Graphic TSP in Cubic Graphs*

Zdeněk Dvořák[1], Daniel Král'[2], and Bojan Mohar[3]

1 **Computer Science Institute, Charles University, Prague, Czech Republic**
`rakdver@iuuk.mff.cuni.cz`

2 **Mathematics Institute, DIMAP and Department of Computer Science, University of Warwick, Coventry, UK**
`d.kral@warwick.ac.uk`

3 **Department of Mathematics, Simon Fraser University, Burnaby, Canada**
`mohar@sfu.ca`

Abstract

We present a polynomial-time 9/7-approximation algorithm for the graphic TSP for cubic graphs, which improves the previously best approximation factor of 1.3 for 2-connected cubic graphs and drops the requirement of 2-connectivity at the same time. To design our algorithm, we prove that every simple 2-connected cubic n-vertex graph contains a spanning closed walk of length at most $9n/7 - 1$, and that such a walk can be found in polynomial time.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

Keywords and phrases Graphic TSP, approximation algorithms, cubic graphs

Digital Object Identifier 10.4230/LIPIcs.STACS.2017.27

1 Introduction

The Travelling Salesperson Problem (TSP) is one of the most central problems in combinatorial optimization. The problem asks to find a shortest closed walk visiting each vertex at least once in an edge-weighted graph, or alternatively to find a shortest Hamilton cycle in a complete graph where the edge weights satisfy the triangle inequality. The Travelling Salesperson Problem is notoriously hard. The approximation factor of 3/2 established by Christofides [4] has not been improved for 40 years despite a significant effort of many researchers. The particular case of the problem, the Hamilton Cycle Problem, was among the first problems to be shown to be NP-hard. Moreover, Karpinski, Lampis and Schmied [10] have recently shown that the Travelling Salesperson Problem is NP-hard to approximate within the factor 123/122, improving the earlier inapproximability results of Lampis [12] and of Papadimitriou and Vempala [18]. In this paper, we are concerned with an important special case of the Travelling Salesperson Problem, the graphic TSP, which asks to find a shortest closed walk visiting each vertex at least once in a graph where all edges have unit weight. We will refer to such a walk as to a *TSP walk*.

There has recently been a lot of research focused on approximation algorithms for the graphic TSP, which was ignited by the breakthrough of the 3/2-approximation barrier in

* The first author was supported by the Center of Excellence – Institute for Theoretical Computer Science, project P202/12/G061 of Czech Science Foundation. The work of the second author has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648509). This publication reflects only its authors' view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains. The third author was supported in part by an NSERC Discovery Grant (Canada), by the Canada Research Chairs program, and by a Research Grant of ARRS (Slovenia).

34th Symposium on Theoretical Aspects of Computer Science (STACS 2017).
Editors: Heribert Vollmer and Brigitte Vallée; Article No. 27; pp. 27:1–27:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the case of 3-connected cubic graphs by Gamarnik, Lewenstein and Sviridenko [7]. This was followed by the improvement of the 3/2-approximation factor for the general graphic TSP by Oveis Gharan, Saberi and Singh [16]. Next, Mömke and Svensson [14] designed a 1.461-approximation algorithm for the problem and Mucha [15] showed that their algorithm is actually a 13/9-approximation algorithm. This line of research culminated with the 7/5-approximation algorithm of Sebö and Vygen [19].

We here focus on the case of graphic TSP for cubic graphs, which was at the beginning of this line of improvements. The $(3/2 - 5/389)$-approximation algorithm of Gamarnik et al. [7] for 3-connected cubic graphs was improved by Aggarwal, Garg and Gupta [1], who designed a 4/3-approximation algorithm. Next, Boyd et al. [2] found a 4/3-approximation algorithm for 2-connected cubic graphs. The barrier of the 4/3-approximation factor was broken by Correa, Larré and Soto [5] who designed a $(4/3 - 1/61236)$-approximation algorithm for this class of graphs. The currently best algorithm for 2-connected cubic graphs is the 1.3-approximation algorithm of Candráková and Lukoťka [3], based on their result on the existence of a TSP walk of length at most $1.3n - 2$ in 2-connected cubic n-vertex graphs. We improve this result as follows. Note that we obtain a better approximation factor and Theorem 2 also applies to a larger class of graphs.

