

# A $4/3$ -Approximation Algorithm for the Minimum 2-Edge Connected Multisubgraph Problem in the Half-Integral Case

**Sylvia Boyd**

School of Electrical Engineering and Computer Science, University of Ottawa, Canada  
sboyd@uottawa.ca

**Joseph Cheriyan**

Department of Combinatorics and Optimization, University of Waterloo, Canada  
jcheriyan@uwaterloo.ca

**Robert Cummings**

Department of Combinatorics and Optimization, University of Waterloo, Canada

**Logan Grout**

Department of Combinatorics and Optimization, University of Waterloo, Canada

**Sharat Ibrahimpur** 

Department of Combinatorics and Optimization, University of Waterloo, Canada  
sharat.ibrahimpur@uwaterloo.ca

**Zoltán Szigeti**

University Grenoble Alpes, CNRS, G-SCOP, France  
zoltan.szigeti@grenoble-inp.fr

**Lu Wang**

Department of Combinatorics and Optimization, University of Waterloo, Canada

---

## Abstract

Given a connected undirected graph  $\overline{G}$  on  $n$  vertices, and non-negative edge costs  $c$ , the 2ECM problem is that of finding a 2-edge connected spanning multisubgraph of  $\overline{G}$  of minimum cost. The natural linear program (LP) for 2ECM, which coincides with the subtour LP for the Traveling Salesman Problem on the metric closure of  $\overline{G}$ , gives a lower bound on the optimal cost. For instances where this LP is optimized by a half-integral solution  $x$ , Carr and Ravi (1998) showed that the integrality gap is at most  $\frac{4}{3}$ : they show that the vector  $\frac{4}{3}x$  dominates a convex combination of incidence vectors of 2-edge connected spanning multisubgraphs of  $\overline{G}$ .

We present a simpler proof of the result due to Carr and Ravi by applying an extension of Lovász's splitting-off theorem. Our proof naturally leads to a  $\frac{4}{3}$ -approximation algorithm for half-integral instances. Given a half-integral solution  $x$  to the LP for 2ECM, we give an  $O(n^2)$ -time algorithm to obtain a 2-edge connected spanning multisubgraph of  $\overline{G}$  whose cost is at most  $\frac{4}{3}c^T x$ .

**2012 ACM Subject Classification** Mathematics of computing → Paths and connectivity problems; Mathematics of computing → Approximation algorithms; Theory of computation → Routing and network design problems

**Keywords and phrases** 2-Edge Connectivity, Approximation Algorithms, Subtour LP for TSP

**Digital Object Identifier** 10.4230/LIPIcs.APPROX/RANDOM.2020.61

**Category** APPROX

**Funding** *Joseph Cheriyan*: Supported by NSERC RGPIN-2019-04197.

*Sharat Ibrahimpur*: Supported in part by NSERC grant 327620-09.

**Acknowledgements** SB, JC, and SI thank BIRS, Canada for organizing the workshop on *The Traveling Salesman Problem* (2018), where some of the results in our references were presented.



© Sylvia Boyd, Joseph Cheriyan, Robert Cummings, Logan Grout, Sharat Ibrahimpur, Zoltán Szigeti, and Lu Wang;

licensed under Creative Commons License CC-BY

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2020).

Editors: Jarosław Byrka and Raghv Meka; Article No. 61; pp. 61:1–61:12



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

## 1 Introduction

The 2-edge connected multisubgraph (2ECM) problem is a fundamental problem in survivable network design where one wants to be resilient against a single edge failure. In this problem, we are given an undirected graph  $\overline{G} = (\overline{V}, \overline{E})$  with non-negative edge costs  $c$  and we want to find a 2-edge connected spanning multisubgraph of  $\overline{G}$  of minimum cost. Below we give an integer linear program for 2ECM. The variable  $x_e$  denotes the number of copies of edge  $e$  that are used in a feasible solution. For any  $S \subset V$ ,  $\delta(S) := \{e = uv \in \overline{E} : u \in S, v \notin S\}$  denotes the cut induced by  $S$ . For any  $F \subseteq E$  and vector  $x \in \mathbb{R}^E$ , we use  $x(F)$  as a shorthand for  $\sum_{e \in F} x_e$ . Also, for any graph  $H$  with edge costs  $c$ , we sometimes use  $c(H)$  as a shorthand for  $c(E(H))$ .

$$\min \sum_{e \in \overline{E}} c_e x_e \quad (1)$$

$$(2\text{ECM-IP}) \quad \text{subject to} \quad x(\delta(S)) \geq 2 \quad \forall \emptyset \subsetneq S \subsetneq \overline{V}, \quad (2)$$

$$x_e \geq 0 \quad \forall e \in \overline{E}, \quad (3)$$

$$x_e \text{ integral} \quad \forall e \in \overline{E}. \quad (4)$$

It is easy to see that an optimal solution for 2ECM never has to use more than two copies of an edge. As is discussed in [6], since we are allowed to use more than one copy of an edge, without loss of generality, we may assume that  $\overline{G}$  is complete by performing the metric completion: for each  $u, v \in \overline{V}$  we set the new cost of the edge  $uv$  to be the shortest path distance between  $u$  and  $v$  in  $\overline{G}$ . In the sequel, we assume that  $\overline{G}$  is a complete graph and that the cost function  $c$  is metric i.e.,  $c \geq 0$  and for every  $u, v, w \in \overline{V}$ , we have  $c_{uw} \leq c_{uv} + c_{vw}$ .

