Improving on the 1.5-Approximation of a Smallest 2-Edge Connected Spanning Subgraph

J. Cheriyan * A. Sebő † Z. Szigeti ‡

September 21, 2000

Abstract

We give a $\frac{17}{12}$-approximation algorithm for the following NP-hard problem:

> Given a simple undirected graph, find a 2-edge connected spanning subgraph that has the minimum number of edges.

The best previous approximation guarantee was $\frac{3}{2}$. If the well known $\frac{4}{3}$ conjecture for the metric TSP holds, then the optimal value (minimum number of edges) is at most $\frac{4}{3}$ times the optimal value of a linear programming relaxation. Thus our main result gets half-way to this target.

1 Introduction

Given a simple undirected graph, consider the problem of finding a 2-edge connected spanning subgraph that has the minimum number of edges. The problem is NP-hard, via a reduction from the Hamiltonian cycle problem. A number of recent papers have focused on approximation algorithms for this and other related problems, [3]. An *α-approximation algorithm* for a combinatorial optimization problem runs in polynomial time and delivers a solution whose value is always within the factor α of the optimum value. The quantity α is called the *approximation guarantee* of the algorithm. We use the abbreviation 2-ECSS for 2-edge connected spanning subgraph.

Here is an easy 2-approximation algorithm for the problem:

> Take an ear decomposition of the given graph (see Section 2 for definitions), and discard all 1-ears (ears that consist of one edge). Then the resulting graph is 2-edge connected and has at most $2n-3$ edges, while the optimal subgraph has $\geq n$ edges, where n is the number of nodes.

Khuller & Vishkin [11] were the first to improve on the approximation guarantee of 2. They gave a simple and elegant algorithm based on depth-first search that achieves an approximation guarantee of 1.5. We improve Khuller & Vishkin's $\frac{18}{12}$-approximation guarantee to $\frac{17}{12}$. If the well known $\frac{4}{3}$ conjecture for the metric TSP holds, then the optimal value (minimum number of edges) is at most

*jcheriyan@math.uwaterloo.ca Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

†Andras.Sebo@imag.fr Departement de Mathematiques Discretes, CNRS, Laboratoire LEIBNIZ (CNRS, INPG, UJF) - IMAG, 46 Avenue Felix Viallet, 38000 Grenoble Cedex, France.

‡Zoltan.Szigeti@ecp6.jussieu.fr Equipe Combinatoire, Université Paris VI, 4 place Jussieu, Couloir 45-46 3e, 75252 Paris, France.

$\frac{4}{3}$ times the optimal value of a linear programming relaxation, see Theorem 5.1. Thus our main result gets half-way to this target.

Let $G = (V, E)$ be the given simple undirected graph, and let n and m denote $|V|$ and $|E|$. Assume that G is 2-edge connected.

Our method is based on a matching-theory result of András Frank, namely, there is a good characterization for the minimum number of even-length ears over all possible ear decompositions of a graph, and moreover, an ear decomposition achieving this minimum can be computed efficiently, [5]. Recall that the 2-approximation heuristic starts with an arbitrary ear decomposition of G. Instead, if we start with an ear decomposition that maximizes the number of 1-ears, and if we discard all the 1-ears, then we will obtain the optimal solution. In fact, we start with an ear decomposition that maximizes the number of odd-length ears. Now, discarding all the 1-ears gives an approximation guarantee of 1.5 (see Proposition 3.4 below). To do better, we repeatedly apply "ear splicing" steps to the starting ear decomposition to obtain a final ear decomposition such that the number of odd-length ears is the same, and moreover, the internal nodes of distinct 3-ears are nonadjacent. We employ two lower bounds to show that discarding all the 1-ears from the final ear decomposition gives an approximation guarantee of $\frac{17}{12}$. The first lower bound is the "component lower bound" due to Garg et al [8, Lemma 4.1], see Proposition 2.4 below. The second lower bound comes from the minimum number of even-length ears in an ear decomposition of G, see Proposition 3.3 below.

A preliminary version of this paper has appeared in [2]. After this paper was submitted for journal publication, an improved approximation guarantee for the minimum-size 2-ECSS problem has been reported by Vempala and Vetta [15].

After developing some preliminaries in Sections 2 and 3, we present our heuristic in Section 4. Section 5.1 shows the relation of the well known $\frac{4}{3}$ conjecture for the metric TSP to the problem of finding a $\frac{4}{3}$-approximation algorithm for a minimum-size 2-ECSS, see Theorem 5.1. Section 5.2 has two examples showing that our analysis of the heuristic is tight. Section 5.2 also compares the two lower bounds with the optimal value.

A Useful Assumption.

For our heuristic to work, it is essential that the given graph be 2-node connected. Hence, in Section 4 of the paper where our heuristic is presented, we will assume that the given graph G is 2-node connected. Otherwise, if G is not 2-node connected, we compute the blocks (i.e., the maximal 2-node connected subgraphs) of G, and apply the algorithm separately to each block. We compute a 2-ECSS for each block, and output the union of the edge sets as the edge set of a 2-ECSS of G. The resulting graph has no cut edges since the subgraph found for each block has no cut edge. Moreover, if an approximation guarantee of α holds for each block, then it holds for G, because the optimal value for G equals the sum of the optimal values for the blocks.

2 Preliminaries

Except in Section 5.1, all graphs are simple, that is, there are no loops nor multiedges. A closed path means a cycle, and an open path means that all the nodes are distinct.

