^*| + \alpha(S)$.*

Proof. We prove the claim by induction on $|Q^*|$. As already observed, F^* must include at least 4 edges with at least one endpoint in each T_i , hence $|Q^*| \geq 4k$. The base case is $|Q^*| = 4k$. In this case there are precisely 4 edges incident to the nodes $V(T_i) = \{u_i, v_i, w_i\}$ of each T_i . These edges must induce either a path of type a_i, u_i, v_i, w_i, b_i or a cycle of type a_i, u_i, v_i, w_i, a_i , where $a_i, b_i \in V(C)$. In particular Q^* is a valid choice for Q , hence $\alpha(Q^*) \geq \alpha(S) = 4k - |Q^*| + \alpha(S)$.

In the inductive case $|Q^*| = h > 4k$ there must exist a triangle T_i such that at least 5 edges of Q^* are incident to T_i . We distinguish 4 cases depending on the number $t_i \in \{0, 1, 2, 3\}$ of these edges with both endpoints in T_i .

($t_i = 3$) Here we distinguish two subcases. Suppose first that there exist two nodes in T_i , say u_i and v_i , with degree at least 3 in Q^* . Let $Q' = Q^* \setminus \{u_i v_i\}$. Notice that $F' := E(C) \cup Q'$ is a feasible 2EC spanning subgraph and $\alpha(Q') = \alpha(Q^*)$. By inductive hypothesis $\alpha(Q') \geq 4k - |Q^*| + \alpha(S) + 1$, hence the claim.

It remains to consider the case that there exists exactly one node of T_i , say u_i , with degree at least 3 in Q^* . Since there must be at least 2 edges in Q^* with exactly one endpoint in T_i , this implies $\deg_{Q^*}(u_i) \geq 4$. Let a_i be a node in C that is adjacent to u_i in Q^* . Notice that, since u_i is not a cut node, there must exist some edge from some other node in T_i to C , say $v_i b_i$. Let $Q' = Q^* \cup \{v_i b_i\} \setminus \{u_i a_i, u_i v_i\}$. Notice that $F' = E(C) \cup Q'$ is a feasible 2EC spanning subgraph, $|Q'| = |Q^*| - 1$, and $\alpha(Q') \leq \alpha(Q^*) + 1$. Therefore, by the inductive hypothesis on Q' , $\alpha(Q^*) \geq \alpha(Q') - 1 \geq 4k - |Q^*| + \alpha(S)$.

($t_i = 2$) The 2 edges in $E(T_i) \cap Q^*$ form a path P of length 2, say v_i, u_i, w_i . There must be a node v in T_i such that $\deg_{Q^*}(v) \geq 3$. Indeed there are at least 3 edges in Q^* incident to nodes of T_i and to C . If one of them is incident to u_i , say $a_i u_i$, then consider $Q' = Q^* \setminus \{a_i u_i\}$. Again $F' = E(C) \cup Q'$ is a feasible 2EC spanning subgraph and the claim follows similarly to previous cases. Otherwise w.l.o.g assume there are two edges in $Q^* \setminus E(T_i)$ incident to v_i . Let $Q' = Q^* \setminus \{e\}$, where e is an edge incident to v_i not in P . As usual $F' = E(C) \cup Q'$ is a feasible 2EC spanning subgraph and $\alpha(Q') \leq \alpha(Q^*) - 1$. Thus, by the inductive hypothesis on Q' , $\alpha(Q^*) \geq \alpha(Q') - 1 \geq 4k - |Q^*| + \alpha(S)$.

($t_i = 1$) Assume w.l.o.g. that $v_i w_i \in Q^*$, hence u_i is not adjacent to v_i nor to w_i in Q^* . We distinguish two sub-cases. If v_i does not belong to the same connected component of u_i in (V, Q^*) , then we let $Q' = Q^* \cup \{u_i v_i\} \setminus \{e'\}$ where e' is any other edge e' incident to u_i . Otherwise, we define Q' similarly but taking care of selecting an $e' \neq u_i v_i$ incident to u_i along a path from u_i to v_i in (V, Q^*) . In both cases we have $|Q'| = |Q^*|$, $\alpha(Q') = \alpha(Q^*)$, and $F' = E(C) \cup Q'$ is a feasible 2EC spanning subgraph. Now Q' contains 2 edges of $E(T_i)$, hence the claim follows from the previous analysis.

