Available online at www.sciencedirect.com

ScienceDirect

Discrete Applied Mathematics 00 (2014) 1–13

Discrete Applied Math

A $\frac{5}{4}$-approximation for subcubic 2EC using circulations and obliged edges[1]

Sylvia Boyd, Yao Fu, Yu Sun

School of Electric Engineering and Computer Science (EECS), University of Ottawa
Ottawa, Ontario K1N 6N5, Canada
{sboyd,yfu099,ysun007}@uottawa.ca

Abstract

In this paper we study the NP-hard problem of finding a minimum size 2-edge-connected spanning subgraph (henceforth 2EC) in cubic and subcubic multigraphs. We present a new $\frac{5}{4}$-approximation algorithm for 2EC for subcubic bridgeless multigraphs, improving upon the current best approximation ratio of $\frac{5}{4}+\varepsilon$. Our algorithm involves an elegant new method based on circulations which we feel has potential to be more broadly applied. We also study the closely related integrality gap problem, i.e. the worst case ratio between the integer linear program for 2EC and its linear programming relaxation, both theoretically and computationally. We show this gap is at most $\frac{5}{4}$ for subcubic bridgeless multigraphs, and is at most $\frac{9}{8}$ for all subcubic bridgeless graphs with up to 16 nodes. Moreover, we present a family of graphs that demonstrate the integrality gap for 2EC is at least $\frac{8}{7}$, even when restricted to subcubic bridgeless graphs. This represents an improvement over the previous best known bound of $\frac{9}{8}$.

Keywords:
minimum 2-edge-connected subgraph problem, approximation algorithm, circulations, integrality gap, subcubic graphs

1. Introduction

Given an unweighted bridgeless multigraph $G = (V, E)$, $|V| = n$, the *minimum size 2-edge-connected spanning subgraph problem* (henceforth $2EC(G)$) consists of finding a 2-edge-connected spanning subgraph H of G with the minimum number of edges. Note that a *2-edge-connected* (or *bridgeless*) graph $G = (V, E)$ is one that remains connected after the removal of any edge. In a solution for $2EC(G)$, multiple copies of an edge $e \in E$ are not allowed (and also not necessary).

The problem $2EC(G)$ is one of the most extensively studied problems in network design. It relates to the optimal design of a network that can survive the loss of a link, and thus has many real world applications. However, it is known to be NP-hard and also MAX SNP-hard even for subcubic graphs [1], where a graph is cubic if every node has degree 3, and subcubic if every node has degree at most 3. Thus research has focused on finding good approximation algorithms. Unfortunately, finding improved approximation algorithms seems to be difficult for the more general weighted version of $2EC(G)$ where, as with the closely related travelling salesman problem (TSP), the best known approximation ratio for metric weights has remained at $\frac{3}{2}$ [2] without any improvement for over 30 years.

Given the difficulty of this problem, people have turned to the study of approximation algorithms for special cases, which has proven to be a more successful approach for $2EC(G)$ than studying its more general weighted form. In such

[1]This research was partially supported by the Natural Science and Engineering Research Council of Canada (NSERC).

studies, not only improved results were obtained but also new innovative methods which may lead to more general results.

In this paper we focus on the simplest form of $2EC(G)$ that still remains NP-hard, i.e. $2EC(G)$ for subcubic graphs. In Section 2 we describe the framework for a new innovative method for designing approximation algorithms for $2EC(G)$ based on circulations and the concept of obliged edges, i.e. edges that must be included in the final solution. Similar types of circulations were used in [3] in the approximation of graph TSP, however the goal was quite different in that context. In fact, to the best of our knowledge, circulations have not previously been used in the way we describe to approximate $2EC(G)$. In Section 3 we develops a new heuristic method for the problem $2EC(G)$ for bridgeless cubic graphs with obliged edges. In Section 4 we demonstrate the usefulness of this method by using it to develop a new $\frac{5}{4}$-approximation algorithm for $2EC(G)$ on bridgeless subcubic multigraphs. This algorithm improves upon the previous best approximation ratio of $\frac{5}{4} + \varepsilon$ given by Csaba, Karpinski and Krysta [1] for $2EC(G)$ on such graphs. We feel this algorithm not only provides a modest improvement in the approximation ratio, but also, and perhaps more importantly, provides an improvement in the simplicity and elegance of the method and proof.

A related approach for finding approximated $2EC(G)$ solutions is to study the integrality gap $\alpha^{2EC(G)}$, which is the worst case ratio between the optimal value for $2EC(G)$ and the optimal value for its linear programming (henceforth LP) relaxation (see [2] for background). As a critical topic throughout this paper, we study $\alpha^{2EC(G)}$ intensively. There are two main reasons this is useful. First, the integrality gap itself serves as an indicator of the quality of the lower bound given by the LP relaxation. This is important for methods, such as branch and bound and approximation, that depend on good lower bounds for their success. Secondly, an algorithmic proof for $\alpha^{2EC(G)} = k$ yields a k-approximation algorithm for $2EC(G)$ [2]. In this paper, we give an upper bound on the value of $\alpha^{2EC(G)}$ on bridgeless cubic graphs with an algorithmic proof, while lower bounds on the integrality gap of $2EC(G)$ are investigated through computational studies. In Section 4, we show that the integrality gap of $2EC(G)$ is strictly less than $\frac{5}{4}$ for bridgeless subcubic multigraphs, improving on the previous best known bound of $\frac{5}{4}+\varepsilon$ [1]. In Section 5 we describe a computational study we conducted by designing a program that calculates $\alpha^{2EC(G)}$ exactly for all simple graphs $G \in \mathcal{G}$, where $\mathcal{G}$ contains all test cases in three categories: (1) General bridgeless graphs for $3 \leq n \leq 10$; (2) Cubic bridgeless graphs for $6 \leq n \leq 16$; (3) Subcubic bridgeless graphs for $3 \leq n \leq 16$. Using the knowledge gained through the data analysis for the computational study, we obtain a family of subcubic bridgeless graphs G which shows $\alpha^{2EC(G)} \geq \frac{8}{7}$ asymptotically, providing an improvement upon the previous best lower bound of $\frac{9}{8}$ [4].

1.1. Literature review on $2EC(G)$

Constant factor approximation algorithms for $2EC(G)$ have been intensely studied. In 1994, Khuller and Vishkin [5] found a $\frac{3}{2}$-approximation, which was improved by Cheriyan, Sebő and Szigeti [6] to $\frac{17}{12}$. The ratio was later improved to $\frac{4}{3}$ in 2000 by Vempala and Vetta [7]. One year later, Krysta and Kumar [8] improved the approximation ratio to $\frac{4}{3} - \varepsilon$ where $\varepsilon = \frac{1}{1344}$. Recently, Sebő and Vygen [4] designed a simpler and more elegant $\frac{4}{3}$-approximation algorithm for $2EC(G)$.