The linear relaxation (2ECM-LP) for 2ECM is obtained by dropping the integrality constraints given by (4). By a result due to Goemans and Bertsimas [8] called the *parsimonious property*, adding the constraint  $x(\delta(v)) = 2$  for each  $v \in \overline{V}$  to (2ECM-LP) does not increase the optimal solution value; here, we require the assumption that the costs form a metric. So, the optimal value of (2ECM-LP) is the same as the optimal value for the well-known subtour elimination LP (Subtour-LP) for the Traveling Salesman Problem (TSP) defined below. Due to this connection, we often refer to an optimal solution for (2ECM-LP) as an optimal solution to (Subtour-LP), and vice versa. Another consequence of the parsimonious property is that for graphs with at least 3 vertices, the constraint  $x_e \leq 1$  is implied by (Subtour-LP): for any  $e = uv$ , we have  $2x_e = x(\delta(u)) + x(\delta(v)) - x(\delta(\{u, v\})) \leq 2$ .

$$\min \sum_{e \in \overline{E}} c_e x_e \quad (5)$$

$$(Subtour-LP) \quad \text{subject to} \quad x(\delta(S)) \geq 2 \quad \forall \emptyset \subsetneq S \subsetneq \overline{V}, \quad (6)$$

$$x(\delta(v)) = 2 \quad \forall v \in \overline{V}, \quad (7)$$

$$x_e \geq 0 \quad \forall e \in \overline{E}. \quad (8)$$

A long-standing open problem called the ‘‘four-thirds conjecture’’ states that the integrality gap of (Subtour-LP) is  $\frac{4}{3}$ . Besides the importance of 2ECM in the field of survivable network design, the connection between (2ECM-LP) and (Subtour-LP) has spurred interest in determining the integrality gap for (2ECM-LP) as a means to gaining useful lower bounds on the integrality gap for (Subtour-LP). The general version of metric TSP has resisted all attempts at proving an upper bound better than  $\frac{3}{2}$  on the integrality gap, so a great deal of research has focused on obtaining improvements for important special cases. In [13], the authors conjecture that the integrality gap for (Subtour-LP) is achieved on instances where

an optimal (fractional) solution to (Subtour-LP) is half integral i.e.,  $2x_e \in \mathbb{Z}_{\geq 0}$  for all  $e$ . We refer to such instances as *half integral instances*. More than two decades ago, Carr and Ravi [6] proved that the integrality gap of (2ECM-LP) is at most  $\frac{4}{3}$  in the half-integral case. They show that  $\frac{4}{3}x$  dominates a convex combination of 2-edge connected spanning multisubgraphs of  $\overline{G}$ . This supports the four-thirds conjecture for TSP since the (integer) optimal value for 2ECM lower bounds the (integer) optimal value for TSP. However, the proof of Carr and Ravi does not give a polynomial-time algorithm for 2ECM. Very recently, in [12], Karlin, Klein, and Oveis Gharan gave a randomized approximation algorithm for half-integral instances of TSP whose (expected) approximation factor is  $\frac{3}{2} - 0.00007$ . This immediately implies a better than  $\frac{3}{2}$ -approximation algorithm, albeit randomized, for 2ECM as well.

We note that the result of Carr and Ravi mentioned above does not apply to the strict variant of 2ECM (henceforth denoted by 2ECS) where we are allowed to pick at most one copy of an edge in  $\overline{G}$ , i.e. where we are considering subgraphs of  $\overline{G}$  rather than multisubgraphs; similarly, our main result does not apply to 2ECS.

## 1.1 Our Work

Our main contribution is a deterministic approximation algorithm for 2ECM on half-integral instances that matches the existence result in [6].

► **Theorem 1.** *Let  $x$  denote an optimal half-integral solution to an instance  $(\overline{G}, c)$  of (Subtour-LP) (and (2ECM-LP)). There is an  $O(|V(\overline{G})|^2)$ -time algorithm for computing a 2-edge connected spanning multisubgraph of  $\overline{G}$  with cost at most  $\frac{4}{3}c^T x$ .*

Given a half-integral solution  $x$  to (Subtour-LP) for  $\overline{G}$ , let  $G = (V, E)$  denote the multigraph induced by  $2x$ . Formally, the vertex-set  $V := \overline{V}$ , and for each edge  $e \in \overline{E}$ , the edge-set  $E$  has  $2x_e$  copies of the edge  $e$ . Note that if  $|\overline{V}| \geq 3$ , then  $2x_e \in \{0, 1, 2\}$  for all  $e \in \overline{E}$ , and if  $|\overline{V}| = 2$ , then  $2x_e = 4$  for the unique edge  $e \in \overline{E}$ . With a slight abuse of notation, we use the same cost function  $c$  to denote the edge costs in  $G$  i.e.,  $c_f := c_e$  where  $e \in \overline{E}$  gave rise to the edge  $f \in E$ . By (7) and (6),  $G$  is a 4-regular 4-edge connected multigraph. Theorem 1 follows from the following result applied to the graph  $G$  induced by  $2x$ .

► **Theorem 2.** *Let  $G = (V, E)$  be a 4-regular 4-edge connected multigraph on  $n$  vertices. Let  $c : E \rightarrow \mathbb{R}$  be an arbitrary cost function on the edges of  $G$  (negative costs on the edges are allowed), and let  $e$  be an arbitrary edge in  $G$ . Then, in  $O(n^2)$  time, we can find a 2-edge connected spanning subgraph  $H$  of  $G - e$  satisfying:*

- (i)  $c(H) \leq \frac{2}{3}c(G - e)$ ; and
- (ii) each multiedge of  $G$  appears at most once in  $H$  (multiedges may arise in  $H$  due to multiedges in  $G$ ).

For any  $F \subseteq E$ , let  $\chi^F \in \{0, 1\}^E$  denote the characteristic vector of  $F$ :  $\chi_e^F = 1$  if and only if  $e \in F$ . Note that distinct multiedges in  $E$  correspond to distinct coordinates in  $\chi^F$ . As mentioned before, Carr and Ravi prove the existence of such a subgraph  $H$  by showing that for any 4-regular 4-edge connected multigraph  $G$ , there exists a finite collection  $H_1, \dots, H_k$  of 2-edge connected spanning subgraphs of  $G$  such that  $\frac{2}{3}\chi^{E(G) \setminus \{e\}}$  lies in the convex hull of  $\{\chi^{H_i}\}_i$ . At a high level, their proof is inductive and splits into two cases based on whether  $G$  has a certain kind of a tight set (a cut of size 4). In the first case they construct two smaller instances of the problem by contracting each of the shores of the tight set, and in the second case they perform two distinct splitting-off operations at a designated vertex to obtain two smaller instances of the problem. In either case, the convex combinations from the

two subinstances are merged to obtain a convex combination for  $G$ . The first case requires gluing since the subgraphs obtained from the two subinstances need to agree on a (tight) cut. Merging the convex combinations arising from the second case is rather straightforward as the two subinstances are more or less independent.