An *ear decomposition* of the graph G is a partition of the edge set into open or closed paths, $P_0 + P_1 + \ldots + P_k$, such that P_0 is the trivial path with one node, and each P_i $(1 \le i \le k)$ is a path that has both end nodes in $V_{i-1} = V(P_0) \cup V(P_1) \cup \ldots \cup V(P_{i-1})$ but has no internal nodes in V_{i-1}. A (closed or open) *ear* means one of the (closed or open) paths $P_1, \ldots, P_k$ in the ear decomposition;

note that P_0 is not regarded as an ear. By induction on the number of ears k, it is easily seen that for the number of edges in G',

$$|E'| = |V| + k - 1. \tag{1}$$

For a nonnegative integer ℓ, an *ℓ-ear* means an ear that has ℓ edges. An ℓ-ear is called *even* if ℓ is an even number, otherwise, the ℓ-ear is called *odd.* An *open* ear decomposition $P_0 + P_1 + \ldots + P_k$ is one such that all the ears $P_2, \ldots, P_k$ are open. (The ear P_1 is always closed.)

Proposition 2.1 (Whitney [16]) *(i) A graph is 2-edge connected iff it has an ear decomposition.*
(ii) A graph is 2-node connected iff it has an open ear decomposition.

An *odd* ear decomposition is one such that every ear has an odd number of edges. The graph G is called *factor-critical* if for every node $v \in V(G)$, there is a perfect matching in $G - v$. The next result gives another characterization of factor-critical graphs.

Theorem 2.2 (Lovász [12], Theorem 5.5.1 in [13]) *A graph is factor-critical iff it has an odd ear decomposition.*

It follows that a factor-critical graph is necessarily 2-edge connected. An *open odd* ear decomposition $P_0 + P_1 + \ldots + P_k$ is an odd ear decomposition such that all the ears $P_2, \ldots, P_k$ are open.

Theorem 2.3 (Lovász & Plummer, Theorem 5.5.2 in [13]) *A 2-node connected factor-critical graph has an open odd ear decomposition. Given such a graph $G = (V, E)$, an open odd ear decomposition can be constructed in time $O(|V| \cdot |E|)$.*

Let $\varepsilon(G)$ denote the minimum number of edges in a 2-ECSS of G. For a graph H, let $c(H)$ denote the number of (connected) components of H. Garg et al [8, Lemma 4.1] use the following lower bound on $\varepsilon(G)$.

Proposition 2.4 *Let $G = (V, E)$ be a 2-edge connected graph, and let S be a nonempty set of nodes such that the deletion of S results in a graph with $c = c(G - S)$ components. Then $\varepsilon(G) \geq |V| + c - |S|$.*

Proof: Focus on an arbitrary component D of $G - S$ and note that it contributes $\geq |V(D)| + 1$ edges to an optimal 2-ECSS, because every node in D contributes ≥ 2 edges, and at least two of these edges have exactly one end node in D. Summing over all components of $G - S$ gives the result. $\square$

For the graph $G = (V, E)$, let $L_c^\star(G)$ denote $\max\{|V| + c(G - S) - |S| \,:\, \emptyset \neq S \subseteq V\}$; by Proposition 2.4, $\varepsilon(G) \geq L_c^\star(G)$.

For a set of nodes $S \subseteq V$ of a graph $G = (V, E)$, $\delta(S)$ denotes the set of edges that have one end node in S and one end node in $V - S$. For the singleton node set $\{v\}$, we use the notation $\delta(v)$. For a vector $x : E \to \mathbf{R}$ and an edge set $F \subseteq E$, $x(F)$ denotes $\sum_{e \in F} x_e$.

3 Frank's Theorem and a New Lower Bound for ε

For a 2-edge connected graph $G = (V, E)$, let $\varphi(G)$ (or φ) denote the minimum number of even ears over all possible ear decompositions. For example: $\varphi(G) = 0$ if G is a factor-critical graph (e.g., G is an odd clique $K_{2\ell+1}$ or an odd cycle $C_{2\ell+1}$), $\varphi(G) = 1$ if G is an even clique $K_{2\ell}$ or an even cycle $C_{2\ell}$, and $\varphi(G) = \ell - 1$ if G is the complete bipartite graph $K_{2,\ell}$ ($\ell \geq 2$). Let $L_\varphi(G)$ denote $|V| + \varphi(G) - 1$.

A *join* of a graph G is an edge set $J \subseteq E(G)$ such that for (the edge set of) every cycle $Q \subseteq E(G)$ we have $|J \cap Q| \leq |Q|/2$. For example, any matching is a join. Let $\mu(G)$ denote the maximum size of a join of the graph G.

The proof of the next result appears in [5], see Theorem 4.5 and Section 2 of [5].

Theorem 3.1 (A. Frank [5]) *Let $G = (V, E)$ be a 2-edge connected graph. An ear decomposition $P_0 + P_1 + \ldots + P_k$ of G having $\varphi(G)$ even ears can be computed in time $O(|V| \cdot |E|)$. Moreover, $L_\varphi(G) = 2\mu(G)$.*

Proposition 3.2 *For every 2-node connected graph G, there exists an* open *ear decomposition $P_0 + P_1 + \ldots + P_k$ that has $\varphi(G)$ even ears. Such an ear decomposition can be computed in time $O(|V| \cdot |E|)$.*

Proof: Apply Theorem 3.1 to construct an ear decomposition having $\varphi(G)$ even ears (the ears may be open or closed). Subdivide one edge in each even ear by adding one new node and one new edge. The resulting ear decomposition is odd. Hence, the resulting graph G' is factor critical, and also, G' is 2-node connected since G is 2-node connected. Apply Theorem 2.3 to construct an open odd ear decomposition of G'. Finally, in the resulting ear decomposition, "undo" the $\varphi(G)$ edge subdivisions to obtain the desired ear decomposition $P_0 + P_1 + \ldots + P_k$ of G.

The running time for constructing $P_0 + P_1 + \ldots + P_k$ is $O(|V| \cdot |E|)$. Note that there are constructive proofs for both Theorems 3.1 and 2.3, and each construction can be implemented in time $O(|V| \cdot |E|)$. $\square$

Frank's theorem gives the following lower bound on the minimum number of edges in a 2-ECSS.