In the meantime, research on $2EC(G)$ has also been conducted for special classes of graphs, especially on cubic and subcubic bridgeless graphs, on which $2EC(G)$ still remains NP-hard [1] . In 2001, along with their $(\frac{4}{3} - \varepsilon)$-approximation algorithm for $2EC(G)$ on general graphs, Krysta and Kumar [8] also presented an approximation algorithm for $2EC(G)$ on cubic graphs with the approximation ratio of $\frac{21}{16} + \varepsilon$. One year later, Csaba, Karpinski and Krysta [1] designed a $(\frac{5}{4} + \varepsilon)$-approximation algorithm for $2EC(G)$ on subcubic bridgeless graphs. In 2004, Huh [9] presented an algorithm yielding a $\frac{5}{4}$-approximation on cubic 3-edge-connected graphs. A more recent improvement came from Boyd, Iwata and Takazawa [10] with a $\frac{6}{5}$-approximation algorithm for $2EC(G)$ on cubic 3-edge-connected graphs.

Concerning the integrality gap $\alpha^{2EC(G)}$ of $2EC(G)$, Csaba, Karpinski and Krysta [1] proved that for maximum degree 3 graphs, the integrality gap of the LP relaxation for $2EC(G)$ is at most $\frac{5}{4} + \varepsilon$ for any fixed $\varepsilon > 0$. It was also stated in [1] that the best known lower bound on $\alpha^{2EC(G)}$ is $\frac{10}{9}$ for maximum degree 3 graphs (and thus subcubic graphs). In 2013, Boyd, Iwata and Takazawa [10] showed $\alpha^{2EC(G)} \leq \frac{6}{5}$ for 3-edge-connected cubic graphs. Around the same time, Sebő and Vygen [4] proved that $\frac{9}{8} \leq \alpha^{2EC(G)} \leq \frac{4}{3}$ in general.

1.2. Notation and background

For the purpose of this paper, any graph $G = (V, E)$ is considered to be a multigraph without loops. We use n to denote $|V|$, and sometimes use $V(G)$ to denote V and $E(G)$ to denote E. For any $S \subset V$, $\delta(S)$ is the set of edges with one end in S and the other end not in S, and for any $F \subseteq E$, the notation $x(F)$ is used to denote $\sum_{e \in F} x_e$.

Denoted by $ILP(G)$, the integer linear program of $2EC(G)$ for a graph $G = (V, E)$ is given as follows:

$$\begin{array}{lllll} \text{Minimize} & \sum_{e \in E} x_e & & & (1) \\ \text{subject to} & x(\delta(S)) \geq 2 & \text{for all} & \emptyset \subset S \subset V, & (2) \\ & 0 \leq x_e \leq 1 & \text{for all} & e \in E, & (3) \\ & x_e \text{ integer} & \text{for all} & e \in E. & (4) \end{array}$$

.

By relaxing the integrality constraints (4) of $ILP(G)$, the LP relaxation of $ILP(G)$, denoted by $LP(G)$, is obtained. We use the notation $OPT(G)$ and $OPT_{LP}(G)$ to denote the optimal objective value for $ILP(G)$ and $LP(G)$ respectively.

1.3. Problem $2EC(G)$ with obliged edges

In order to enhance the clarity and usefulness of our results for $2EC(G)$, we will consider a more general form of the problem which includes a specification of a subset of edges that are obliged to be in the final $2EC(G)$ solution. More specifically, given an unweighted bridgeless multigraph $G = (V, E)$ and set $\Omega \subseteq E$ of obliged edges, the problem *2EC with obliged edges* (henceforth $2EC(G, \Omega)$) consists of finding a 2-edge-connected spanning subgraph $H = (V, J)$ of G with the minimum number of edges, and such that $\Omega \subseteq J$. Note that $2EC(G) = 2EC(G, \Omega)$ when $\Omega = \emptyset$. We obtain the ILP for $2EC(G, \Omega)$ by adding the following constraints to $ILP(G)$:

$$x_e = 1 \text{ for all } e \in \Omega.$$

We denote this ILP by $ILP(G, \Omega)$ and denote its corresponding LP relaxation by $LP(G, \Omega)$. We use the notation $OPT(G, \Omega)$ and $OPT_{LP}(G, \Omega)$ to denote the optimal objective values of $ILP(G, \Omega)$ and $LP(G, \Omega)$ respectively.

2. Using circulations to obtain 2-edge-connected spanning subgraphs with obliged edges

In this section we outline a new method for finding solutions for the more general problem $2EC(G, \Omega)$ which is based on circulations and depth first search (DFS) trees. In later sections we use these ideas to obtain a $\frac{5}{4}$-approximation for $2EC(G)$ for subcubic bridgeless graphs. Note that similar ideas were used in [3] for approximation for graph TSP, but not in the same way or for the same purpose as they are used here. To the best of our knowledge, these ideas represent a new framework for $2EC(G)$ approximation.

Given a digraph $D = (V, A)$, $f \in \mathbb{R}^A$ is called a *circulation* for D if $f(\delta^{in}(v)) = f(\delta^{out}(v))$ for all $v \in V$. For an arc $e \in A$, f_e is called the *flow* of e. Given arc demands $d \in \mathbb{R}^A$ and arc capacities $u \in \mathbb{R}^A$, a circulation is called *feasible* if $d_e \leq f_e \leq u_e$ for all $e \in A$. The *support graph* of a circulation f is the subgraph $D_f = (V, A_f)$, where A_f is the set of arcs $a \in A$ for which $f_a > 0$. Finally, given arc costs $c \in \mathbb{R}^A$, the *minimum cost circulation problem* is as follows:

$$\begin{array}{lll} \text{minimize} & cf & \\ \text{subject to} & f(\delta^{in}(v)) = f(\delta^{out}(v)) & \text{for all } v \in V, \\ & d_e \leq f_e \leq u_e & \text{for all } e \in A. \end{array}$$

The following is well known for circulations ([11], see also [12]).

Theorem 1. *Given a minimum cost circulation problem for which d and u are integer-valued and for which there exists a feasible circulation f, there exists an optimal circulation f^* which is integer-valued and can be found in polynomial time.*

Given a 2-edge-connected multigraph $G = (V, E)$ and a set of obliged edges $\Omega \subseteq E$, we now define a minimum cost circulation problem $P(G, \Omega)$ based on G and Ω. To begin, give G an orientation by growing a spanning tree from an arbitrary root $r \in V$ using DFS. Call the edges in the tree *tree edges* and the rest of the edges in E *back edges*. Let the directed graph $D = (V, A)$ be the orientation of G obtained in the usual way using the DFS tree, i.e. by directing all the tree edges away from r and all the back edges towards r. Let $A = T \cup B$ where T is the set of directed tree edges and B is the set of directed back edges. Note that the arc set T forms a spanning arborescence of D, and that the edges uv of G are in one to one correspondence with the arcs (u, v) of D, a fact that we exploit by referring to these edges and arcs interchangeably.