Our first insight in this work is that the case from Carr and Ravi’s proof that requires the gluing step can be completely avoided, thereby unifying the analysis. This is discussed in Section 2. Our proof relies on an extension of Lovász’s splitting-off theorem that is due to Bang-Jensen et al., [2]. For further discussion on splitting-off theorems, see [7, Chapter 8]. The challenge in efficiently finding a cheap subgraph  $H$  from the above convex combination construction is that each inductive step requires solving two subinstances of the problem, each with one fewer vertex, leading to an exponential-time algorithm. Having said that, an (expected) polynomial-time Las Vegas randomized algorithm can be easily designed that randomly recurses on one of the two subinstances and produces a 2-edge connected spanning subgraph whose expected cost is at most  $\frac{2}{3}c(G - e)$ . Our second insight, which is used in derandomizing the above procedure, is that it is easy to recognize which of the two subinstances leads to a “cheaper” solution, so we recurse only on the cheaper subinstance. Complementing this step, we lift the solution back to the original instance. This operation can lead to two different outcomes so the cost analysis must account for the worst outcome. There is a choice of defining the costs in the subinstance such that the cost of the lifted subgraph is the same irrespective of the outcome. Such a choice can lead to negative costs, but this is not a hindrance for our inductive step because Theorem 2 allows arbitrary real-valued edge costs. This generality of cost functions is crucial to our algorithm.

In Section 4 we consider a well-studied special case of the 2ECM problem. We present a simple  $O(n^3)$ -time algorithm that given a 3-regular 3-edge connected graph  $G$ , finds a 2-edge connected spanning multisubgraph of cost at most  $\frac{7}{8}c(G)$  (see Theorem 10). The proof is inspired by that of Haddadan, Newman, and Ravi in [11] where they give a polynomial-time algorithm for this problem with a factor  $\frac{15}{17}(> \frac{7}{8})$ . In [10, Theorem 1.1], Haddadan and Newman improve this result to a factor  $\frac{7}{8}$ , and very recently, in [9, Theorem 1.20], Haddadan claims a stronger result with a factor of  $\frac{41}{47} = \frac{7}{8} - \frac{1}{376}$ . We remark that these proofs are longer and/or more complicated than that of Theorem 10. Another motivation for Section 4 is to illustrate the potential of Theorem 2 in giving simpler proofs for results that may not have any explicit half-integrality restrictions.

## 1.2 Related Work

The 2ECM problem has been intensively studied in network design and several works have tried to bound the integrality gap  $\alpha 2\text{ECM}$  of (2ECM-LP). For the general case with metric costs, we have  $\frac{6}{5} \leq \alpha 2\text{ECM} \leq \frac{3}{2}$ , where the lower bound is from [1] and the upper bound follows from the polyhedral analysis of Wolsey [16] and Shmoys and Williamson [15] (this analysis also gives a  $\frac{3}{2}$ -approximation algorithm). It is generally conjectured that  $\alpha 2\text{ECM} = \frac{4}{3}$ , however in [1], Alexander et al., study  $\alpha 2\text{ECM}$  and conjecture that  $\alpha 2\text{ECM} = \frac{6}{5}$  based on their findings. As mentioned before, Carr and Ravi [6] show that the integrality gap of (2ECM-LP) is at most  $\frac{4}{3}$  in the half-integral case. In [5] Boyd and Legault consider a more restrictive collection of instances called half-triangle instances where the optimal LP solution is half-integral and the graph induced by the half-edges is a collection of disjoint triangles. They prove that  $\alpha 2\text{ECM} = \frac{6}{5}$  in this setting. Half-triangle solutions are of interest as there is evidence that the integrality gap of (2ECM-LP) is attained at such solutions (see [1]). When the costs come from a graphic metric (i.e., we want to find a minimum-size 2-edge connected spanning multisubgraph of a given unweighted graph), we have  $\frac{8}{7} \leq \alpha 2\text{ECM} \leq \frac{4}{3}$  (see [3, 14]).

## 2 A Simpler Proof of a Result of Carr and Ravi

In this section, we give a simplified proof of the following result from [6]. As mentioned before, avoiding the case involving the gluing operation is useful for our algorithm in Section 3. For notational convenience, for any subgraph  $K$  of some graph, we use  $\chi^K$  to denote  $\chi^{E(K)}$  whenever the underlying graph is clear from the context.

► **Theorem 3** (Statement 1 from [6]). *Let  $G = (V, E)$  be a 4-regular 4-edge connected multigraph and  $e = uv$  be an arbitrary edge in this graph. There exists a finite collection  $\{H_1, \dots, H_k\}$  of 2-edge connected spanning subgraphs of  $G - e$  such that for some nonnegative  $\mu_1, \dots, \mu_k$  with  $\sum_i \mu_i = 1$ , we have  $\frac{2}{3}\chi^{E \setminus \{e\}} = \sum_{i=1}^k \mu_i \chi^{H_i}$ . Moreover, we may assume that none of the  $H_i$ 's use more than one copy of an edge in  $E$ ;  $H_i$  may have multiedges as long as they come from distinct edges in  $G$ .*

### 2.1 Operations involving splitting-off at a vertex

The following tools on the splitting-off operation will be useful. In keeping with standard terminology, we designate a vertex  $v$  (one of the endpoints of  $e$  in the theorem statement) at which the splitting-off operation is applied. For a multigraph  $H = (V, E)$  and  $x, y \in V$ , let  $\lambda_H(x, y)$  denote the size of a minimum  $(x, y)$ -cut in  $H$ , and let  $\deg_H(x)$  denote the degree of  $x$  in  $H$ . Note that each multiedge is counted separately towards the degree of a vertex and the size of a cut.

► **Definition 4.** *Given a multigraph  $G$  and two edges  $sv$  and  $vt$  that share an endpoint  $v$ , the graph  $G_{s,t}$  obtained by splitting off the pair  $(sv, vt)$  at  $v$  is given by  $G + st - sv - vt$ .*

► **Definition 5.** *Given a multigraph  $G$  and a vertex  $v$  of  $G$  of even degree, a complete splitting at  $v$  is a sequence of  $\frac{1}{2}\deg_G(v)$  splitting off operations that result in vertex  $v$  having degree zero in the resulting graph.*

► **Definition 6.** *Let  $k \geq 2$  be an integer and let  $G$  be a multigraph such that for all  $x, y \in V \setminus \{v\}$ ,  $\lambda_G(x, y) \geq k$ . Let  $e = sv$  and  $vt$  be two edges incident to  $v$ . We say that the pair  $(sv, vt)$  is admissible if for all  $x, y \in V \setminus \{v\}$ ,  $\lambda_{G_{s,t}}(x, y) \geq k$ , and for a particular edge  $e \in \delta(v)$ , we let  $A_e$  denote the set of edges  $f \in \delta(v) \setminus \{e\}$  such that  $(e, f)$  is an admissible pair.*

The following result due to Bang-Jensen et al., [2] shows that in our setting with a 4-regular 4-edge connected multigraph at least two distinct edges incident to  $v$  form an admissible pair with  $e = uv$ . Using this we can perform a complete splitting at  $v$  in two distinct ways.