Proposition 3.3 *Let $G = (V, E)$ be a 2-edge connected graph. Then $\varepsilon(G) \geq L_\varphi(G) = 2\mu(G)$.*

Proof: Consider an arbitrary 2-ECSS $G' = (V, E')$ of G. Note that G' contains all nodes of G, but there may be several edges in $E - E'$. If G' has an ear decomposition with fewer than $\varphi(G)$ even ears, then we can obtain an ear decomposition of G with fewer than $\varphi(G)$ even ears as follows: we start with the ear decomposition of G', and for each edge $vw \in E - E'$, we add the 1-ear v, w. This contradiction to the definition of $\varphi(G)$ shows that every ear decomposition of G' has $\geq \varphi(G)$ even ears. Let $P_0 + P_1 + \ldots + P_k$ be an ear decomposition of the 2-ECSS G', where $k \geq \varphi(G)$. By (1), $|E'| = |V| + k - 1 \geq |V| + \varphi(G) - 1$. The result follows. $\square$

The next result is not useful for our main result, but we include it for completeness.

Proposition 3.4 *Let $G = (V, E)$ be a 2-edge connected graph. Let $P_0 + P_1 + \ldots + P_k$ be an ear decomposition of G that has $\varphi(G)$ even ears, and let $G' = (V, E')$ be obtained by discarding all the 1-ears from $P_0 + P_1 + \ldots + P_k$. Then $|E'|/\varepsilon(G) \leq 1.5$.*

Proof: Let t be the number of internal nodes in the odd ears of $P_0 + P_1 + \ldots + P_k$. (Note that the node in P_0 is not counted by t.) Then, the number of edges contributed to E' by the odd ears is $\leq 3t/2$, and the number of edges contributed to E' by the even ears is $\leq \varphi + |V| - t - 1 = L_\varphi(G) - t$. By applying Proposition 3.3 (and the fact that $\varepsilon(G) \geq |V|$) we get,

$$|E'|/\varepsilon(G) \leq (t/2 + L_\varphi(G))/\max(|V|, L_\varphi(G)) \leq (t/2|V|) + L_\varphi(G)/L_\varphi(G) \leq 1.5.$$

□

4 Approximating ε via Frank's Theorem

For a graph H and an ear decomposition $P_0 + P_1 + \ldots + P_k$ of H, we call an ear P_i of length ≥ 2 *pendant* if none of the internal nodes of P_i is an end node of another ear P_j of length ≥ 2. In other words, if we discard all the 1-ears from the ear decomposition, then one of the remaining ears is called pendant if all its internal nodes have degree 2 in the resulting graph.

Let $G = (V, E)$ be the given graph, and let $\varphi = \varphi(G)$. Recall the assumption from Section 1 that G is 2-node connected. By an *evenmin ear decomposition* of G, we mean an ear decomposition that has $\varphi(G)$ even ears. Our method starts with an open evenmin ear decomposition $P_0 + P_1 + \ldots + P_k$ of G, see Proposition 3.2, i.e., for $2 \leq i \leq k$, every ear P_i has distinct end nodes, and the number of even ears is minimum possible. The method performs a sequence of "ear splicings" to obtain another (evenmin) ear decomposition $Q_0 + Q_1 + \ldots + Q_k$ (the ears Q_i may be either open or closed) such that the following holds:

Property (α)

(0) the number of even ears is the same in $P_0 + P_1 + \ldots + P_k$ and in $Q_0 + Q_1 + \ldots + Q_k$,

(1) every 3-ear Q_i is a pendant ear,

(2) for every pair of 3-ears Q_i and Q_j, there is no edge between an internal node of Q_i and an internal node of Q_j, and

(3) every 3-ear Q_i is open, where $Q_i \neq Q_1$.

See Figure 1 for an illustration of several cases in an "ear splicing" step.

Proposition 4.1 *Let $G = (V, E)$ be a 2-node connected graph. Let $P_0 + P_1 + \ldots + P_k$ be an open evenmin ear decomposition of G. There is a linear-time algorithm that given $P_0 + P_1 + \ldots + P_k$, finds an ear decomposition $Q_0 + Q_1 + \ldots + Q_k$ satisfying property (α).*

Proof: The following procedure adds one ear P_j, $(j = 0, \ldots, k)$ at a time using ear splicing if necessary.

Suppose that $Q'_0 + \ldots + Q'_{j-1}$ have property (α). We will find in linear time in the length of P_j an ear-decomposition $Q_0 + \ldots + Q_j$ of $Q'_0 + \ldots + Q'_{j-1} + P_j$, that satisfies (α). Let T' denote here the set of internal nodes of the ears of length 3. If P_j has no endpoint in T', or has one endpoint in T' and the length of P_j is 1, we simply add it, that is, $Q_i := Q'_i$ $(i = 0, \ldots, j-1)$, $Q_j = P_j$.

If P_j has one endpoint in T' and has length bigger than 1, or has two endpoints in T' and these are in different 3-ears, then we add the appropriate subpaths of length 2 of these 3-ear(s) to P_j (Figure 1a and c). We get from the 3-ear(s) and P_j by this ear-splicing an ear Q_j and