($t_i = 0$) We use the same construction as in the case $t_i = 1$, considering again nodes u_i and v_i . In particular, by adding the edge $u_i v_i$ we reduce to the case $t_i = 1$, which is addressed above. \square

Recall that we used the lower bound $\text{opt}(G) \geq 4k$. We can next use the following stronger lower bound.

COROLLARY E.1. *Let S be a core-triangle 2-edge-cover of a structured graph G with k triangles. Then $\text{opt}(G) \geq 4k + \alpha(S) - 1$.*

Proof. Consider the edges Q^* of $\text{OPT}(G)$ with at least one endpoint not in the core C of S . By Lemma E.3, $(V(G), Q^*)$ has at least $\alpha(Q^*) \geq 4k - |Q^*| + \alpha(S)$ connected components. Therefore $\text{OPT}(G)$, which is connected, must contain at least $4k - |Q^*| + \alpha(S) - 1$ edges, besides Q^* to connect the connected components of $(V(G), Q^*)$. The claim follows. \square

We are now ready to prove Lemma E.1.

Proof. [Proof of Lemma E.1] Recall that we computed two feasible solutions APX and APX' satisfying

$$|APX| = |E(C)| + 4k \leq \left(\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b\right)|H| + \frac{k}{2} \quad \text{and} \quad |APX'| \leq 4k + \alpha(S) - 1 + \frac{1}{2}|E(C)|.$$

By Corollary E.1 we have that $\text{opt} = \text{opt}(G) \geq 4k + \alpha(S) - 1$. We will also use $\text{opt} \geq |H|$ as usual. Notice that

$$|S| = 3k + |E(C)| \leq \left(\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b\right)|H| - \frac{k}{2} =: f(t, b)|H| - \frac{k}{2},$$

hence

$$\text{opt} \geq |H| \geq \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}$$

Observe that $\frac{7}{6} \leq f(t, b) \leq 2$. Recall that $|APX_B| = |S| + k = |C| + 4k$. The overall approximation factor is upper bounded by

$$\text{apx}(t, b) := \frac{\min\{|APX|, |APX'|\}}{\max\{|H|, 4k + \alpha(S) - 1\}} \leq \frac{4k + \min\{|E(C)|, \alpha(S) - 1 + \frac{1}{2}|E(C)|\}}{\max\{\frac{\frac{7}{2}k + |E(C)|}{f(t, b)}, 4k + \alpha(S) - 1\}}.$$

We distinguish 4 cases depending on which is the smallest upper bound on the cost of the approximate solution and the largest lower bound on the optimal cost:

(1) $\alpha(S) - 1 \geq \frac{1}{2}|E(C)|$ **and** $4k + \alpha(S) - 1 \geq \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}$. In this case:

$$\text{apx}(t, b) \leq \frac{4k + |E(C)|}{4k + \alpha(S) - 1} \leq \frac{4k + |E(C)|}{\max\{4k + \frac{1}{2}|E(C)|, \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}\}}.$$

The righthand side of the above inequality is an increasing function of $|E(C)|$ for $4k + \frac{1}{2}|E(C)| \geq \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}$, i.e. for $|E(C)| \leq \frac{8f(t, b) - 7}{2 - f(t, b)}k$, and a decreasing function of $|E(C)|$ in the complementary case. Thus the worst case is achieved for $|E(C)| = \frac{8f(t, b) - 7}{2 - f(t, b)}k$, leading to

$$\text{apx}(t, b) \leq \frac{4k + \frac{8f(t, b) - 7}{2 - f(t, b)}k}{4k + \frac{1}{2} \frac{8f(t, b) - 7}{2 - f(t, b)}k} = \frac{2 + 8f(t, b)}{9}.$$

(2) $\alpha(S) - 1 \geq \frac{1}{2}|E(C)|$ **and** $4k + \alpha(S) - 1 < \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}$. This imposes $\frac{\frac{7}{2}k + |E(C)|}{f(t, b)} - 4k > \frac{1}{2}|E(C)|$, i.e. $|E(C)| > \frac{8f(t, b) - 7}{2 - f(t, b)}k$. In this case

$$\text{apx}(t, b) \leq \frac{4k + |E(C)|}{\frac{\frac{7}{2}k + |E(C)|}{f(t, b)}}.$$

The right hand side of the above inequality is a growing function of $|E(C)|$ due to the fact that $f(t, b) \geq \frac{7}{6} \geq \frac{8}{7}$. Hence the worst case is achieved for $|E(C)|$ going to infinity, leading to

$$\text{apx}(t, b) \leq f(t, b).$$

Notice that $\frac{2 + 8f(t, b)}{9} \geq f(t, b)$ iff $2 \geq f(t, b)$. Hence case (1) dominates case (2) in terms of the worst-case approximation factor.