► **Lemma 7** (Theorem 2.12 from [2]). *Let  $k \geq 2$  be an even integer. Let  $G$  be a multigraph such that for all  $x, y \in V \setminus \{v\}$ ,  $\lambda_G(x, y) \geq k$ . Let  $\deg_G(v)$  be even (each multiedge is counted separately towards the degree). Then,  $|A_{uv}| \geq \frac{1}{2}\deg_G(v)$ .*

► **Lemma 8.** *Let  $G$  be a 4-regular 4-edge connected multigraph and  $e = vx$  be an edge incident to  $v$ . Then, (i)  $|A_e| \geq 2$ ; and (ii) if  $(e, f)$  is an admissible pair for some  $f = vy \in \delta(v) \setminus \{e\}$ , then the remaining two edges in  $\delta(v) \setminus \{e, f\}$  form an admissible pair in  $G_{x,y}$ .*

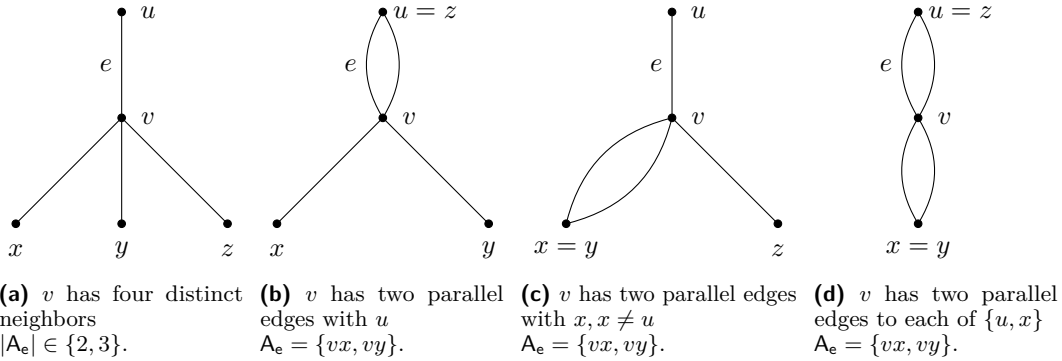
**Proof.** Conclusion (i) follows from Lemma 7 since  $G$  is 4-regular and 4-edge connected. For conclusion (ii), let  $f \in \delta(v) \setminus \{e\}$  be such that  $(e, f)$  forms an admissible pair in  $G$ . Let  $G_{x,y}$  denote the graph obtained by splitting off the pair  $(e = vx, f = vy)$  i.e.,

$G_{x,y} = G - vx - vy + xy$ . Observe that the hypothesis of Lemma 7 still holds for  $G_{x,y}$  with  $k = 4$  because (a) we performed a splitting off operation using an admissible pair of edges; and (b)  $\deg_{G_{x,y}}(v) = 2$  is even. Let  $g$  denote one of the two remaining edges in  $\delta(v) \setminus \{e, f\}$ . By Lemma 7, the other unique edge  $h \in \delta(v) \setminus \{e, f, g\}$  forms an admissible pair with  $g$  in  $G_{x,y}$ . ◀

Equipped with the above tools, we give a proof of Theorem 3.

**Proof of Theorem 3.** Let  $G = (V, E)$  be a 4-regular 4-edge connected multigraph and let  $e = uv$  be an arbitrary edge in  $G$ . We prove this theorem via induction on  $n := |V(G)|$ . The base case  $n = 2$  corresponds to a pair of vertices having four parallel edges, call them  $e, f, g, h$ . Observe that  $\frac{2}{3}\chi^{E \setminus \{e\}} = \frac{1}{3}(\chi^{\{f,g\}} + \chi^{\{f,h\}} + \chi^{\{g,h\}})$ , so the induction hypothesis is true for the base case.

For the induction step, suppose that  $n \geq 3$  and the hypothesis holds for all 4-regular 4-edge connected multigraphs with at most  $n - 1$  vertices and for all choices of  $e \in E$ . Consider a 4-regular 4-edge connected multigraph  $G$  on  $n$  vertices and an arbitrary edge  $e = uv \in E(G)$ . Besides  $e$ , let  $vx, vy, vz$  be the other three edges incident to  $v$ . With a relabeling of vertices, by Lemma 8, we may assume that  $(uv, vx)$  and  $(uv, vy)$  form an admissible pair in  $G$  (see Figure 1).



■ **Figure 1** Four configurations of edges in  $\delta(v) = \{uv, vx, vy, vz\}$  that can arise in our proof.

By the second conclusion of Lemma 8,  $(vy, vz)$  is an admissible pair in  $G_{u,x}$ , and  $(vx, vz)$  is an admissible pair in  $G_{u,y}$ . Consider the graph  $G_1$  obtained by splitting off the pair  $(vy, vz)$  in  $G_{u,x}$  i.e.,  $G_1 = G - v + \{ux, yz\}$ ; it is customary to drop the vertex  $v$  after all its edges have been split off. Similarly, let  $G_2$  be the graph obtained by splitting off the pair  $(vx, vz)$  in  $G_{u,y}$  i.e.,  $G_2 = G - v + \{uy, xz\}$ .

Since we only split off admissible pairs, both  $G_1$  and  $G_2$  are 4-regular 4-edge connected multigraphs on  $n - 1$  vertices. Recall that for any subgraph  $K$  of some graph,  $\chi^K$  is a shorthand for  $\chi^{E(K)}$  whenever the underlying graph is clear from the context. Applying the induction hypothesis to  $G_1$  with the designated edge  $e_1 = ux$  gives:

$$\frac{2}{3} \cdot \chi^{E(G_1) \setminus \{e_1\}} = \frac{2}{3} \cdot \chi^{(E \setminus \delta(v)) \cup \{yz\}} = \sum_{i=1}^{k_1} \mu_i^1 \chi^{H_i^1}, \quad (\text{ConvexComb-}G_1)$$

where  $\{\mu_i^1\}_i$  denote the coefficients in a convex combination, and  $\{H_i^1\}_i$  are 2-edge connected spanning subgraphs of  $G_1$  such that none of them use more than one copy of an edge in  $G_1$ .