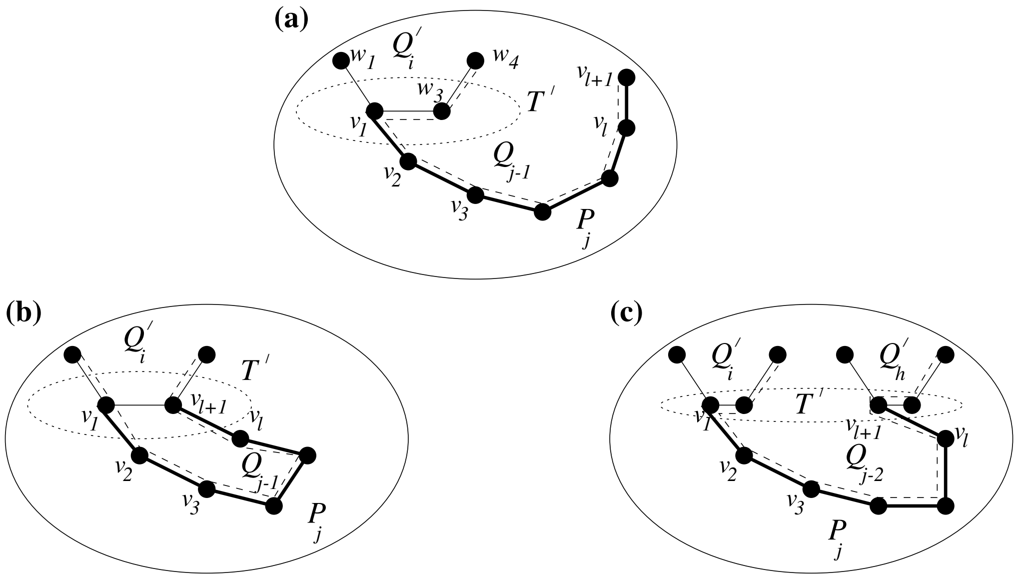

Figure 1: Illustration of the proof of Proposition 4.1. **(a)**, **(b)** The first and second cases. Ears P_j and Q_i' are indicated by solid lines, and ear Q_{j-1} is indicated by dashed lines. **(c)** The third case. Ears P_j, Q_i', Q_h' are indicated by solid lines, and ear Q_{j-2} is indicated by dashed lines.

1-ear(s). Clearly, Q_j has the same parity as P_j. (The two endpoints of Q_j may be equal, but we don't care, since its length is bigger than 3.) Therefore, leaving the other ears as they are, the ear decomposition we got has property (α).

Finally if P_j has two endpoints in T' and both in the same ear (this is the only remaining case), then we add the first and last edge of the 3-ear to P_j to get Q_j, and the middle edge becomes a 1-ear (Figure 1b). The two new ears (Q_j and the 1-ear) should replace the 3-ear and P_j in the decomposition $Q_0' + \ldots + Q_{j-1}' + P_j$, and property (α) is again maintained. $\square$

Remark: In the induction step, which applies for $j \geq 2$ (but not for $j = 1$), it is essential that the ear P_j is open, though Q_i' (and Q_h') may be either open or closed. Note that Q_1 is not a 3-ear provided $|V| \neq 3$. Our main result (Theorem 4.3) does not use part (3) of property (α).

Our approximation algorithm for a minimum-size 2-ECSS computes the ear decomposition $Q_0 + Q_1 + \ldots + Q_k$ satisfying property (α), starting from an open evenmin ear decomposition $P_0 + P_1 + \ldots + P_k$. Then, the algorithm discards all the edges in 1-ears. Let the resulting graph be $G' = (V, E')$. G' is 2-edge connected by Proposition 2.1.

Let T denote the set of internal nodes of the 3-ears of $Q_0 + Q_1 + \ldots + Q_k$, and let $t = |T|$. (Note that the node in Q_0 is not counted by t.) Property (α) implies that in the subgraph of G induced by T, $G[T]$, every (connected) component has exactly two nodes. Consider the approximation guarantee for G', i.e., the quantity $|E'|/\varepsilon(G)$.

Lemma 4.2 *$\varepsilon(G) \geq L_c^\star(G) \geq 3t/2$.*

Proof: Apply Proposition 2.4 with $S = V - T$ (so $|S| = n - t$) and $c = c(G - S) = t/2$ to get $\varepsilon(G) \geq n - (n - t) + (t/2)$. $\square$

Theorem 4.3 *Given a 2-edge connected graph $G = (V, E)$, the above algorithm finds a 2-ECSS $G' = (V, E')$ such that $|E'|/\varepsilon(G) \le \frac{17}{12}$. The algorithm runs in time $O(|V| \cdot |E|)$.*

Proof: By the previous lemma and Proposition 3.3,

$$\varepsilon(G) \ge \max(3t/2, L_\varphi(G)). \tag{2}$$

We claim that

$$|E'| \le \frac{t}{4} + \frac{5L_\varphi(G)}{4}. \tag{3}$$

To see this, let us denote by k, k_e, k_3, k_o the number of ears, the number of even ears, the number of 3-ears and the number of odd ears of length > 3. Note that $k = k_e + k_3 + k_o$, $k_e = \varphi(G)$, $k_3 = t/2$ and $k_o \le (n-t-1)/4$. Thus, by (1), $|E'| = n+k-1 = n+k_e+k_3+k_o-1 \le n+\varphi(G)+t/2+(n-t-1)/4-1 \le t/4 + 5(n+\varphi(G)-1)/4 = t/4 + 5L_\varphi(G)/4$.

The approximation guarantee follows since, by (2) and (3),

$$\frac{|E'|}{\varepsilon(G)} \le \frac{t/4 + 5L_\varphi(G)/4}{\max(3t/2, L_\varphi(G))} \le \frac{t}{4}\,\frac{2}{3t} + \frac{5L_\varphi(G)}{4L_\varphi(G)} = \frac{17}{12}.$$

$\Box$

The next result follows from the proof of Theorem 4.3.

Corollary 4.4 *For a 2-edge connected graph $G = (V, E)$,*

$$\varepsilon(G) \le \frac{5}{4} L_\varphi + \frac{1}{6} L_c^\star \le \frac{17}{12} \max(L_c^\star, L_\varphi).$$

5 Conclusions

5.1 Lower Bounds for ε and the Relation to the TSP $\frac{4}{3}$ Conjecture

This subsection has a comparison of several lower bounds for $\varepsilon(G)$; throughout, $G = (V, E)$ denotes an arbitrary 2-edge connected graph. The best of these lower bounds is given by a linear programming relaxation based on cut constraints. Moreover, we show that $\varepsilon(G)$ is at most $\frac{17}{12}$ times this lower bound. Theorem 5.1 below implies that if the well known TSP $\frac{4}{3}$ conjecture is true, then we have $\frac{4}{3}$ rather than $\frac{17}{12}$ in the previous statement.