Define the minimum cost circulation problem $P(G, \Omega)$ as follows:

$$d_e = \begin{cases} 1 \text{ for } e \in T \cup \Omega \\ 0 \text{ otherwise,} \end{cases} \qquad c_e = \begin{cases} 1 \text{ for } e \in B \\ 0 \text{ otherwise,} \end{cases} \qquad u_e = \begin{cases} 1 \text{ for } e \in B \\ \infty \text{ otherwise.} \end{cases}$$

Let f be any feasible circulation for $P(G, \Omega)$. By Theorem 1, there exists an integer feasible circulation f^* of cost at most cf. The support graph $D_{f^*} = (V, A_{f^*})$ of f^* will consist of the edges of T plus the edges of $e \in B$ with $f^*_e = 1$. Thus $|A_{f^*}| \leq (n-1) + cf$. Moreover, the edges in G corresponding to A_{f^*} include the edges in Ω and form a 2-edge-connected spanning subgraph $H = (V, J)$ of G. To see this, consider any cut $\delta(S)$ in G. Clearly $\delta(S)$ contains at least one edge of T. If it contains two edges of T, then both of these edges are in J as well. If it contains only one edge h of T, then the demand of one for h ensures that at least one edge $e \in B \cap \delta(S)$ has $f^*_e = 1$, so $\delta(S)$ contains at least 2 edges in J, namely h and e.

Given the circulation problem $P(G, \Omega)$ and a lower bound β for $OPT_{LP}(G, \Omega)$, the above suggests the following scheme for finding a k-approximation for $2EC(G, \Omega)$:

(1) Show there exists a feasible (perhaps fractional) circulation f for $P(G, \Omega)$ such that $(n-1) + cf \leq k\beta$.

(2) Find an optimal integer circulation f^* for $P(G, \Omega)$. The support graph of f^* provides a 2-edge-connected spanning subgraph $H = (V, J)$ of G with $\Omega \subseteq J$ and $|J| = (n-1) + cf^* \leq (n-1) + cf \leq k\beta$. Since $\beta \leq OPT_{LP}(G, \Omega)$, $OPT_{LP}(G, \Omega) \leq OPT(G, \Omega)$ and f^* can be found in polynomial time, we have a k-approximation algorithm for $2EC(G, \Omega)$.

The above ideas are illustrated in Figure 1. In Figure 1a) a graph $G = (V, E)$ and edge set Ω are shown, with the edges in Ω denoted in bold. In Figure 1b) a corresponding digraph $D = (V, A)$ obtained by DFS is shown, along with the edges in T (denoted by solid lines) and the edges in B (denoted by dashed lines). In Figure 1c) the values on the arcs represent a feasible circulation f for the circulation problem $P(G, \Omega)$. Note that for this circulation, $cf = 4$. In Figure 1d) an optimal integer circulation f^* is shown for which $cf^* = 3$. Finally, in Figure 1e), the 2-edge-connected subgraph $H = (V, J)$ obtained from f^* is shown, where $|J| = (n-1) + cf^* = 9+3 = 13$.

Next we provide a very useful lower bound for $OPT_{LP}(G, \Omega)$ to be used in the above framework.

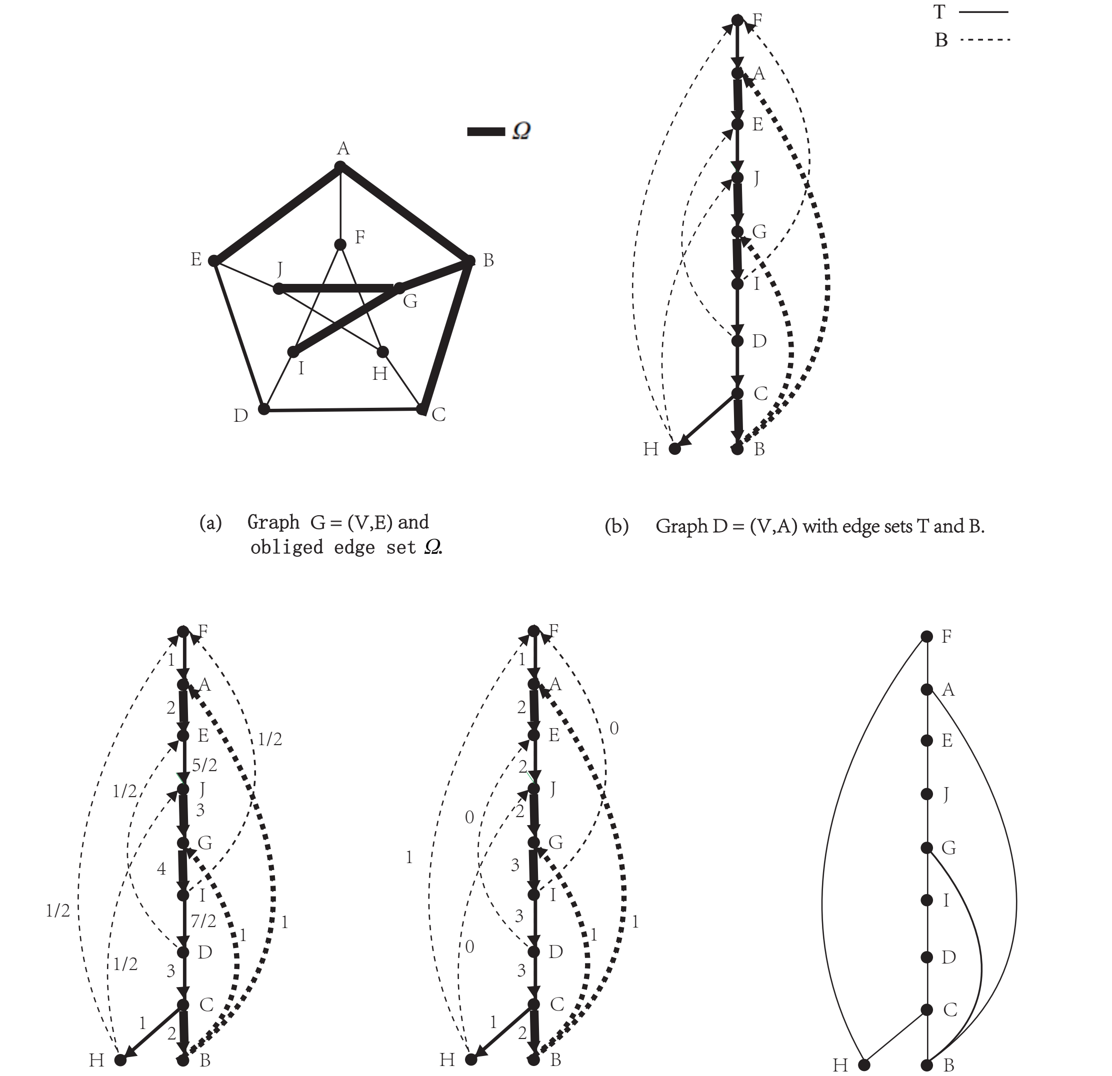
Lemma 1. *Let graph $G = (V, E)$ be a 2-edge-connected multigraph, and let F represent the set of edges in G which are in 2-edge cuts. Then we have the following lower bound for $OPT_{LP}(G, \Omega)$:*