Repeating the same argument for  $G_2$  with the designated edge  $e_2 = uy$  gives:

$$\frac{2}{3} \cdot \chi^{E(G_2) \setminus \{e_2\}} = \frac{2}{3} \cdot \chi^{(E \setminus \delta(v)) \cup \{xz\}} = \sum_{i=1}^{k_2} \mu_i^2 \chi^{H_i^2}, \quad (\text{ConvexComb-}G_2)$$

where  $\{\mu_i^2\}_i$  denote the coefficients in the other convex combination arising from  $\{H_i^2\}_i$ . It remains to combine (ConvexComb- $G_1$ ) and (ConvexComb- $G_2$ ) to obtain such a representation for  $G$  with the designated edge  $e$ . We mimic the strategy from [6].

For each  $i \in \{1, \dots, k_1\}$ , we lift  $H_i^1$  to a spanning subgraph  $\hat{H}_i^1$  of  $G - e$ . Define  $\hat{H}_i^1$  as follows:

$$\hat{H}_i^1 := \begin{cases} H_i^1 - yz + vy + vz & \text{if } yz \in E(H_i^1), \\ H_i^1 + vy + vx & \text{if } yz \notin E(H_i^1). \end{cases} \quad (\text{Lift-}G_1)$$

Similarly, for each  $i \in \{1, \dots, k_2\}$ , we define  $\hat{H}_i^2$  as the following spanning subgraph of  $G - e$ :

$$\hat{H}_i^2 := \begin{cases} H_i^2 - xz + vx + vz & \text{if } xz \in E(H_i^2), \\ H_i^2 + vx + vy & \text{if } xz \notin E(H_i^2). \end{cases} \quad (\text{Lift-}G_2)$$

We finish the proof of Theorem 3 by arguing that the following convex combination meets all the requirements:

$$q := \frac{1}{2} \sum_{i=1}^{k_1} \mu_i^1 \chi^{\hat{H}_i^1} + \frac{1}{2} \sum_{i=1}^{k_2} \mu_i^2 \chi^{\hat{H}_i^2}. \quad (\text{ConvexComb-}G)$$

Many of our arguments are the same for  $G_1$  and  $G_2$  so we just mention them in the context of  $G_1$ . First of all, by the induction hypothesis and (Lift- $G_1$ ) it is clear that  $e(= uv), yz, ux \notin E(\hat{H}_i^1)$ , where  $yz$  and  $ux$  refer to the edges that originated from the splitting off operations applied at  $v$ . Next, we argue that  $\hat{H}_i^1$  is a spanning subgraph of  $G$  that uses no more than one copy of any edge in  $G$ . By the induction hypothesis, none of the subgraphs  $H_i^1$  use more than one copy of an edge in  $G_1$ , and  $H_i^1$  spans  $V \setminus \{v\}$ . By the way we lift  $H_i^1$  to  $\hat{H}_i^1$ , it is clear that  $\hat{H}_i^1$  uses no more than one copy of any multiedge in  $G$ , and that it is spanning. To see that  $\hat{H}_i^1$  is 2-edge connected, observe that the two cases of lifting may be viewed as either (i) subdividing the edge  $yz$  by a node  $v$  when  $yz \in E(H_i^1)$ , or (ii) adding an edge  $yx$  and subdividing it by a node  $v$  when  $yz \notin E(H_i^1)$ . Clearly, these operations preserve 2-edge connectivity, hence,  $\hat{H}_i^1$  is 2-edge connected.

It remains to argue that the vector  $q$  in the expression (ConvexComb- $G$ ) matches the vector  $\frac{2}{3} \chi^{E(G) \setminus \{e\}}$ . Since  $\{\mu_i^1\}_i$  and  $\{\mu_i^2\}_i$  denote coefficients in a convex combination, taking an unweighted average of these two combinations gives us another convex combination. Since none of the edges in  $E(G) \setminus \delta(v)$  are modified in the lifting step,  $q_f = 2/3$  for any such edge  $f$ . Next, consider the edge  $vy$ . Observe that  $\hat{H}_i^1$  always contains the edge  $vy$ , whereas  $\hat{H}_i^2$  contains  $vy$  only when  $xz \notin E(H_i^2)$  (this happens with weight  $1/3$ ). Therefore,  $q_{vy} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3}$ . The analysis for  $vx$  is symmetric. Lastly, consider the edge  $vz$ . It appears in  $\hat{H}_i^1$  ( $\hat{H}_i^2$ ) if and only if  $yz \in E(H_i^1)$  (respectively,  $xz \in E(H_i^2)$ ). Therefore,  $q_{vz} = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{3}$ . This completes the proof of Theorem 3. ◀

### 3 Our Algorithm and the Proof of Theorem 2

In this section we give a proof of Theorem 2 which we use to obtain a  $\frac{4}{3}$ -approximation algorithm for 2ECM on half-integral instances (Theorem 1). We apply the same splitting-off theorem of [2] together with an induction scheme that is captured in Theorem 2. A key feature of this theorem is that we allow edges of negative cost, although the edge costs in any instance of 2ECM are non-negative.

Consider a 4-regular 4-edge connected multigraph  $G = (V, E)$  on  $n$  vertices, and let  $e = uv$  be an edge in  $G$ . Let  $c : E \rightarrow \mathbb{R}$  be an arbitrary real-valued cost function. Our goal is to obtain a 2-edge connected spanning subgraph  $H$  of  $G$  whose cost is at most  $\frac{2}{3}c(G - e)$  while ensuring that  $H$  uses no more than one copy of any multiedge in  $G$ . Observe that if we had access to the collection  $\{H_1, \dots, H_k\}$  of 2-edge connected spanning subgraphs from Theorem 3 for some  $k$  that is polynomial in  $|V(G)|$ , then we would be done: for any cost function  $c$ , the cheapest subgraph in this collection (w.r.t. cost  $c$ ) is one such desired subgraph. It is not clear how to efficiently obtain such a collection; a naive algorithm that follows the proof of Theorem 3 does not run in polynomial time.