▶ **Theorem 1.** *There exists a polynomial-time algorithm that for a given 2-connected subcubic n-vertex graph with n_2 vertices of degree two outputs a TSP walk of length at most*

$$\frac{9}{7}n + \frac{2}{7}n_2 - 1 .$$

▶ **Theorem 2.** *There exists a polynomial-time 9/7-approximation algorithm for the graphic TSP for cubic graphs.*

Note that our approximation factor matches the approximation factor for cubic bipartite graphs in the algorithm Karp and Ravi [9], who designed a 9/7-approximation algorithm for the graphic TSP for cubic bipartite graphs. However, van Zuylen [20] has recently found a 5/4-approximation algorithm for this class of graphs. Both the result of Karp and Ravi, and the result of van Zuylen are based on finding a TSP walk of length of at most $9n/7$ and $5n/4$, respectively, in an n-vertex cubic bipartite graph. On the negative side, Karpinski and Schmied [11] showed that the graphic TSP is NP-hard to approximate within the factor of 535/534 in the general case and within the factor 1153/1152 in the case of cubic graphs.

Our contribution in addition to improving the approximation factor for graphic TSP for cubic graphs is also in bringing several new ideas to the table. The proof of our main result, Theorem 1, differs from the usual line of proofs in this area. In particular, to establish the existence of a TSP walk of length at most $9n/7 - 1$ in a 2-connected cubic n-vertex graph, we allow subcubic graphs as inputs and perform reductions in this larger class of graphs. While we cannot establish the approximation factor of 9/7 for this larger class of graphs, we are still able to show that our technique yields the existence of a TSP walk of length at most $9n/7 - 1$ for cubic n-vertex graphs. At this point, we should remark that we have not attempted to optimize the running time of our algorithm.

We conclude with a brief discussion on possible improvements of the bound from Theorem 1. In Section 5, we give a construction of a 2-connected cubic n-vertex graph with no TSP walks of length smaller than $\frac{5}{4}n - 2$ (Proposition 16) and a 2-connected subcubic n-vertex graph with $n_2 = \Theta(n)$ vertices of degree two with no TSP walks of length smaller than $\frac{5}{4}n + \frac{1}{4}n_2 - 1$ (Proposition 14); the former construction was also found independently by Mazák and Lukoťka [13]. We believe that these two constructions provide the tight examples for an improvement of Theorem 1 and conjecture the following. We also refer to a more detailed discussion at the end of Section 5.

▶ **Conjecture 3.** *Every 2-connected subcubic n-vertex graph with n_2 vertices of degree has a TSP walk of length at most*

$$\frac{5}{4}n + \frac{1}{4}n_2 - 1 .$$

Note that Conjecture 3 is of interest for small values of n_2; in particular, its statement is known to be true if $n_2 \geq n/3$ [14]. We would like to stress that it is important that Conjecture 3 deals with simple graphs, i.e., graphs without parallel edges. Indeed, consider the cubic graph G obtained as follows: start with the graph that has two vertices of degree three that are joined by three paths, each having 2ℓ internal vertices of degree two, and replace every second edge of these paths with a pair of parallel edges to get a cubic graph. The graph G has $n = 6\ell + 2$ vertices but no TSP walk of length shorter than $8\ell + 2$.

2 Preliminaries

In this section, we fix the notation used in the paper and make several simple observations on the concepts that we use.

All graphs considered in this paper are *simple*, i.e., they do not contain parallel edges. When we allow parallel edges, we will always emphasize this by referring to a considered graph as to a *multigraph*. We will occasionally want to stress that a graph obtained during the proof has no parallel edges and we will do so by saying that it is simple even if saying so is superfluous. The *underlying graph* of a multigraph H is the graph obtained from H by suppressing parallel edges, i.e., replacing each set of parallel edges by a single edge.