Recall that $L_\varphi(G) = |V| + \varphi(G) - 1 = 2\mu(G)$ is a lower bound on $\varepsilon(G)$, where $\mu(G)$ is the maximum size of a join of G, see Proposition 3.3.

Garg et al [8, Theorem 4.2] introduced another lower bound on $\varepsilon(G)$ that we denote by L_c. Let

$$L_c(G) = \max\{\sum_{i=1}^{\ell} c(G - S_i) \ : \ S_1, S_2, \ldots, S_\ell \text{ is a partition of } V, \text{ where } \ell \text{ is any integer} \ge 1\}.$$

(We remark that in the lower bound in [8, Theorem 4.2] $S_1, S_2, \ldots, S_\ell$ is a subpartition rather than a partition, but it can be seen that this lower bound equals L_c.) Clearly, $L_c \ge |V|$, by the partition of V into singleton sets. Notice that the lower bound in Proposition 2.4, $L_c^\star(G) = \max\{c(G-S) + |V - S| \ : \ \emptyset \ne S \subseteq V\}$, is $\le L_c$; to see this, apply the definition of L_c with $S_1 = S$ and $S_2, \ldots, S_\ell$ being singleton sets of $V - S$.

Let $L_z(G)$ denote the optimal value of the following linear programming relaxation of the minimum-size 2-ECSS problem. There is one nonnegative variable x_e for each edge e in G, and the other constraints state that every (nontrivial) cut has x-weight at least two. Let $\mathbf{1}$ be a vector of "1"s with $|E|$ entries.

$$\begin{array}{llrll} L_z(G) = & \text{minimize} & \mathbf{1}\cdot x & & \\ & \text{subject to} & x(\delta(S)) & \geq 2, & \forall S \subset V,\ \emptyset \neq S \neq V \\ & & x & \geq 0, & \\ & & x & \in \mathbf{R}\ . & \end{array}$$

Clearly, $L_z(G)$ is a lower bound on $\varepsilon(G)$ since the incidence vector of a minimum-size 2-ECSS satisfies all the constraints. We may have arbitrary coefficients $c : E \to \mathbf{R}$ in the objective function rather than unit coefficients, and then we will use $L_z(G, c)$ to denote the optimal value. Note that the optimal value of the LP (linear program) is computable in polynomial time, e.g., via the Ellipsoid method.

Now consider the metric TSP (traveling salesman problem). Let $G' = K_n$ be a complete graph and let $c' : E(G') \to \mathbf{R}$ assign metric costs to the edges (so for every triple of nodes i, j, k we have $c'(ij) \leq c'(ik) + c'(kj)$). Let $tsp(G', c')$ denote the minimum cost of a Hamiltonian cycle (or TSP tour) of G', c'. Clearly, the above linear program, but with objective vector c' instead of $\mathbf{1}$, is a relaxation of the TSP, so $L_z(G', c') \leq tsp(G', c')$.

The $\frac{4}{3}$ conjecture for the TSP is: if c' is a metric, then $tsp(G', c') \leq \frac{4}{3} L_z(G', c')$.

Remark: It is known that $tsp(G', c') \leq 1.5\, L_z(G', c')$, see [4, 9, 17]. The conjecture actually refers to the optimal value of the linear programming relaxation that has the additional constraints $x(\delta(v)) = 2$ for each node v; however, if the edge costs are metric, then the addition of the new constraints does not change the optimal value, see [14, 9].

Theorem 5.1 *Let $G = (V, E)$ be a 2-edge connected graph. Then*

$$L_\varphi = 2\mu \leq L_c \leq L_z \leq \varepsilon \leq \frac{17}{12} L_c \leq \frac{17}{12} L_z.$$

Moreover, if the $\frac{4}{3}$ conjecture for the metric TSP holds, then

$$\varepsilon \leq \frac{4}{3} L_z.$$

Proof: To prove the first statement, we will derive the first two inequalities.

- ($2\mu \leq L_c$) Let J be a join of G with $|J| = \mu$. By [6, Theorem 8'] there exists a partition $V_1, \ldots, V_r$ ($r \geq 1$) of V such that $2|J| \leq \sum_{i=1}^{r} c(G - V_i)$. Therefore,

$$2\mu(G) \leq \max\{\sum_{i=1}^{\ell} c(G - S_i)\ :\ S_1, S_2, \ldots, S_\ell \text{ is a partition of } V\} = L_c.$$

Note that the maximization is over ℓ as well as the sets $S_1, S_2, \ldots, S_\ell$ in the partition.

- ($L_c \leq L_z$) Let $S_1, S_2, \ldots, S_\ell$ denote the optimal partition in the definition of L_c, so $L_c = \sum_{i=1}^{\ell} c(G - S_i)$. We sum up the following constraints (inequalities) from the linear program defining L_z: $x(\delta(V(D))) \geq 2$ for each component D of $G - S_i$, for each $i = 1, \ldots, \ell$. Let the resulting inequality be $\sum_{e \in E} a_e x_e \geq b_0$. Clearly, $b_0 = 2\sum_{i=1}^{\ell} c(G - S_i)$. Moreover, note that every coefficient a_{vw} ($vw \in E$) is ≤ 2. To see this, we consider two cases: v, w are in different sets, say $v \in S_i$, $w \in S_j$ ($i \neq j$), or v, w are in the same set S_i. Consider the first case in detail; the inequality for the component of $G - S_i$ containing w contributes x_{vw}, and similarly for the component of $G - S_j$ containing v, so $a_{vw} = 2$. In the second case, $a_{vw} = 0$. Hence,

$$2x(E) \geq \sum_{e \in E} a_e x_e \geq b_0 = 2 \sum_{i=1}^{\ell} c(G - S_i) = 2L_c,$$

where $x : E \rightarrow \mathbf{R}$ is any feasible solution to the linear program. If x is an optimal solution of the linear program, then we get $2L_z = 2x(E) \geq 2L_c$. This proves the second inequality in the theorem. Moreover, by Corollary 4.4 and the fact that $L_c^\star \leq L_c$, we have $\varepsilon(G) \leq \frac{17}{12}\max(L_c^\star, L_\varphi) \leq \frac{17}{12} L_c \leq \frac{17}{12} L_z$. Hence, the first statement in the theorem follows.