(3) $\alpha(S) - 1 < \frac{1}{2}|E(C)|$ **and** $4k + \alpha(S) - 1 \geq \frac{\frac{7}{2}k + |E(C)|}{f(t, b)}$. This imposes $\frac{\frac{7}{2}k + |E(C)|}{f(t, b)} - 4k < \frac{1}{2}|E(C)|$, i.e. $\frac{2 - f(t, b)}{8f(t, b) - 7}|E(C)| < k$. In this case one has

$$(E.1) \quad \text{apx}(t, b) \leq \frac{\alpha(S) - 1 + \frac{1}{2}|E(C)|}{\alpha(S) - 1 + 4k} \leq \frac{\alpha(S) - 1 + \frac{1}{2}|E(C)|}{\alpha(S) - 1 + 4 \frac{2 - f(t, b)}{8f(t, b) - 7}|E(C)|}.$$

For $\frac{1/2}{4\frac{2-f(t,b)}{8f(t,b)-7}} = \frac{8f(t,b)-7}{8(2-f(t,b))} \geq 1$, i.e. for $f(t,b) \geq \frac{23}{16}$, the righthand side of (E.1) is decreasing in $\alpha(S) - 1$. Hence the worst case is achieved for $\alpha(S) - 1 = \frac{\frac{7}{2}k + |E(C)|}{f(t,b)} - 4k$ leading to

$$apx(t,b) \leq \frac{\frac{\frac{7}{2}k + |E(C)|}{f(t,b)} - 4k + \frac{1}{2}|E(C)|}{\frac{\frac{7}{2}k + |E(C)|}{f(t,b)} - 4k + 4\frac{2-f(t,b)}{8f(t,b)-7}|E(C)|}.$$

For $f(t,b) \geq \frac{23}{16}$ the righthand side of the above inequality is increasing in $|E(C)|$, hence that quantity is upper bounded by setting $|E(C)| = \frac{8f(t,b)-7}{2-f(t,b)}k$, leading to

$$apx(t,b) \leq \frac{\frac{7k}{2f(t,b)} - 4k + \frac{2+f(t,b)}{2f(t,b)}\frac{8f(t,b)-7}{2-f(t,b)}k}{\frac{7k}{2f(t,b)} + \frac{1}{f(t,b)}\frac{8f(t,b)-7}{2-f(t,b)}k} = \frac{16f(t,b) - 14}{9}.$$

For $f(t,b) < \frac{23}{16}$, the righthand side of (E.1) is increasing in $\alpha(S) - 1$. Hence we can upper bound it by setting $\alpha(S) - 1 = \frac{1}{2}|E(C)|$, leading to

$$apx(t,b) \leq \frac{|E(C)|}{\frac{1}{2}|E(C)| + 4\frac{2-f(t,b)}{8f(t,b)-7}|E(C)|} = \frac{2(8f(t,b) - 7)}{8f(t,b) - 7 + 8(2 - f(t,b))} = \frac{16f(t,b) - 14}{9}$$

We observe that $\frac{16f(t,b)-14}{9} \leq \frac{2+8f(t,b)}{9}$ for $f(t,b) \leq 2$. Hence case (3) is dominated by case (1).

(4) $\alpha(S) - 1 < \frac{1}{2}|E(C)|$ **and** $4k + \alpha(S) - 1 < \frac{\frac{7}{2}k + |E(C)|}{f(t,b)}$. In this case one has

$$apx(t,b) \leq \frac{\alpha(S) - 1 + \frac{1}{2}|E(C)|}{\frac{\frac{7}{2}k + |E(C)|}{f(t,b)}} \leq \frac{\min\{\frac{1}{2}|E(C)|, \frac{\frac{7}{2}k + |E(C)|}{f(t,b)} - 4k\} + \frac{1}{2}|E(C)|}{\frac{\frac{7}{2}k + |E(C)|}{f(t,b)}}.$$