If G is a graph, its vertex set is denoted by $V(G)$ and its edge set by $E(G)$. Further, the number of vertices of G is denoted by $n(G)$ and the number of its vertices of degree two by $n_2(G)$. If w a vertex of G, then $G - w$ is a graph obtained by deleting the vertex w and all the edges incident with w. Similarly, if W is a set of vertices of G, then $G - W$ is the graph obtained by deleting all vertices of W and edges incident with them. Finally, if F is a set of its edges, then $G \setminus F$ is the graph obtained from G by removing the edges of F but none of the vertices.

A graph with all vertices of degree at most three is called *subcubic*. We say that a graph G is *k-connected* if it has at least $k + 1$ vertices and $G - W$ is connected for any $W \subseteq V(G)$ containing at most $k - 1$ vertices (here, we deviate from the standard terminology since we do not consider K_2 to be 2-connected). If G is connected but not 2-connected, then a vertex v such that $G - v$ is not connected is called a *cut-vertex*. Maximal 2-connected subgraphs of G and edges that are not contained in any 2-connected subgraphs of G are called *blocks*. A subset F of the edges of a graph G is an *edge-cut* if the graph $G \setminus F$ have more components than G and F is minimal with this property. Such a subset F containing exactly k edges will also be referred to as *k-edge-cut*. An edge forming a 1-edge-cut is called a *cut-edge*. A graph G is *k-edge-connected* if it has no ℓ-edge-cut for $\ell < k$. Note that a subcubic graph G with at least two vertices is 2-connected if and only if it is 2-edge-connected.

A *θ-graph* is a simple graph obtained from the pair of vertices joined by three parallel edges by subdividing some of the edges several times. In other words, a θ-graph is a graph that contains two vertices of degree three joined by three paths formed by vertices of degree two such that at most one of these paths is trivial, i.e., it is a single edge. In our consideration, we will need to consider a special type of cycles of length six in subcubic graphs, which

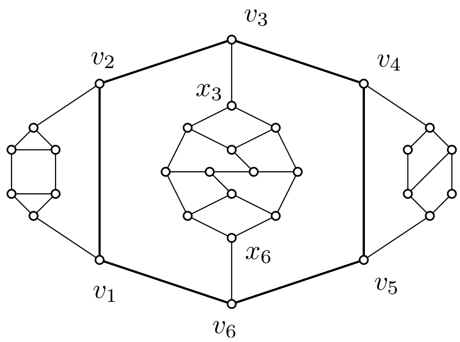

Figure 1 A θ-cycle with poles v_3 and v_6.

resembles θ-graphs. A cycle $K = v_1 \dots v_6$ of length six in a subcubic graph G is a *θ-cycle*, if all vertices $v_1, \dots, v_6$ have degree three, their neighbors $x_1, \dots, x_6$ outside of K are pairwise distinct, and $G - V(K)$ has three connected components, one containing x_1 and x_2, one containing x_4 and x_5, and one containing x_3 and x_6. See Figure 1 for an example. The vertices v_3 and v_6 of the cycle K will be referred to as the *poles* of the θ-cycle K.

We say that a multigraph is *Eulerian* if all its vertices have even degree; note that we do not require the multigraph to be connected, i.e., a multigraph has an Eulerian tour if and only if it is Eulerian and connected. A subgraph is *spanning* if it contains all vertices of the original graphs, possibly some of them as isolated vertices, i.e., vertices of degree zero. It is easy to relate the length of the shortest TSP walk in a graph G to the size of Eulerian multigraphs using edges of G as follows. To simplify our presentation, let $\mathrm{tsp}(G)$ denote the length of the shortest TSP walk in a graph G. The proof of the next observation is omitted because of the space constraints.