Focus on the second statement in the theorem. The multiedge (or uncapacitated) version of our minimum-size 2-ECSS problem is: Given $G = (V, E)$ as above, compute $\widetilde{\varepsilon}(G)$, the minimum size (counting multiplicities) of a 2-edge connected spanning submultigraph $H = (V, F)$, where F is a multiset of edges of E. (To give an analogy, if we take $\varepsilon(G)$ to correspond to the f-factor problem, then $\widetilde{\varepsilon}(G)$ corresponds to the f-matching problem.)

Fact 5.2 *If G is a 2-edge connected graph, then $\widetilde{\varepsilon}(G) = \varepsilon(G)$.*

Proof: Let $H = (V, F)$ give the optimal solution for $\widetilde{\varepsilon}(G)$. If H uses two copies of an edge vw, then we can replace one of the copies by some other edge of G in the cut given by $H - \{vw, vw\}$. In other words, if S is the node set of one of the two components of $H - \{vw, vw\}$, then we replace one copy of vw by some edge from $\delta_G(S) - \{vw\}$. □

Remark: The above is a lucky fact. It fails to generalize, both for minimum-cost (rather than minimum-size) 2-ECSS, and for minimum-size k-ECSS, $k \geq 3$.

Given an n-node graph $G = (V, E)$ together with edge costs c (possibly c assigns unit costs), define its *metric completion* G', c' to be the complete graph $K_n = G'$ with c'_{vw} ($\forall\, v, w \in V$) equal to the minimum-cost of a v-w path in G, c.

Fact 5.3 *Let G be a 2-edge connected graph, and let c assign unit costs to the edges. The minimum cost of the TSP on the metric completion of G, c, satisfies $tsp(G', c') \geq \widetilde{\varepsilon}(G) = \varepsilon(G)$.*

Proof: Let T be an optimal solution to the TSP. We replace each edge $vw \in E(T) - E(G)$ by the edges of a minimum-cost v-w path in G, c. The resulting multigraph H is obviously 2-edge connected, and has $tsp(G', c') = c(H) \geq \widetilde{\varepsilon}(G)$. □

The previous two facts show that $\varepsilon(G) \leq tsp(G', c')$. Moreover, note that for the metric completion G', c', $L_z(G, \mathbf{1})$ equals $L_z(G', c')$, since every feasible solution of the LP on G', c' gives a feasible solution of the LP on G, $\mathbf{1}$ of the same objective value and vice versa. Hence, if the TSP $\frac{4}{3}$ conjecture holds, then we have $\varepsilon(G) \leq tsp(G', c') \leq \frac{4}{3} L_z(G', c') = \frac{4}{3} L_z$. □

5.2 Tight Examples

Our analysis of the heuristic is (asymptotically) tight. We give two example graphs. Each is an n-node Hamiltonian graph $G = (V, E)$, where the heuristic (in the worst case) finds a 2-ECSS $G' = (V, E')$ with $17n/12 - \Theta(1)$ edges.

Here is the first example graph, $G = (V, E)$ (see Figure 2 (top)). The number of nodes is $n = 3 \times 5^q$, and $V = \{0, 1, 2, ..., 3 \times 5^q - 1\}$. The "first node" 0 will be also denoted 3×5^q.