For $\frac{1}{2}|E(C)| \leq \frac{\frac{7}{2}k + |E(C)|}{f(t,b)} - 4k$, i.e. for $|E(C)| \geq k\frac{8f(t,b)-7}{2-f(t,b)}$, the righthand side of the above inequality is $\frac{f(t,b)|E(C)|}{\frac{7}{2}k + |E(C)|}$, which is a growing function of $|E(C)|$. Hence the worst-case value is $f(t,b)$. Like in case (2), this case is dominated by case (1). Otherwise, i.e. for $k > \frac{2-f(t,b)}{8f(t,b)-7}|E(C)|$, the righthand side of the above inequality is $1 + f(t,b)\frac{\frac{1}{2}|E(C)| - 4k}{\frac{1}{2}|E(C)| + \frac{7}{2}k}$. This is a decreasing function of k , hence it is upper bounded by setting $k = \frac{2-f(t,b)}{8f(t,b)-7}|E(C)|$. This leads to

$$apx(t,b) \leq 1 + f(t,b)\frac{\frac{1}{2}|E(C)| - 4\frac{2-f(t,b)}{8f(t,b)-7}|E(C)|}{|E(C)| + \frac{7}{2}\frac{2-f(t,b)}{8f(t,b)-7}|E(C)|} = \frac{16f(t,b) - 14}{9}.$$

This matches the upper bound achieved in case (3), hence in particular it is dominated by case (1). Altogether

$$apx(t,b) \leq \frac{2 + 8f(t,b)}{9} = \frac{1}{9}\left(2 + 8\left(\frac{3}{2} - \frac{1}{3}t + \frac{1}{2}b\right)\right) = \frac{14}{9} - \frac{8}{27}t + \frac{4}{9}b.$$

□

F Some Previous Attempts to Breach the 4/3 Approximation Barrier

To the best of our knowledge, there are only two (conference) papers which claim a better than $\frac{4}{3}$ approximation in the literature¹⁵: Krysta and Kumar [30] claim a $\frac{4}{3} - \frac{1}{1344}$ approximation and Jothi, Raghavachari, and Varadarajan [26] a $\frac{5}{4}$ approximation. Both papers (after about two decades) have no journal version nor any public full version to the best of our knowledge. We remark that the two papers presenting a 4/3 approximation algorithm that we mentioned earlier, namely [24, 34], do not include [26, 30] in their references, despite the fact that [24, 34] were

¹⁵More recently, Civril made several failed attempts to improve on 4/3. However the last version of his work on ArXiv [11] only claims an alternative 4/3 approximation for the general case.

published several years later and claim a worse approximation factor w.r.t. [26, 30]. In particular, [34] states: “Better approximation ratios have been claimed, but to the best of our knowledge, no correct proof has been published”. We next give more details about the gaps in [26, 30].

[30] gives a rather vague description of the algorithm and of its analysis (only about 4 pages altogether), with no formal proof. We are not even close to being able to fill in the gaps in [30]. The most critical part in our opinion is Section 3.3 “2-Connecting Paths”. Here even the description of the algorithm is not fully detailed, as mentioned by the authors: “The above cases have been stated in a highly oversimplified manner”. The subsequent case analysis is only sketched at a very high level. In our experience, most of the difficulty in case-analysis-based algorithms (like [24, 26, 30] and this paper) is due to handling very specific extreme cases. Hence it is very hard for the reader to verify the correctness without seeing the full details. Let us mention that [30] heavily builds upon the algorithm and analysis in [37], which is the conference version of [24]. [37] contains some gaps (as declared by a superset of the authors in [24]). This makes even more critical that the authors of [30] publish a full and self-contained version of their paper. Finally, we remark that the approximation factor in [30] is worse than the one claimed in this paper (hence here the only unclear point is which is the first paper to breach the $4/3$ approximation barrier).

[26] is much more detailed than [30], still major parts of their analysis are missing. The most important such gap (but not the only one), in our opinion, is in Section 4.1 “Finding a 5-cycle”. The authors give a full proof for the case that the leaf vertex v_1 corresponds to a single node of G (Lemma 4.1). The authors claim (Lemma 4.2) that they can extend that lemma to the case where v_1 is a “2-connected component that has been shrunk into a single node”, but no proof is provided. The authors mention: “The proof is more complicated . . . There are many different cases to consider, and we omit the proof here due to lack of space”. Proving Lemma 4.2, if possible, is in our opinion highly non-trivial and we are not able to reconstruct their proof on our own. As mentioned before, missing even one case might lead to a too optimistic (hence wrong) approximation factor. So it is particularly important that the community has the chance to verify this type of case analysis.