▶ **Observation 4.** *For every graph G, $\mathrm{tsp}(G)$ is equal to the minimum number of edges of a connected Eulerian multigraph H such that the underlying graph of H is a spanning subgraph of G.*

We now explore the link between Eulerian spanning subgraphs and the minimum length of a TSP walk further. For a graph G, let $c(F)$ denote the number of non-trivial components of F, i.e., components formed by two or more vertices, and let $i(F)$ be the number of isolated vertices of F. We define the *excess* of a graph F as

$$\mathrm{exc}(F) = 2c(F) + i(F).$$

If G is a subcubic graph, we define

$$\mathrm{minexc}(G) = \min \{\mathrm{exc}(F) : F \text{ spanning Eulerian subgraph of } G\}.$$

Note that any subcubic Eulerian graph F is a union of $c(F)$ cycles and $i(F)$ isolated vertices, i.e., the spanning subgraph F of a subcubic graph G with $\mathrm{exc}(F) = \mathrm{minexc}(G)$ must also have this structure. The values of $\mathrm{minexc}(G)$ for simple-structured graphs are given in the next observation.

▶ **Observation 5.** *The following holds.*
1. *If G is a cycle, then* $\mathrm{minexc}(G) = 2 < \frac{n(G)+n_2(G)}{4} + 1$.
2. *If $G = K_4$, then* $\mathrm{minexc}(G) = 2 = \frac{n(G)+n_2(G))}{4} + 1$.

3. *If G is θ-graph with k_1, k_2 and k_3 vertices of degree two on the paths joining its two vertices of degree three and $k_1 \leq k_2 \leq k_3$ (note that $k_2 \neq 0$ by the definition of a θ-graph),* $\mathrm{minexc}(G) = 2 + k_1 \leq \frac{n(G)+n_2(G)}{4} + 1$.

We next relate the quantity $\mathrm{minexc}(G)$ to the length of the shortest TSP walk in G.

▶ **Observation 6.** *Let G be a connected subcubic n-vertex graph, and let F be a spanning Eulerian subgraph F of G. There exists a polynomial-time algorithm that finds a TSP walk of length $n-2+\mathrm{exc}(F)$. In addition, the minimum length of a TSP walk in G is equal to*

$$\mathrm{tsp}(G) = n - 2 + \mathrm{minexc}(G) .$$

Proof. Let F be a spanning Eulerian subgraph of G. We aim to construct a TSP walk of length $n-2+\mathrm{exc}(F)$. The subgraph F has $c(F)+i(F)$ components. Since F is subcubic, each of the $c(F)$ non-trivial components of F is a cycle, which implies that F has $n-i(F)$ edges. Since G is connected, there exists a subset S of the edges of G such that $|S| = c(F)+i(F)-1$ and F together with the edges of S is connected. Clearly, such a subset S can be found in linear time. Let H be the multigraph obtained from F by adding each edge of S with multiplicity two. Since H is a connected Eulerian multigraph whose underlying graph is a spanning subgraph of G, the proof of Observation 4 yields that it corresponds to an Eulerian tour of length

$$|E(H)| = |E(F)| + 2|S| = n - i(F) + 2(c(F) + i(F) - 1) = n - 2 + \mathrm{exc}(F),$$

which can be found in linear time. In particular, it holds that $\mathrm{tsp}(G) \leq n-2+\mathrm{exc}(F)$. Since the choice of F was arbitrary, we conclude that $\mathrm{tsp}(G) \leq n-2+\mathrm{minexc}(G)$.