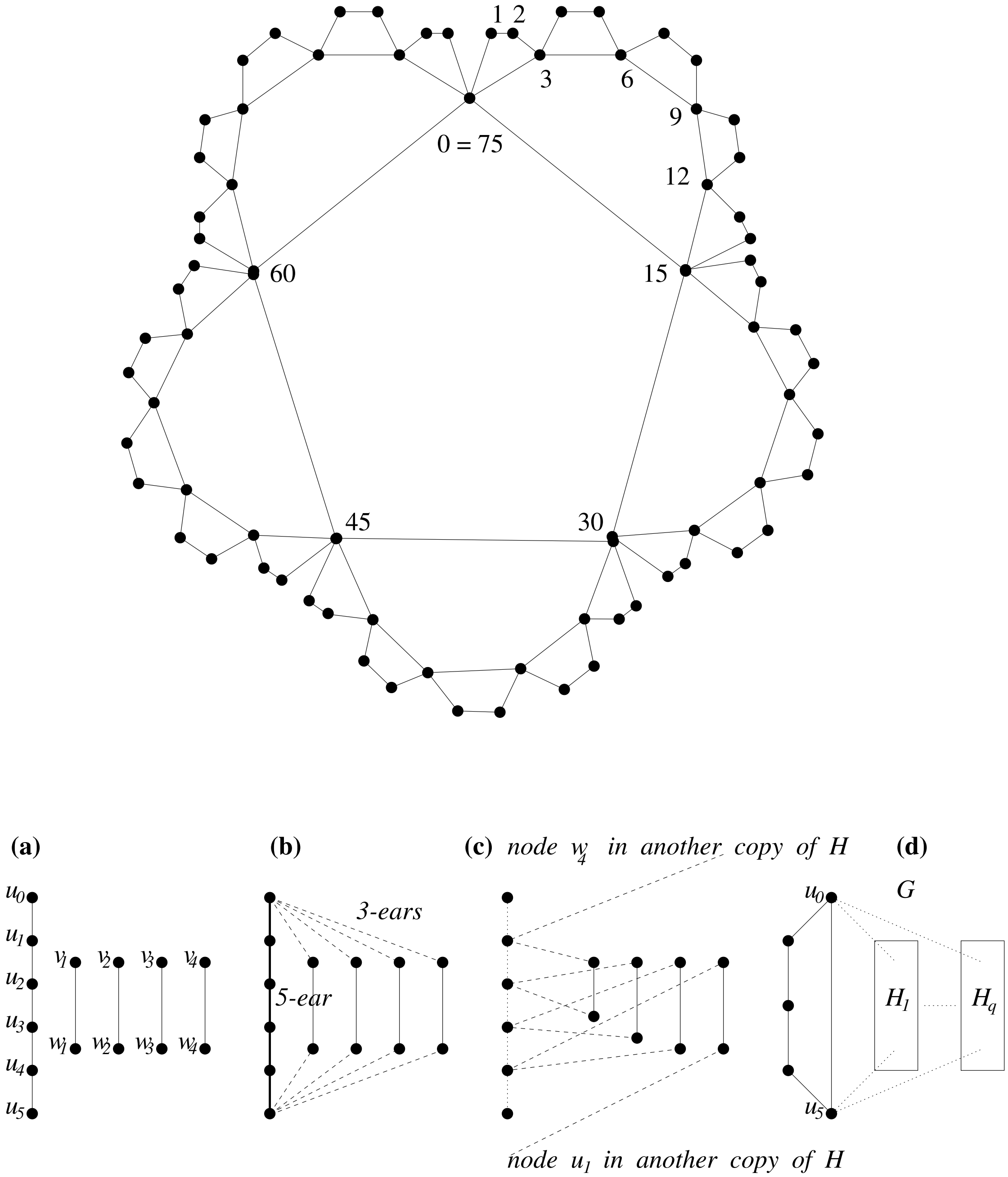

Figure 2: Tight examples for our 17/12-approximation algorithm for minimum-size 2-ECSS.
Top: The first example, with $q = 2$. Some of the node labels $0, 1, 2, \ldots, 3 \times 5^q - 1$ are indicated.
Bottom: The second example. **(a)** The graph H; **(b)** "covering" the nodes of H by a 5-ear and four 3-ears; **(c)** "covering" the nodes of $H - \{u_0, u_5\}$ by a subpath of a Hamiltonian cycle; **(d)** the graph G.

The edge set E consists of (the edge set of) a Hamiltonian cycle together with (the edge sets of) "shortcut cycles" of lengths $n/3, n/(3 \times 5), n/(3 \times 5^2), \ldots, 5$. In detail, $E = \{i(i+1) : 0 \leq i \leq n-1\} \cup \{(3 \times 5^j \times i)(3 \times 5^j \times (i+1)) : 0 \leq j \leq q-1,\ 0 \leq i \leq 5^{q-j}-1\}$. Note that $|E| = 3 \times 5^q + 5^q + 5^{q-1} + \ldots + 5 = (17 \times 5^q - 5)/4$. In the worst case, the heuristic initially finds 5-ears, and finally finds 3-ears, and so the 2-ECSS (V, E') found by the heuristic has all the edges of G. Hence, we have $|E'|/\varepsilon(G) = |E|/n = 17/12 - 1/(12 \times 5^{q-1})$.

The second example graph, G, (see Figure 2 (bottom)) is constructed by "joining" many copies of the following graph H: H consists of a 5-edge path $u_0, u_1, u_2, u_3, u_4, u_5$, and 4 disjoint edges $v_1w_1, v_2w_2, v_3w_3, v_4w_4$. We take q copies of H and identify the node u_0 in all copies, and identify the node u_5 in all copies. Then we add all possible edges u_iv_j, and all possible edges u_iw_j, i.e., we add the edge set of a complete bipartite graph on all the u-nodes and all the v-nodes, and we add the edge set of another complete bipartite graph on all the u-nodes and all the w-nodes. Finally, we add 3 more nodes u'_1, u'_2, u'_3 and 5 more edges to obtain a 5-edge cycle $u_0, u'_1, u'_2, u'_3, u_5, u_0$. Clearly, $\varepsilon(G) = n = 12q + 5$. If the heuristic starts with the closed 5-ear $u_0,u'_1,u'_2,u'_3,u_5,u_0$, and then finds the 5-ears u_0,u_1,u_2,u_3,u_4,u_5 in all the copies of H, and finally finds the 3-ears $u_0v_jw_ju_5$ $(1 \leq j \leq 4)$ in all the copies of H, then we have $|E'| = 17q + 5$.