$$OPT_{LP}(G, \Omega) \geq \frac{1}{2} \sum_{v \in V} max(2, |\delta(v) \cap (F \cup \Omega)|).$$

Proof. Let x' be any feasible solution for $LP(G, \Omega)$. For any edge in $F \cup \Omega$ we must have $x'_e = 1$, thus for any node $v \in V$, $x'(\delta(v)) \geq max(2, |\delta(v) \cap (F \cup \Omega)|)$. Since $\sum_{e \in E} x'_e = \frac{1}{2} \sum_{v \in V} x'(\delta(v))$, the result follows. □

3. An algorithm for $2EC(G, \Omega)$ for cubic bridgeless graphs

In this section we use the ideas presented in the previous section to provide a heuristic algorithm for $2EC(G, \Omega)$ for cubic bridgeless multigraphs and also provide a performance guarantee for the solutions provided by the algorithm. In the next section we will show how to use this algorithm to obtain a $\frac{5}{4}$-approximation algorithm for $2EC(G)$ for subcubic bridgeless multigraphs. We begin with some preliminaries.



(a) Graph G = (V,E) and obliged edge set Ω.

(b) Graph D = (V,A) with edge sets T and B.

(c) Feasible circulation f for P(G,Ω), cf = 4.

(d) Optimal integer circulation f* for P(G,Ω), cf* = 3.

(e) Final 2-edge-connected subgraph H=(V,J), |J|=(n-1)+cf*.

Figure 1. Illustration of the steps used to find the 2-edge-connected subgraph $H = (V, J)$.

Given a graph $G = (V, E)$, a cut $\delta(S)$ for $S \subset V$ is called *proper* if $2 \leq |S| \leq n-2$. For $S \subset V$, let $\overline{S} = V \setminus S$. Given $S \subseteq V$, let $G[S]$ be the graph with node set S and edge set $\{uv \in E : u \in S, v \in S\}$.

For the remainder of this section let $G = (V, E)$ be a cubic bridgeless multigraph, let $\Omega \subseteq E$ and, as before, let F be the set of edges in E which are in 2-edge cuts. Let $V^*(G) = \{v \in V : |\delta(v) \cap (F \cup \Omega)| = 3\}$. By Lemma 1 we have

$$OPT_{LP}(G, \Omega) \geq n + \frac{|V^*(G)|}{2}.$$

We now construct the minimum cost circulation problem $P(G, \Omega)$ as described in Section 2, using the same notation. Recall T is the set of tree edges formed by the DFS and B is the set of back edges. For each edge $e \in T$, let $\delta_{T_e} \subset B$ be the set of back edges in the unique cut in D for which e is the only tree edge in the cut. Each edge b in δ_{T_e} forms a unique directed cycle C_b in $T \cup \{b\}$ which contains tree edge e. Let $f' \in \mathbb{R}^A$ be the circulation for $P(G, \Omega)$ obtained by setting the flow to $\frac{1}{2}$ around each cycle C_b for $b \in B \setminus (F \cup \Omega)$ and to 1 around each cycle C_b for $b \in B \cap (F \cup \Omega)$, and then summing these cycle flows. More specifically, f' is defined as follows:

$$f'_a = \begin{cases} \frac{1}{2} \text{ for } a \in B \setminus (F \cup \Omega), \\ 1 \text{ for } a \in B \cap (F \cup \Omega), \\ \sum(f'_b : b \in \delta_{T_a}) \text{ for } a \in T. \end{cases}$$

Note that the feasible circulation f shown in Figure 1c) gives an illustration of the circulation f' described above.

Lemma 2. *The circulation f' is a feasible circulation for $P(G, \Omega)$ and has cost $cf' = \frac{n}{4} + \frac{|B \cap (F \cup \Omega)|}{2} + \frac{1}{2}$.*

Proof. Since every tree edge e not in F must be in the cycle C_e for at least 2 back edges, f' is clearly a feasible circulation for $P(G, \Omega)$. Moreover, since $c_a = 0$ for $a \in T$ and $c_a = 1$ for $a \in B$, $cf' = \frac{1}{2}|B \setminus (F \cup \Omega)| + |B \cap (F \cup \Omega)| = \frac{|B|}{2} + \frac{|B \cap (F \cup \Omega)|}{2}$. Since $|E| = \frac{3n}{2}$ for a cubic graph, we have $|B| = |E| - |T| = \frac{3n}{2} - (n-1) = \frac{n}{2} + 1$ and the result follows. □

We now describe our recursive heuristic algorithm for $2EC(G, \Omega)$, which is based on the ideas from Section 2 along with a careful specification of how we grow the DFS tree.

Algorithm 2EC(G, Ω):
Input: A cubic bridgeless multigraph $G = (V, E)$ and obliged edge set $\Omega \subseteq E$.
Output: A 2-edge-connected spanning subgraph $H = (V, J)$ of G with $\Omega \subseteq J$.

Case (a): {Base Case} G is 3-edge-connected (i.e. $F = \emptyset$).

1. Grow a DFS tree T in G. We grow the DFS tree according to the following two rules:

   Rule 1 Start growing T from a node $r \notin V^*(G)$, if possible.

   Rule 2 If we have a choice for the next edge to add to the tree, we always add one in Ω, if possible.

2. Construct the minimum cost circulation problem $P(G, \Omega)$ using the DFS tree T above, and find the optimal solution f^* for $P(G, \Omega)$ which is integer-valued. Let J be the edges in G corresponding to the edges of the support graph of f^*. Return graph $H = (V, J)$.