To finish the proof, we need to show that $n-2+\mathrm{minexc}(G) \leq \mathrm{tsp}(G)$. By Observation 4, there exists a connected Eulerian multigraph H with $|E(H)| = \mathrm{tsp}(G)$ such that its underlying graph is a spanning subgraph of G. By the minimality of $|E(H)|$, every edge of H has multiplicity at most two (otherwise, we can decrease its multiplicity by 2 while keeping the multigraph Eulerian and connected). Similarly, removing any pair of parallel edges of H disconnects H (as the resulting multigraph would still be Eulerian), i.e., the edge in the underlying graph of H corresponding to a pair of parallel edges is a cut-edge. Let F be the graph obtained from H by removing all the pairs of parallel edges. The number of components of F is equal to

$$c(F) + i(F) = \frac{|E(H)| - |E(F)|}{2} + 1.$$

Since F is subcubic, it is a union of $c(F)$ cycles and $i(F)$ isolated vertices, which implies that $|E(F)| = n - i(F)$. Consequently, we get that

$$c(F) + i(F) = \frac{|E(H)| - (n - i(F))}{2} + 1,$$

which yields the desired inequality

$$n - 2 + \mathrm{minexc}(G) \leq n - 2 + \mathrm{exc}(F) = n - 2 + 2c(F) + i(F) = |E(H)| = \mathrm{tsp}(G). \qquad ◀$$

3 Reductions

In this section, we present a way of reducing a 2-connected subcubic graph to a smaller one such that a spanning Eulerian subgraph of the smaller graph yields a spanning Eulerian subgraph of the original graph with few edges. We now define this process more formally. For subcubic graphs G and G', let

$$\delta(G, G') = (n(G) + n_2(G)) - (n(G') + n_2(G')) \ .$$

We say that a 2-connected subcubic graph G' is a reduction of a 2-connected subcubic graph G if $n(G') < n(G)$, $\delta(G, G') \geq 0$, and there exists a linear-time algorithm that turns any spanning Eulerian subgraph F' of G' into a spanning Eulerian subgraph F of G satisfying

$$\mathrm{exc}(F) \leq \mathrm{exc}(F') + \tfrac{1}{4} \cdot \delta(G, G'). \tag{1}$$

For the proof of our main result, it would be enough to prove the lemmas in this section with $\frac{1}{4}$ replaced by $\frac{2}{7}$ in (1). However, this would not simplify most of our arguments and we believe that the stronger form of (1) can be useful in an eventual proof of Conjecture 3.

The reductions that we present are intended to be applied to an input subcubic 2-connected graph until the resulting graph is simple or it has a special structure. A subcubic 2-connected graph is *basic* if it is a cycle, a θ-graph, or K_4. A subcubic 2-connected graph that is not basic will be referred to as *non-basic*. The reductions involve altering a subgraph K of a graph G such that K has some additional specific properties. This subgraph sometimes needs to be provided as a part of an input of an algorithm that constructs G'. We say that a reduction is a *linear-time reduction with respect to a subgraph K* if there exists a linear-time algorithm that transforms G to G' given G and a subgraph K with the specific properties. We will say that a reduction is a *linear-time reduction* if there exists a linear-time algorithm that both finds a suitable subgraph K and performs the reduction. If a graph G admits such a reduction, we will say that G has a linear-time reduction or that G has a linear-time reduction with respect to a subgraph K.

As an example of our reductions, we state and prove a lemma describing the simplest among our reductions; we leave the details of all other reductions because of the space constraints.

▶ **Lemma 7.** *Every non-basic 2-connected subcubic graph G that contains a cycle K with at most two vertices of degree three has a linear-time reduction.*

Proof. Since G is neither a cycle nor a θ-graph, it follows that $V(G) \neq V(K)$. Since G is 2-connected, K contains exactly two vertices of degree three, say v_1 and v_2. Let x_1 and x_2 be their neighbors outside of K, and let k_1 and k_2 be the the number of the internal vertices of the two paths between v_1 and v_2 in K. We can assume that $k_1 \leq k_2$ by symmetry. If $x_1 = x_2$, then either G is a θ-graph or x_1 is incident with a cut-edge; since neither of these is possible, it holds that $x_1 \neq x_2$.