As alluded to before, for the purposes of obtaining a cheap 2-edge connected subgraph, it suffices to only recurse on one of two subinstances that arise in the proof of Theorem 3. This insight comes from working backwards from (ConvexComb- $G$ ). Since this convex combination for  $G$  is a simple average of convex combinations from two subinstances (see (ConvexComb- $G_1$ ) and (ConvexComb- $G_2$ )), it is judicious to only recurse on the “cheaper” subinstance. Combining (ConvexComb- $G_1$ ) and (Lift- $G_1$ ), we get that the first subinstance gives rise to a convex combination for  $\frac{2}{3}\chi^{E(G)\setminus\{e\}} + \frac{1}{3}(\chi^{\{vy\}} - \chi^{\{vx\}})$ . On the other hand, the second subinstance gives rise to a convex combination for  $\frac{2}{3}\chi^{E(G)\setminus\{e\}} + \frac{1}{3}(\chi^{\{vx\}} - \chi^{\{vy\}})$ . Thus, we should recurse on  $G_1$  if  $c_{vx} \geq c_{vy}$ , and  $G_2$  otherwise. For the sake of argument, suppose that we are recursing on  $G_1$ . So far, we have ignored an important detail in the recursion: the splitting-off operation creates a new edge  $yz$  that was not originally present in  $G$ , so we need to assign it some cost to apply the algorithm recursively. Depending on how we choose the cost of  $yz$ , it might either be included or excluded from the subgraph obtained for the smaller instance, so to bound the cost of the lifted solution we must have a handle on both outcomes of the lift operation. Setting  $c_{yz} := c_{vz} - c_{vx}$  balances the cost of both outcomes. Note that  $c_{yz}$  could possibly be negative, but this is permissible since the statement of Theorem 2 allows for arbitrary edge costs. We formalize the above ideas.

In the recursive step, we pick one end vertex  $v$  of  $e$  and apply a complete splitting off operation at  $v$  to obtain a 4-regular 4-edge connected graph on  $n - 1$  vertices; this can be implemented in  $O(n)$  time. The running time of the algorithm is  $O(n^2)$ , since we apply the induction step  $O(n)$  times.

Let  $T = \{v, x, y, z\}$  be the four neighbors of  $v$  and let  $e = uv$ . Recall that  $A_e$  denotes the set of edges  $f \in \delta(v) \setminus \{e\}$  such that  $(e, f)$  is an admissible pair (see Definition 6).

► **Lemma 9.** *For  $vx \in \delta(v) \setminus \{e\}$ , we can check whether  $vx \in A_e$  in  $O(n)$  time.*

**Proof.** We may suppose that the elements of the set  $T$  of neighbors of  $v$  are all distinct. Otherwise, by Lemma 8, we know exactly which pairs are admissible, see Figure 1. Consider the graph  $\hat{G} = (G_{u,x})_{y,z}$  obtained by splitting off the pairs  $(uv, vx)$  and  $(yv, vz)$  at  $v$ . Let  $G^*$  be the graph obtained from  $\hat{G}$  by contracting  $ux$  to a single vertex  $s$  and contracting  $yz$  to a single vertex  $t$ . Then we apply a max  $s - t$  flow computation to check whether  $G^*$  has  $\geq 4$  edge-disjoint  $s - t$  paths; otherwise,  $G^*$  has an  $(s, t)$ -cut  $\delta(S)$  of size  $\leq 3$ . In the latter case, it is clear that our trial splitting is not admissible.



In the former case, we claim that our trial splitting is admissible. Suppose that  $\hat{G}$  is not 4-edge connected. Then there exists a non-empty, proper vertex set  $S$  in  $\hat{G}$  such that  $|T \cap S| \leq |T \setminus S|$  and  $|\delta_{\hat{G}}(S)| < 4$ . Clearly,  $|S \cap T| \leq 2$ , and if  $|S \cap T| = 2$ , then we have  $|S \cap \{u, x\}| = 1$  and  $|S \cap \{y, z\}| = 1$  (otherwise,  $S$  would give an  $(s, t)$ -cut of  $G^*$  of size  $\leq 3$ ). Since the size of the cut of  $S$  is the same in  $G$  and in  $\hat{G}$ , we have, by 4-edge connectivity of  $G$ ,  $4 > |\delta_{\hat{G}}(S)| = |\delta_G(S)| \geq 4$ , a contradiction.

To see that the running time is linear, observe that  $G^*$  has  $\leq 2n$  edges, an  $s - t$  flow of value  $\geq 4$  can be computed by finding 4 augmenting paths, and each augmenting path can be found in linear time.  $\blacktriangleleft$

**Proof of Theorem 2.** First, consider the base case in the recursion when  $n = 2$ . The only such 4-regular 4-edge connected multigraph is given by four parallel edges between  $u$  and  $v$ , of which  $e$  is one. Picking the two cheapest edges from the remaining three edges gives the desired subgraph.

For the induction step, suppose that  $n \geq 3$  and the induction hypothesis holds for all 4-regular 4-edge connected multigraphs with at most  $n - 1$  vertices and for all choices of edge  $e$ . Consider a 4-regular 4-edge connected multigraph  $G$  on  $n$  vertices and an edge  $e = uv$  in  $G$ .