Case (b): {Recursive Step} There is a 2-edge cut $\delta(S) = \{e, e'\}$ in G.
Let $e = uv$ and $e' = u'v'$, with $u \in S$, $u' \in S$, $v \in \overline{S}$ and $v' \in \overline{S}$. Since G is cubic and bridgeless, it follows that $|S| \geq 2$ and $|\overline{S}| \geq 2$, and thus $\delta(S)$ is a proper cut. We split the problem into two new smaller problems $2EC(G_1, \Omega_1)$ and $2EC(G_2, \Omega_2)$ defined as follows. Let $G_1 = (S, E_1)$ be the graph obtained by taking $G[S] \cup \{uu'\}$, and let $\Omega_1 = (\Omega \cap E_1) \cup \{uu'\}$. Similarly, let $G_2 = (\overline{S}, E_2)$ be the graph obtained by taking $G[\overline{S}] \cup \{vv'\}$, and let $\Omega_2 = (\Omega \cap E_2) \cup \{vv'\}$. Apply Algorithm $2EC(G, \Omega)$ recursively to G_1 and G_2 to obtain graphs $H_1 = (S, J_1)$ and $H_2 = (\overline{S}, J_2)$ respectively. Combine these graphs into a graph $H = (V, J)$ by removing the edges uu' and vv' and adding back the edges e and e'. Return graph $H = (V, J)$.

Theorem 2. *The graph $H = (V, J)$ from Algorithm $2EC(G, \Omega)$ is a 2-edge-connected spanning subgraph of G such that $\Omega \subseteq J$ and $|J| \leq \frac{5}{4}n + \frac{|V^*(G)|}{8} + \frac{|\Omega|}{4} - \frac{1}{2}$.*

Proof.
Case (a): {Base Case} G is 3-edge-connected (i.e. $F = \emptyset$).
Since G is cubic and T is a DFS tree, each node $v \in V, v \neq r$ has at most one edge in B directed into it. Moreover, since G is also 3-edge-connected, r has exactly two edges in B directed into it.

Now consider $e = (u, v) \in B \cap \Omega$. If $v \neq r$, or if $v = r$ and $r \notin V^*(G)$, then e is the unique arc in $B \cap \Omega$ directed into v, and there is a unique arc $h \in T$ directed out of v. By Rule 2 we must have $h \in \Omega$. From this we can conclude that if $r \notin V^*(G)$, every arc in $B \cap \Omega$ can be paired with a unique arc in $T \cap \Omega$, and thus

$$|B \cap \Omega| \leq \frac{|\Omega|}{2}. \tag{5}$$

If $r \in V^*(G)$, then there are 2 arcs in $B \cap \Omega$ incident with r and one arc in $T \cap \Omega$ incident with r, hence every arc in $B \cap \Omega$ except one can be paired with a unique arc in $T \cap \Omega$. Thus

$$|B \cap \Omega| \leq \frac{(|\Omega| - 1)}{2} + 1 = \frac{|\Omega|}{2} + \frac{1}{2}. \tag{6}$$

We also know by Rule 1 that if $r \in V^*(G)$, then $V^*(G) = V$. Since $|V| \geq 2$ for G, it follows that if $r \in V^*(G)$ then

$$\frac{1}{2} \leq \frac{|V^*(G)|}{4}. \tag{7}$$

Combining (5), (6) and (7), we obtain, for all cases,

$$|B \cap \Omega| \leq \frac{|\Omega|}{2} + \frac{|V^*(G)|}{4}. \tag{8}$$

Now consider our final solution $H = (V, J)$. We know $|J| = (n-1) + cf^*$. Thus by Theorem 1 and Lemma 2 it follows that

$$\begin{aligned} |J| &\leq (n-1) + \frac{n}{4} + \frac{|B \cap \Omega|}{2} + \frac{1}{2} \\ &= \frac{5n}{4} + \frac{|B \cap \Omega|}{2} - \frac{1}{2}. \end{aligned}$$

Applying (8), we obtain

$$|J| \leq \frac{5n}{4} + \frac{|\Omega|}{4} + \frac{|V^*(G)|}{8} - \frac{1}{2},$$

as required. This completes the proof of Case (a).

Case (b): { Recursive Step }: There is a proper 2-edge cut $\delta(S) = \{e, e'\}$.
Consider the two new graphs $G_1 = (S, E_1)$ with obliged edges Ω_1 and $G_2 = (\overline{S}, E_2)$ with obliged edges Ω_2. As G is cubic and bridgeless, it follows that $|S| \geq 2, |\overline{S}| \geq 2$, and the nodes u, v, u' and v' are all distinct. Moreover, G_1 and G_2 are also cubic bridgeless graphs with fewer nodes than G, thus we can apply Algorithm $2EC(G, \Omega)$ recursively to them to obtain 2-edge connected spanning subgraphs $H_1 = (S, J_1)$ and $H_2 = (\overline{S}, J_2)$ respectively such that $\Omega_1 \subseteq J_1, \Omega_2 \subseteq J_2$, and

$$|J_1| \leq \frac{5|S|}{4} + \frac{|V^*(G_1)|}{8} + \frac{|\Omega_1|}{4} - \frac{1}{2} \text{ and } |J_2| \leq \frac{5|\overline{S}|}{4} + \frac{|V^*(G_2)|}{8} + \frac{|\Omega_2|}{4} - \frac{1}{2}.$$

Now consider the subgraph $H = (V, J)$ for G we obtain by removing $uu' \in \Omega_1$ from H_1, $vv' \in \Omega_2$ from H_2, and adding the edges e and e'. Graph H is a 2-edge connected spanning subgraph of G. Moreover, $n = |S| + |\overline{S}|$, $|J| = |J_1| + |J_2|$, $|V^*(G)| = |V^*(G_1)| + |V^*(G_2)|$ and $|\Omega_1| + |\Omega_2| \leq |\Omega| + 2$. Hence

$$|J| = |J_1| + |J_2| \leq \frac{5}{4}n + \frac{|V^*(G)|}{8} + \frac{|\Omega|}{4} - \frac{1}{2},$$

as required. □

4. A $\frac{5}{4}$-approximation algorithm for $2EC(G)$ for subcubic bridgeless graphs

In this section we use the results presented in the previous section to provide a $\frac{5}{4}$-approximation algorithm for $2EC(G)$ for subcubic bridgeless multigraphs and also show that $\alpha^{2EC(G)} < \frac{5}{4}$ for such graphs.

For the remainder of this section, let $G = (V, E)$, $|V| = n$, be a subcubic bridgeless 2-edge-connected multigraph, and as before let F be the set of edges in E which are in 2-edge cuts. Note that $\Omega = \emptyset$ here. By Lemma 1 we have

$$OPT_{LP}(G) \geq n + \frac{|V^*(G)|}{2}, \tag{9}$$

where $V^*(G) \subseteq V$ is the set of nodes in G incident with three edges in F.

We now describe our approximation algorithm for $2EC(G)$, which is based on the results from Section 3 along with the replacement of paths in G by obliged edges. Note that we assume that we have at least one node of degree three in G, as otherwise the problem is trivial.