Suppose that $k_1 = 0$ and $k_2 = 1$, i.e., K is a triangle. Let z be the vertex of K distinct from v_1 and v_2, and let $G' = G - z$. Note that G' is a 2-connected subcubic graph. We claim that G' is a reduction of G. Since $n(G') = n(G) - 1$ and $n_2(G') = n_2(G) + 1$, it follows $\delta(G, G') = 0$. Consider a spanning Eulerian subgraph F' of G'. If F' contains the edge v_1v_2, then let F be the spanning Eulerian subgraph of G obtained from F' by removing the edge v_1v_2 and adding the path v_1zv_2. If F' does not contain the edge v_1v_2, i.e., v_1 and v_2 are isolated vertices of F, then let F be the spanning Eulerian subgraph of G obtained from F' by adding the cycle K. It holds that $\mathrm{exc}(F) = \mathrm{exc}(F')$ in both cases.

It remains to consider the case $k_1 + k_2 \geq 2$. Let G' be obtained from $G - V(K)$ by adding a path x_1wx_2 where w is a new vertex; note that w has degree two in G' and $\delta(G, G') = 2(k_1 + k_2)$. Since $x_1 \neq x_2$, G' is simple. We show that G' is a reduction of G. Let F' be a spanning Eulerian subgraph of G'; we will construct a spanning Eulerian subgraph

F of G. If F' contains the path x_1wx_2, then let F be obtained from $F'-w$ by adding the vertices of K and the edges x_1v_1, x_2v_2, and the path in K between v_1 and v_2 with k_2 internal vertices. Note that the k_1 vertices of the other path between v_1 and v_2 in K are isolated in F. Observe that

$$\text{exc}(F) = \text{exc}(F') + k_1 \leq \text{exc}(F') + \frac{k_1+k_2}{2},$$

since $k_1 \leq k_2$. If w is an isolated vertex of F', then let F be obtained from $F'-w$ by adding the cycle K. In this case, we get that

$$\text{exc}(F) = \text{exc}(F') + 1 \leq \text{exc}(F') + \frac{k_1+k_2}{2}.$$

Since it holds that $\text{exc}(F) \leq \text{exc}(F) + \frac{1}{4}\delta(G,G')$ in both cases, the proof of the lemma is finished. ◀

We next summarize the facts that can be established using our reduction techniques; the details are omitted because of the space constraints. We say that a 2-connected subcubic graph G is a *proper* graph if G is non-basic, has no cycle with at most four vertices of degree three, and has no cycle of length five or six with five vertices of degree three. We will call a non-basic 2-connected subcubic graph G clean if none of the lemmas that we have proven can be applied to G. Formally, a 2-connected subcubic graph G is *clean* if it is proper and

(CT1) no cycle of length at most 7 in G contains a vertex of degree two,

(CT2) every cycle of length six in G that is not a θ-cycle is disjoint from all other cycles of length six,

(CT3) every cycle $K = v_1 \dots v_m$ of length $m \leq 7$ in G satisfies that if each of the edges v_1v_m and v_2v_3 is contained in a 2-edge-cut, then the edges v_1v_m and v_2v_3 themselves form a 2-edge-cut, and

(CT4) every cycle $K = v_1 \dots v_6$ of length six in G satisfies at least one of the following

- **(a)** K is a θ-cycle, or
- **(b)** each edge exiting K is contained in a 2-edge-cut but no two of them together form a 2-edge-cut, or
- **(c)** each edge exiting K is contained in a 2-edge-cut, and there exists exactly one pair i and j with $1 \leq i < j \leq 6$ such that the edges v_ix_i and v_jx_j form a 2-edge-cut, and this pair satisfies $j - i = 3$, or,
- **(d)** precisely one edge exiting K, say v_1x_1, is not contained in a 2-edge-cut, and there exists a partition A and B of the vertices of $G - V(K)$ such that $x_1, x_2, x_6 \in A$, $x_3, x_4, x_5 \in B$, there is exactly one edge between A and B, and both A and B induce connected subgraphs of $G - V(K)$,

where x_i is the neighbor of the vertex v_i outside the cycle K, $i \in \{1, \dots, 6, \}$.

The next theorem summarizes our reduction results.