Our algorithm proceeds as follows. By Lemmas 8 and 9, we can find in  $O(n)$ -time two neighbors of  $v$ , say  $x$  and  $y$ , such that  $vx, vy \in A_e$  and  $c_{vx} \geq c_{vy}$ . Next, we construct the graph  $\hat{G} := (G_{u,x})_{y,z} = G - v + \{ux, yz\}$  and extend the cost function  $c$  to the new edge  $yz$  as  $c_{yz} := c_{vz} - c_{vx}$  (note that the cost of  $ux$  is inconsequential and that  $c_{yz}$  may be negative or non-negative). We recursively find a 2-edge connected spanning subgraph  $\hat{H}$  of  $\hat{G}$  with cost at most  $\frac{2}{3}c(\hat{G} - ux)$ . Then, we lift  $\hat{H}$  to obtain a spanning subgraph  $H$  of  $G$ :

$$H := \begin{cases} \hat{H} - yz + vy + vz & \text{if } yz \in E(\hat{H}), \\ \hat{H} + vy + vx & \text{if } yz \notin E(\hat{H}). \end{cases}$$

We analyze the cost of this subgraph. Regardless of the cases above, our choice of  $c_{yz}$  implies that  $c(H) = c(\hat{H}) + c_{vy} + c_{vx}$ . Therefore,

$$\begin{aligned} c(H) &\leq \frac{2}{3}c(\hat{G} - ux) + c_{vy} + c_{vx} = \frac{2}{3}\{c(G - e) - c_{vx} - c_{vy} - c_{vz} + (c_{vz} - c_{vx})\} + c_{vy} + c_{vx} \\ &= \frac{2}{3}c(G - e) + \frac{1}{3}(c_{vy} - c_{vx}) \leq \frac{2}{3}c(G - e), \end{aligned}$$

where the last inequality follows from our choice of  $vx, vy$  to satisfy  $c_{vx} \geq c_{vy}$ .

It remains to argue that  $H$  is a 2-edge connected spanning subgraph of  $G - e$  that uses no more than one copy of any multiedge in  $G$ . It is clear that the following hold: (a)  $e \notin E(H)$ ; (b)  $H$  is a spanning subgraph of  $G$ ; and (c) each multiedge of  $G$  appears at most once in  $H$ . Since  $\hat{H}$  is 2-edge connected and adding and/or subdividing an edge preserves 2-edge connectivity,  $H$  is 2-edge connected. Overall, in  $O(n^2)$ -time we have constructed a 2-edge connected spanning subgraph  $H$  of  $G - e$  whose cost is at most  $\frac{2}{3}c(G - e)$ , thereby proving Theorem 2.  $\blacktriangleleft$

Using Theorem 2, we give a deterministic  $\frac{4}{3}$ -approximation algorithm for 2ECM on half-integral instances.

**Proof of Theorem 1.** Let  $x$  be an optimal half-integral solution to (Subtour-LP) (and (2ECM-LP)) for an instance given by an  $n$ -vertex graph  $\bar{G} = (\bar{V}, \bar{E})$  and a metric cost function  $c$ . Let  $G = (V, E)$  denote the graph induced by  $2x$  where for each  $e \in \bar{E}$  we include

$2x_e$  copies of the edge  $e$  in  $G$ . Since  $x$  has (fractional) degree 2 at each vertex and it is fractionally 2-edge connected,  $G$  is a 4-regular 4-edge connected multigraph. With a slight abuse of notation, we use the same cost function for the edges of  $E$ : for any  $e \in E$ ,  $c_e := c_f$ , where  $f$  denotes the edge in  $\bar{E}$  that gave rise to  $e$ . We invoke Theorem 2 on  $G$  and some edge  $e \in E$ . This gives us a 2-edge connected spanning subgraph  $H$  of  $G - e$  satisfying  $c(H) \leq \frac{2}{3}c(G - e)$ . Lifting the subgraph  $H$  to  $\bar{G}$  gives a 2-edge connected spanning multisubgraph  $\bar{H}$  (of  $\bar{G}$ ); note that  $\bar{H}$  uses at most two copies of any edge in  $\bar{G}$ . By the first conclusion of Theorem 2 and the non-negativity of  $c$ ,  $c(\bar{H}) = c(H) \leq \frac{2}{3}c(G - e) \leq \frac{2}{3}c(G) = \frac{4}{3}c^T x$ , where the last equality follows by recalling that  $G$  is induced by  $2x$ . Besides invoking Theorem 2 we only perform trivial graph operations so the running time is  $O(n^2)$ . ◀

#### 4 2ECM for 3-Regular 3-Edge Connected Graphs

Let  $G = (V, E)$  be a 3-regular 3-edge connected graph with non-negative edge costs  $c \in \mathbb{R}_{\geq 0}^E$ . In this section we consider an analogous problem to that of Theorem 2, namely the problem of finding a polynomial-time algorithm which gives a 2-edge connected spanning multisubgraph of  $G$  of cost at most  $\beta c(G)$  for some  $\beta \geq 0$ . Note that the everywhere  $\frac{2}{3}$  vector for  $G$  is feasible for (Subtour-LP). For any costs  $c$  for which the everywhere  $\frac{2}{3}$  vector is also optimal for (Subtour-LP) (such as for the graphic metric), such an algorithm would provide a  $\frac{3\beta}{2}$ -approximation for 2ECM. The conjecture that  $\alpha 2\text{ECM} = \frac{4}{3}$  would then imply  $\beta = \frac{4}{3} \cdot \frac{2}{3} = \frac{8}{9}$  should be possible, and the  $\frac{6}{5}$  conjecture for  $\alpha 2\text{ECM}$  would imply  $\beta = \frac{6}{5} \cdot \frac{2}{3} = \frac{4}{5}$  should be possible. In [5] a constructive algorithm for  $\beta = \frac{4}{5}$  is given, however it does not run in polynomial time.

In [11, Theorem 2], Haddadan, Newman, and Ravi show that it is possible to do better than  $\frac{8}{9}$  for this problem, and provide an efficient algorithm for  $\beta = \frac{15}{17}$ . In fact, they show that the everywhere  $\frac{15}{17}$  vector can be expressed as a convex combination of 2-edge connected spanning multisubgraphs of  $G$  and this convex combination can be found in polynomial time. They remark that combining their ideas with an efficient algorithm for Theorem 3 would imply the result for  $\beta = \frac{7}{8} (< \frac{15}{17})$ . Although a polynomial-time algorithm for Theorem 3 is not currently known, it is possible to use our result in Theorem 2 to obtain  $\beta = \frac{7}{8}$ , as follows.