Algorithm Approx 5_4:
Input: A bridgeless subcubic multigraph $G = (V, E)$.
Output: A 2-edge-connected spaning subgraph $H = (V, J)$ of G.

1. Graph G can be viewed as a cubic graph $G' = (V', E')$ with some edges subdivided into paths of length 2 or more. Let P be the set of such paths, let $P' \subseteq E'$ be the edges in G' corresponding to such paths, and let F' be the set of edges in G' in 2-edge cuts. We then consider the problem $2EC(G', \Omega')$ for graph $G' = (V', E')$ with $\Omega' = P'$.

2. Apply Algorithm $2EC(G, \Omega')$ to graph G' and obliged edge set Ω'. Let the resulting 2-edge-connected subgraph be $H' = (V', J')$. Note that $P' \subseteq J'$.

3. Obtain the subgraph $H = (V, J)$ for G from H' by replacing each edge $e \in P'$ in H' by its corresponding path in P.

4. Output $H = (V, J)$.

Theorem 3. *The graph $H = (V, J)$ from Algorithm Approx 5_4 is a 2-edge-connected spanning subgraph of G such that $|J| \leq \frac{5}{4} OPT_{LP}(G) - \frac{1}{2}$.*

Proof. Clearly H is 2-edge-connected subgraph of G. By Theorem 2,

$$|J'| \leq \frac{5|V'|}{4} + \frac{|V^*(G')|}{8} + \frac{|\Omega'|}{4} - \frac{1}{2}. \tag{10}$$

Let $V_2 \subset V$ be the set of nodes in G of degree 2. We have $|J| = |J'| + |V_2|$, $n = |V'| + |V_2|$ and $V^*(G') = V^*(G)$. Substituting these into (10) gives

$$\begin{aligned} |J| &\leq |V_2| + \frac{5}{4}(n - |V_2|) + \frac{|V^*(G)|}{8} + \frac{|\Omega'|}{4} - \frac{1}{2} \\ &= \frac{5}{4}n + \frac{|V^*(G)|}{8} - \frac{1}{2} + \frac{|\Omega'| - |V_2|}{4}. \end{aligned} \tag{11}$$

Noting that $|V_2| \geq |\Omega'|$, (11) gives

$$\begin{aligned} |J| &\leq \frac{5}{4}n + \frac{|V^*(G)|}{8} - \frac{1}{2} \\ &\leq \frac{5}{4}(n + \frac{|V^*|}{2}) - \frac{1}{2}. \end{aligned}$$

Applying (9), the result $|J| \leq \frac{5}{4} OPT_{LP}(G) - \frac{1}{2}$ now follows. □

The next two corollaries follow immediately from Theorem 3.

Corollary 1. *Algorithm 5_4 is a $\frac{5}{4}$-approximation algorithm for $2EC(G)$ for subcubic bridgeless multigraphs.*

Corollary 2. *The integrality gap $\alpha^{2EC(G)}$ is less than $\frac{5}{4}$ for subcubic bridgeless multigraphs.*

5. Computational Study on the Integrality Gap of $2EC(G)$

In this section we report on a computational study where we investigate the worst-case ratio between $ILP(G)$ and $LP(G)$ for graphs with a small number of nodes, thus obtaining a lower bound on $\alpha^{2EC(G)}$ for those types of graphs. Here we give a brief summary of the methods used and results obtained. For more details, see [13].

5.1. Methodology

It is known that the computational complexity for solving $ILP(G)$ is NP-hard. However, it is practically possible to solve $ILP(G)$ in reasonable time for graphs G of small size. Therefore the graphs in this experimental study were limited to the following three sets:

- General simple graphs $\mathbb{G} = \bigcup_{k=3}^{10} \mathbb{G}_k$, where $\mathbb{G}_k$ denotes the set of all non-isomorphic 2-edge-connected simple graphs with k nodes;
- Cubic simple graphs $\mathbb{C} = \bigcup_{k=6}^{16} \mathbb{C}_k$, where $\mathbb{C}_k$ denotes the set of all non-isomorphic 2-edge-connected cubic simple graphs with k nodes; and
- Subcubic simple graphs $\mathbb{S} = \bigcup_{k=3}^{16} \mathbb{S}_k$, where $\mathbb{S}_k$ denotes the set of all non-isomorphic 2-edge-connected subcubic simple graphs with k nodes.

We let $\mathcal{G}$ denote the complete set of all graphs studied in this experiment, i.e. $\mathcal{G} = \mathbb{G} \cup \mathbb{C} \cup \mathbb{S}$. With the objective of learning more about the lower bound for the integrality gap $\alpha^{2EC(G)}$ of the LP relaxation for $2EC(G)$, we calculated the ratio, denoted by $\alpha(G)$, between the optimal objective value $OPT(G)$ and $OPT_{LP}(G)$ for all graphs $G \in \mathcal{G}$. Note that the maximum ratio $\alpha(G)$ among all $G \in \mathcal{G}$ provides a lower bound for the value of $\alpha^{2EC(G)}$ for graphs of that type.

By using the `nauty` package (Version 2.4), developed by Brendan D. McKay [14], we were able to obtain all non-isomorphic connected graphs of a certain category (i.e. general, cubic, subcubic), and then eliminate all the graphs with bridges. We then formulated $ILP(G)$ and $LP(G)$ for each graph in our set. Finally, we used Gurobi™ Optimizer (Version 5.0) to obtain solutions to $ILP(G)$ and $LP(G)$ for each G. The program designed for our experiments was developed using the *C* programming language, on a 64-bit system running Mircosoft®Windows 7 Professional, with a Lenovo®Thinkpad X201 laptop equipped with Intel®Core™i5 M480 @ 2.67GHz, and 4.00 GB installed memory (RAM).

5.2. Analysis of Results

Facing a large amount of data, it became difficult for us to analyze all results with the limited resources available. For example, the number of all non-isomorphic 2-edge-connected graphs on 10 nodes is 9,804,368, and it took the program approximately 11 days to finish the experiment process for all graphs $G \in \mathbb{G}_{10}$. In order to learn more about the lower bound for the value of $\alpha^{2EC(G)}$ in general and the upper bound for particular classes and sizes of graphs, more attention was given to the data that resulted in a higher ratio between $OPT(G)$ and $OPT_{LP}(G)$. Figure 2 demonstrates the trend in the changes of the maximum ratios between $OPT(G)$ and $OPT_{LP}(G)$ for $G \in \mathbb{G}_k$, $G \in \mathbb{C}_k$ and $G \in \mathbb{S}_k$. Table 1 gives a summary of the maximum value of the ratio between $OPT(G)$ and $OPT_{LP}(G)$ for graphs in each of the three categories, along with the size (i.e. number of nodes) in the graph that gave the maximum value.