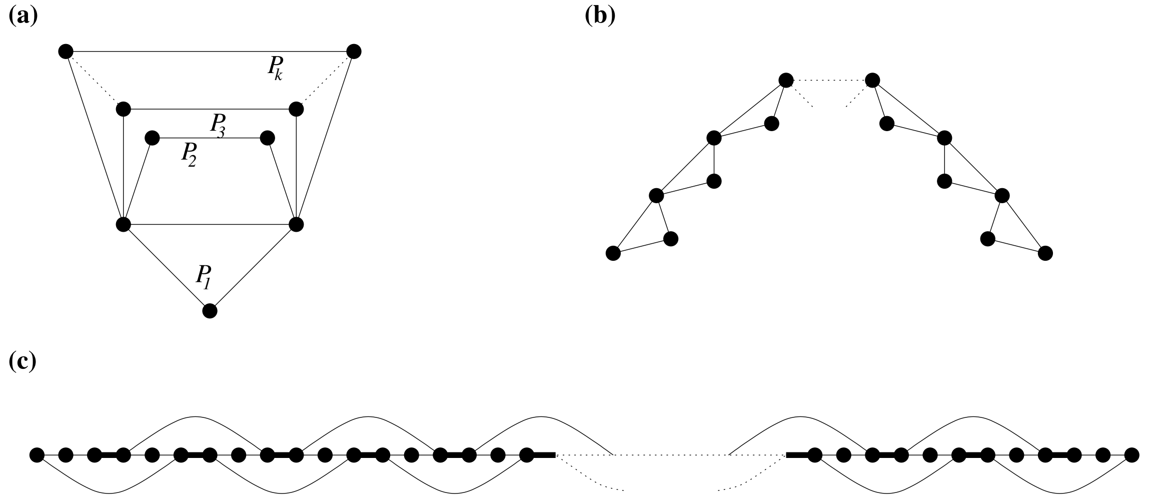

Figure 3: Comparing the lower bounds $L_c^\star$ and L_φ with ε. The graphs achieve the following ratios. **(a)** $\varepsilon/L_\varphi \geq 1.5 - \Theta(1)/n$. **(b)** $\varepsilon/L_c^\star \geq 1.5 - \Theta(1)/n$. **(c)** $\varepsilon/L_c^\star \geq 4/3 - \Theta(1)/n$. To get a graph with $\varepsilon/\max(L_c^\star, L_\varphi) \geq 5/4 - \Theta(1)/n$, subdivide every "thick edge" (3rd edge in path). The resulting graph G has $L_c^\star \leq n+1$ (G has a Hamiltonian path), $L_\varphi \leq n$ (G is factor-critical), $\varepsilon = |E(G)| = (5n-7)/4$.

How do the lower bounds in Proposition 2.4 (namely, $L_c^\star$) and in Proposition 3.3 (namely, L_φ) compare with ε? Let n denote the number of nodes in the graph. There is a 2-node connected graph such that $\varepsilon/L_\varphi \geq 1.5 - \Theta(1)/n$ (see Figure 3(a)). Therefore the upper bound $|E'| \leq 1.5\max(L_\varphi, n)$ of Proposition 3.4 is tight. There is another 2-edge connected (but not 2-node connected) graph such that $\varepsilon/L_c^\star \geq 1.5 - \Theta(1)/n$ and $\varepsilon/L_\varphi \geq 1.5 - \Theta(1)/n$ (see Figure 3(b)). Huh [10] uses the proof of Theorem 3.1 of Garg et al [8] to show that $\varepsilon \leq 1.5\, L_c^\star$. Among 2-node connected graphs, we have a graph with $\varepsilon/L_c^\star \geq 4/3 - \Theta(1)/n$, but we do not know whether there exist graphs that give higher ratios (see Figure 3(c)). There is a 2-node connected graph such that $\varepsilon/\max(L_c^\star, L_\varphi) \geq 5/4 - \Theta(1)/n$, but we do not know whether there exist graphs that give higher ratios (see Figure 3(c)).

Acknowledgment We thank the referees for their comments.

References

[1] R. Carr and R. Ravi, "A new bound for the 2-edge connected subgraph problem," in Proc. Integer Programming and Combinatorial Optimization, 6th International IPCO Conference, Houston, Texas, June 1998, R.E.Bixby, E.A.Boyd, R.Z.Rios-Mercado (Eds.), 112–125, LNCS 1412, Springer, Berlin.

[2] J. Cheriyan, A. Sebő, and Z. Szigeti, "An improved approximation algorithm for minimum size 2-edge connected spanning subgraphs," in Proc. Integer Programming and Combinatorial Optimization, 6th International IPCO Conference, Houston, Texas, June 1998, R.E.Bixby, E.A.Boyd, R.Z.Rios-Mercado (Eds.), 126–136, LNCS 1412, Springer, Berlin.

[3] J. Cheriyan and R. Thurimella, "Approximating minimum-size k-connected spanning subgraphs via matching," *Proc. 37th Annual IEEE Sympos. on Foundat. of Comput. Sci.* (1996), 292–301.

[4] N. Christofides, "Worst-case analysis of a new heuristic for the travelling salesman problem," Technical report, G.S.I.A., Carnegie-Mellon Univ., Pittsburgh, PA, 1976.

[5] A. Frank, "Conservative weightings and ear-decompositions of graphs," *Combinatorica* **13** (1993), 65–81.

[6] A. Frank, A. Sebő and E. Tardos, "Covering directed and odd cuts," *Math. Programming Study* **22** (1984), 99–112.

[7] G. L. Frederickson and J. Ja'Ja', "On the relationship between the biconnectivity augmentation and traveling salesman problems," *Theor. Comp. Sci.* **19** (1982), 189–201.

[8] N. Garg, V. S. Santosh and A. Singla, "Improved approximation algorithms for biconnected subgraphs via better lower bounding techniques," Proc. 4th Annual ACM-SIAM Symposium on Discrete Algorithms 1993, 103–111.

[9] M. X. Goemans and D. J. Bertsimas, "Survivable networks, linear programming relaxations and the parsimonious property," *Mathematical Programming* **60** (1993), 143–166.

[10] T. Huh, "On 2-edge connected spanning subgraphs," M. Math essay, Dept. Combinatorics and Optimization, Univ. Waterloo, Waterloo, Canada, Aug. 1998.

[11] S. Khuller and U. Vishkin, "Biconnectivity approximations and graph carvings," *Journal of the ACM* **41** (1994), 214–235. Preliminary version in: Proc. 24th Annual ACM STOC 1992, 759–770.

[12] L. Lovász, "A note on factor-critical graphs," *Studia Sci. Math. Hungar.*, **7** (1972), 279–280.

[13] L. Lovász and M. D. Plummer, *Matching Theory*, Akadémiai Kiadó, Budapest, 1986.

[14] C. L. Monma, B. S. Munson and W. R. Pulleyblank, "Minimum-weight two-connected spanning networks," *Mathematical Programming* **46** (1990), 153–171.

[15] S. Vempala and A. Vetta, "Factor 4/3 Approximations for Minimum 2-Connected Subgraphs," in Proc. 3rd International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, Max-Planck-Institut für Informatik, Saarbrücken, Germany, September 2000, Springer, Berlin.

[16] H. Whitney, "Nonseparable and planar graphs," *Trans. Amer. Math. Soc.* **34** (1932), 339–362.

[17] L. A. Wolsey, "Heuristic analysis, linear programming and Branch and Bound," *Mathematical Programming Study* **13** (1980), 121–134.