► **Theorem 10.** *Let  $G = (V, E)$  be a 3-regular 3-edge connected graph on  $n$  vertices with non-negative edge costs  $c \in \mathbb{R}_{\geq 0}^E$ . Then in  $O(n^3)$ -time we can find a 2-edge connected spanning multisubgraph  $H$  of  $G$  such that  $c(H) \leq \frac{7}{8}c(G)$ .*

**Proof.** Let  $F$  be a 2-factor of  $G$  that intersects all of the 3-edge cuts and 4-edge cuts of  $G$ . Such a 2-factor can be found in  $O(n^3)$ -time (see [4, Theorem 5.4]). Let  $G'$  be the graph obtained by contracting the cycles of  $F$  and removing any resulting loops, and let  $M := E(G')$ . Clearly  $G'$  is 5-edge connected (by choice of  $F$ ), and thus the vector  $y \in \mathbb{R}_{\geq 0}^M$  defined by  $y_e := \frac{2}{5}$  for all  $e \in M$  is feasible for (2ECM-LP) for  $G'$ . It then follows from the polyhedral analysis of Wolsey [16] and Shmoys and Williamson [15] of the (Subtour-LP) that we can find a 2-edge connected spanning multisubgraph of  $G'$  with edge set  $R$  satisfying  $c(R) \leq \frac{3}{2}c^T y = \frac{3}{5}c(M)$ . Then the graph  $H_1$  induced by  $F \cup R$  is a 2-edge connected spanning multisubgraph of  $G$  such that

$$c(H_1) \leq c(F) + \frac{3}{5}c(M) \leq c(F) + \frac{3}{5}c(E \setminus F). \quad (9)$$

Now consider the vector  $z \in \mathbb{R}_{\geq 0}^E$  where  $z_e := 1/2$  for all  $e \in F$ , and  $z_e := 1$  otherwise. Vector  $z$  is a feasible half-integer solution for (Subtour-LP), and thus by Theorem 2 and the ideas used in the proof of Theorem 1, in  $O(n^2)$ -time we can find a 2-edge connected spanning multisubgraph  $H_2$  of  $G$  such that

$$c(H_2) \leq \frac{2}{3}c(F) + \frac{4}{3}c(E \setminus F). \quad (10)$$

We complete the proof by showing that either  $H_1$  or  $H_2$  has cost at most  $\frac{7}{8}c(G)$ . Using (9) and (10) we have:

$$\min(c(H_1), c(H_2)) \leq \frac{5}{8}c(H_1) + \frac{3}{8}c(H_2) \leq \frac{5}{8}(c(F) + \frac{3}{5}c(E \setminus F)) + \frac{3}{8}(\frac{2}{3}c(F) + \frac{4}{3}c(E \setminus F)) = \frac{7}{8}c(G).$$

◀

---

## References

- 1 Anthony Alexander, Sylvia Boyd, and Paul Elliott-Magwood. On the integrality gap of the 2-edge connected subgraph problem. *Technical Report TR-2006-04, SITE, University of Ottawa, Ottawa, Canada*, 2006. URL: <http://www.site.uottawa.ca/~sylvia/publications/AlexanderBoydElliottmagwood2EC.pdf>.
- 2 Jørgen Bang-Jensen, Harold N. Gabow, Tibor Jordán, and Zoltán Szigeti. Edge-Connectivity Augmentation with Partition Constraints. *SIAM Journal on Discrete Mathematics*, 12(2):160–207, 1999. doi:10.1137/S0895480197324700.
- 3 Sylvia Boyd, Yao Fu, and Yu Sun. A 5/4-approximation for subcubic 2EC using circulations and obliged edges. *Discrete Applied Mathematics*, 209:48–58, 2016. doi:10.1016/j.dam.2015.10.014.
- 4 Sylvia Boyd, Satoru Iwata, and Kenjiro Takazawa. Finding 2-Factors Closer to TSP tours in Cubic Graphs. *SIAM Journal on Discrete Mathematics*, 27(2):918–939, 2013. doi:10.1137/110843514.
- 5 Sylvia Boyd and Philippe Legault. Toward a 6/5 Bound for the Minimum Cost 2-Edge Connected Spanning Subgraph. *SIAM Journal on Discrete Mathematics*, 31(1):632–644, 2017. doi:10.1137/16M1057486.
- 6 Robert Carr and R. Ravi. A new bound for the 2-edge connected subgraph problem. In Robert E. Bixby, E. Andrew Boyd, and Roger Z. Ríos-Mercado, editors, *Integer Programming and Combinatorial Optimization, 6th International IPCO Conference*, volume 1412 of *Lecture Notes in Computer Science*, pages 112–125. Springer, Berlin, 1998. doi:10.1007/3-540-69346-7\_9.
- 7 András Frank. *Connections in Combinatorial Optimization*. Oxford Lecture Series in Mathematics and its Applications, Volume 38. Oxford University Press, 2011.
- 8 Michel X Goemans and Dimitris J Bertsimas. Survivable networks, linear programming relaxations and the parsimonious property. *Mathematical Programming*, 60:145–166, 1993. doi:10.1007/BF01580607.
- 9 Arash Haddadan. *New Bounds on Integrality Gaps by Constructing Convex Combinations*. PhD Thesis. Carnegie Mellon University, 2020. URL: <https://www.cmu.edu/tepper/programs/phd/program/assets/dissertations/2020-joint-programs-algorithms-combinatorics-and-optimization-haddadan-dissertation.pdf>.
- 10 Arash Haddadan and Alantha Newman. Efficient constructions of convex combinations for 2-edge-connected subgraphs on fundamental classes. *CoRR*, abs/1811.09906, 2018. arXiv:1811.09906.
- 11 Arash Haddadan, Alantha Newman, and R. Ravi. Shorter tours and longer detours: uniform covers and a bit beyond. *Mathematical Programming*, 2019. doi:10.1007/s10107-019-01426-8.

- 12 Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. An improved approximation algorithm for TSP in the half integral case. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 28–39. ACM, 2020. doi:10.1145/3357713.3384273.
- 13 Frans Schalekamp, David P. Williamson, and Anke van Zuylen. 2-Matchings, the Traveling Salesman Problem, and the Subtour LP: A Proof of the Boyd-Carr Conjecture. *Mathematics of Operations Research*, 39(2):403–417, 2014. doi:10.1287/moor.2013.0608.
- 14 András Sebő and Jens 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. *Combinatorica*, 34(5):597–629, 2014. doi:10.1007/s00493-014-2960-3.
- 15 David B. Shmoys and David P. Williamson. Analyzing the Held-Karp TSP bound: a monotonicity property with application. *Information Processing Letters*, 35(6):281–285, 1990. doi:10.1016/0020-0190(90)90028-V.
- 16 Laurence A. Wolsey. Heuristic analysis, linear programming and branch and bound. *Mathematical Programming Study* 13, pages 121–134, 1980. doi:10.1007/BFb0120913.