Let $\alpha(\mathbb{G}_k)$, $\alpha(\mathbb{C}_k)$ and $\alpha(\mathbb{S}_k)$ denote the maximum ratios between $OPT(G)$ and $OPT_{LP}(G)$ for $G \in \mathbb{G}_k$, $G \in \mathbb{C}_k$ and $G \in \mathbb{S}_k$ respectively. It is noted from Figure 2 that our result on $\alpha(\mathbb{S}_{16})$ reached the highest value overall, i.e. $\frac{9}{8}$. Note that this was the previous best known lower bound on the integrality gap of the LP relaxation for $2EC(G)$ [4], although in [4] this bound was reached asymptotically by an infinite family of graphs. In addition, it is noted that for each value of k ($3 \leq k \leq 10$), the set of graphs with k nodes that gave the worst ratio among all graphs with the same size always included subcubic graphs. This supports the idea that subcubic graphs are most likely to give $\alpha^{2EC(G)}$ in general.

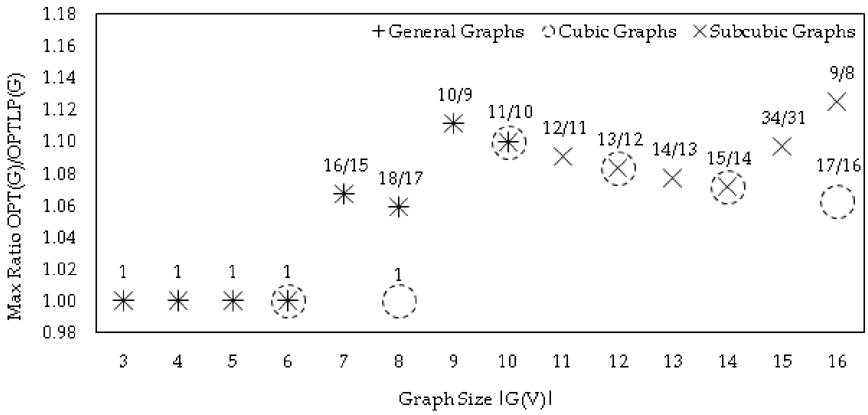


Figure 2. Data analysis of experimental results.

Table 1. Summary of the experimental study.

Graph Category	Max $\alpha(G)$	Corresponding $\|V(G)\|$
$\mathbb{G}_k$ $(3 \leq k \leq 10)$	10/9	9
$\mathbb{C}_k$ $(6 \leq k \leq 16)$	11/10	10
$\mathbb{S}_k$ $(3 \leq k \leq 16)$	9/8	16

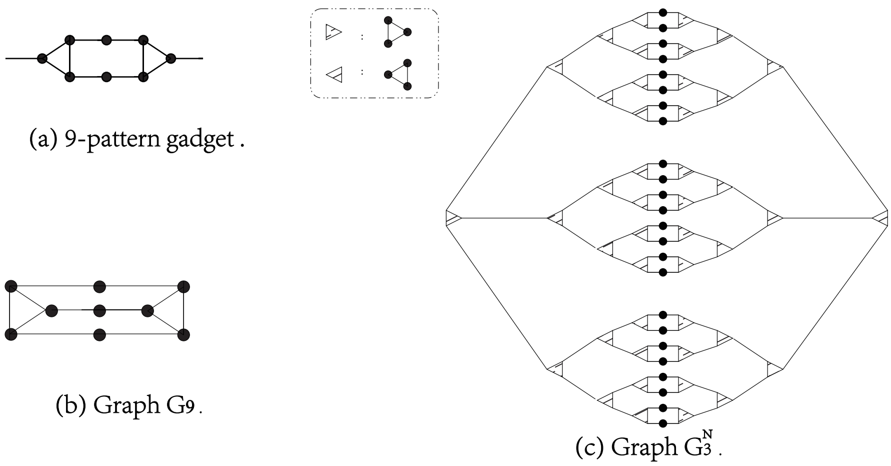


Figure 3. Illustrations for the family G_t^N.

6. New Lower Bounds for the Integrality Gap for $2EC(G)$

In this section we discuss a family of subcubic graphs which asymptotically give a ratio of $\frac{8}{7}$ for $OPT(G)$ and $OPT_{LP}(G)$, thus improving on the previous best known lower bound for $\alpha^{2EC(G)}$ of $\frac{9}{8}$ [4]. In our computational study, we were able to find a pattern which gives relatively high ratio between $OPT(G)$ and $OPT_{LP}(G)$. Inspired by this finding, we designed the following family of graphs by continuously replacing all degree 2 nodes with the 9-pattern gadget shown in Figure 3(a), starting from G_9 (shown in Figure 3(b)). The family of graphs generated from the above operation is referred to as $\mathcal{F}^N$. Executing the replacement for t times gives us the graph $G_t^N \in \mathcal{F}^N (t \geq 0)$. Figure 3(c) shows the graph $G_3^N \in \mathcal{F}^N$, which was obtained by repeating the replacement three times.

Theorem 4. *For the family of graphs $\mathcal{F}^N$, the following hold :*

$$\lim_{t\to\infty} \frac{OPT(G_t^N)}{|V(G_t^N)|} = \frac{8}{7}, \text{ and}$$

$$\lim_{t\to\infty} \frac{OPT(G_t^N)}{OPT_{LP}(G_t^N)} = \frac{8}{7}.$$

Proof. For the convenience of calculation, we refer to every shaded triangle shown in Figure 3 as a *virtual* node, and use $virtual(G_t^N)$ to denote the set of virtual nodes for any graph $G_t^N \in F^N$. If we think of each virtual node as a normal node, it is easy to see that any graph $G_t^N \in \mathcal{F}^N$ can be viewed as a tree of height $(t+1)$, denoted by T_{t+1}, mirrored along its leaves. All internal nodes of T_{t+1} are virtual nodes in G_t^N, while all its leaves are the nodes of degree 2 in G_t^N, which we denote by $d_2(G_t^N)$. Note that T_{t+1} consists of three identical binary trees BT_t of height t connected to one root node.

It is well known that for any binary tree, the number of nodes at depth d is 2^d. Thus letting $leaves(BT_t)$ and $internal(BT_t)$ represent the number of leaves and internal nodes in BT_t, we have

$$|leaves(BT_t)| = 2^t, \tag{12}$$

and

$$\begin{aligned} |internal(BT_t)| &= 1 + 2^1 + \cdots + 2^{t-1} \\ &= \sum_{i=0}^{t-1} 2^i \\ &= 2^t - 1. \end{aligned} \tag{13}$$

In addition, since every internal node has two children, the total number of edges in BT_t is

$$\begin{aligned} |E(BT_t)| &= 1 \times 2 + 2^1 \times 2 + \cdots + 2^{t-1} \times 2 \\ &= 2 \times \sum_{i=0}^{t-1} 2^i \\ &= 2^{t+1} - 2. \end{aligned} \tag{14}$$

Using (13) the number of all virtual nodes in a graph $G_t^N \in \mathcal{F}^N$, which is double the number of all internal nodes in T_{t+1}, is

$$\begin{aligned} |virtual(G_t^N)| &= 2 \times (3 \times |internal(BT_t)| + 1) \\ &= 2 \times [3 \times (2^t - 1) + 1] \\ &= 6 \times 2^t - 4. \end{aligned} \tag{15}$$

Also, using (12) the number of degree two nodes in G_t^N, which is equal to the number of leaves in T_{t+1}, is given by

$$\begin{aligned} |d_2(G_t^N)| &= 3 \times |leaves(BT_t)| \\ &= 3 \times 2^t. \end{aligned} \tag{16}$$

Thus by (15) and (16), it follows that the total number of nodes in G_t^N is

$$\begin{aligned}|V(G_t^N)| &= 3 \times |virtual(G_t^N)| + |d_2(G_t^N)| \\ &= 21 \times 2^t - 12.\end{aligned}$$

On the other hand, since all edges in G_t^N not in a virtual node triangle belong to a 2-edge cut, any feasible ILP or LP solution x must have $x_e = 1$ for such edges. The total number of such edges E', which is twice of the number of edges in T_{t+1}, is

$$\begin{aligned}|E'| &= 2 \times (3 \times |E(BT_t)| + 3) \\ &= 12 \times 2^t - 6.\end{aligned} \tag{17}$$

Note that every edge in E' contributes 1 to both $OPT(G_t^N)$ and $OPT_{LP}(G_t^N)$.

Returning back to the original graph, we need to consider the values of variables on the edges belonging to the virtual node. Shown in Figure 4 are possible feasible solutions for $ILP(G_t^N)$ and $LP(G_t^N)$ for such edges. Hence by

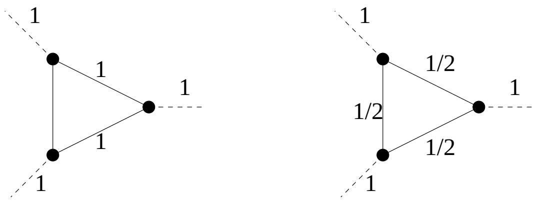


Figure 4. Feasible solutions to $ILP(G_t^N)$ (left) and $LP(G_t^N)$ (right) on a virtual node.

(15) and (17) we have the following:

$$\begin{aligned}OPT(G_t^N) &\leq |E'| + 2|virtual(G_t^N)| \\ &= (12 \times 2^t - 6) + 2 \times (6 \times 2^t - 4) \\ &= 24 \times 2^k - 14,\end{aligned} \tag{18}$$

$$\begin{aligned}OPT_{LP}(G_t^N) &\leq |E'| + \frac{3}{2}|virtual(G_t^N)| \\ &= (12 \times 2^t - 6) + \frac{3}{2} \times (6 \times 2^t - 4) \\ &= 12 \times 2^t - 6 + 9 \times 2^t - 6 \\ &= 21 \times 2^t - 12.\end{aligned} \tag{19}$$

Note that (17) and (19) imply that $OPT_{LP}(G_t^N) = |V(G_t^N)|$, since $OPT_{LP}(G_t^N)$ must be at least $|V(G_t^N)|$. Thus equality holds in (19). Equality holds in (18) as well, as at least two edges of any virtual node triangle must be used in $OPT(G_t^N)$.

From the above it follows that

$$\lim_{t\to\infty} \frac{OPT(G_t^N)}{|V(G_t^N)|} = \lim_{t\to\infty} \frac{OPT(G_t^N)}{OPT_{LP}(G_t^N)} = \lim_{t\to\infty} \frac{24 \times 2^t - 14}{21 \times 2^t - 12} = \frac{8}{7},$$

as required.

□

From the above discussion, Corollary 3 naturally follows from Theorem 4.

Corollary 3. *The integrality gap $\alpha^{2EC(G)}$ for $2EC(G)$ is at least $\frac{8}{7}$, even when restricted to subcubic bridgeless graphs.*

Acknowledgement

The authors would like to thank Andras Sebő for his valuable suggestions and comments.

References

[1] B. Csaba, M. Karpinski, P. Krysta, Approximability of dense and sparse instances of minimum 2-connectivity, tsp and path problems, in: D. Eppstein (Ed.), SODA, ACM/SIAM, 2002, pp. 74–83.
[2] A. Alexander, S. Boyd, P. 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`
[3] T. Mömke, O. Svensson, Approximating graphic tsp by matchings, in: R. Ostrovsky (Ed.), FOCS, IEEE, 2011, pp. 560–569.
[4] A. Sebő, J. Vygen, Shorter tours by nicer ears, CoRR abs/1201.1870.
[5] S. Khuller, U. Vishkin, Biconnectivity approximations and graph carvings, J. ACM 41 (2) (1994) 214–235.
[6] J. Cheriyan, A. Sebő, Z. Szigeti, Improving on the 1.5 approximation of a smallest 2-edge connected spanning subgraph, SIAM J. Discrete Math. 14 (2001) 170–180.
[7] S. Vempala, A. Vetta, Factor 4/3 approximations for minimum 2-connected subgraphs, in: K. Jansen, S. Khuller (Eds.), Approximation Algorithms for Combinatorial Optimization, Third International Workshop, Springer, 2000, pp. 262–273.
[8] P. Krysta, V. S. A. Kumar, Approximation algorithms for minimum size 2-connectivity problems, in: A. Ferreira, H. Reichel (Eds.), STACS, Springer, 2001, pp. 431–442.
[9] W. T. Huh, Finding 2-edge connected spanning subgraphs, Oper. Res. Lett. 32 (3) (2004) 212–216.
[10] S. Boyd, S. Iwata, K. Takazawa, Finding 2-factors closer to tsp tours in cubic graphs, SIAM J. Discrete Math. 27 (2) (2013) 918–939.
[11] A.J.Hoffman, Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, Combinatorial Analysis (1960) 113–127.
[12] A. Schrijver, Combinatorial Optimization, Springer, 2003, Ch. Chapters 11-12.
[13] Y. Sun, Theoretical and experimental studies on the minimum size 2-edge-connected spanning subgraph problem, Master's thesis, University of Ottawa, Ottawa, Canada (2013).
URL `http://hdl.handle.net/10393/24198`
[14] B. D. McKay, Practical graph isomorphism, Congressus Numerantium 30 (1981) 45–87.