▶ **Theorem 8.** *There exists an algorithm running in time $O(n^3)$ that constructs for a given n-vertex 2-connected subcubic graph G a reduction of G that is either basic or clean.*

4 Main result

We need few additional results before we can prove Theorem 1. The first concerns the structure of cycles passing through vertices of a cycle of length six in a clean 2-connected subcubic graph. Let v be a vertex of degree three in a graph G, and let x_1, x_2 and x_3 be its neighbors. The *type* of v is the triple (ℓ_1, ℓ_2, ℓ_3) such that ℓ_1, ℓ_2 and ℓ_3 are the lengths of shortest cycles containing paths x_1vx_2, x_1vx_3 and x_2vx_3. In our consideration, the order of the coordinates of the triple will be irrelevant, so we will always assume that the lengths satisfy that $\ell_1 \le \ell_2 \le \ell_3$. A type $(\ell'_1, \ell'_2, \ell'_3)$ *dominates* the type (ℓ_1, ℓ_2, ℓ_3) if $\ell'_i \ge \ell_i$ for every $i = 1, 2, 3$. If K is a cycle in a graph G and each vertex of K has degree three, then the *type* of the cycle K is the multiset of the types of the vertices of K. Finally, a multiset M_1 of types *dominates* a multiset M_2 types if there exists a bijection between the types contained in M_1 and M_2 such that each type of M_1 dominates the corresponding type in M_2.

We can now show the following lemma (note that all vertices of the cycle K in the lemma must have degrees three since G is assumed to be clean). The proof is omitted because of the space constraints.

▶ **Lemma 9.** *Let G be a clean 2-connected subcubic graph and let $K = v_1v_2 \dots v_6$ be a cycle of length six in G. If K is not a θ-cycle, then the type of K dominates at least one of the following multisets:*

- $\{(6,7,7),\ (6,7,7),\ (6,8,8),\ (6,8,8),\ (6,8,8),\ (6,8,8)\}$,
- $\{(6,7,7),\ (6,7,8),\ (6,7,8),\ (6,8,8),\ (6,8,8),\ (6,8,8)\}$, *or*
- $\{(6,7,7),\ (6,7,8),\ (6,7,9),\ (6,7,9),\ (6,8,8),\ (6,8,8)\}$.

The following lemma follows from the description of the perfect matching polytope by Edmonds [6] and the fact that the perfect matching polytope has a strong separation oracle [17]; see e.g. [8] for further details.

▶ **Lemma 10.** *There exists a polynomial-time algorithm that for a given cubic 2-connected n-vertex graph outputs a collection of $m \le n/2 + 2$ perfect matchings $M_1, \dots, M_m$ and non-negative coefficients $a_1, \dots, a_m$ such that $a_1 + \dots + a_m = 1$ and*

$$\sum_{i=1}^{m} a_i \chi_{M_i} = (1/3, \dots, 1/3) \in \mathbb{R}^{E(G)} \ ,$$

where $\chi_{M_i} \in \mathbb{R}^{E(G)}$ is the characteristic vector of M_i.

Lemma 10 gives the following; we note that variants of Lemma 11 have also been used in [2, 7, 14].

▶ **Lemma 11.** *There exists a polynomial-time algorithm that for a given 2-connected n-vertex subcubic graph outputs a collection of $m \le n/2 + 2$ spanning Eulerian subgraphs $F_1, \dots, F_m$ and probabilities $p_1, \dots, p_m \ge 0$, $p_1 + \dots + p_m = 1$ that satisfy the following. If a spanning Eulerian subgraph F is equal to F_i with probability p_i, $i = 1, \dots, m$, then $\mathbb{P}[e \in E(F)] = 2/3$. In particular, a vertex of degree three is contained in a cycle of F with probability one and a vertex of degree two is isolated with probability $1/3$.*

We now combine Lemmas 9 and 11 to get the following.

▶ **Lemma 12.** *There exists a polynomial-time algorithm that given a clean 2-connected subcubic graph G outputs a spanning Eulerian subgraph F of G such that*

$$\mathrm{exc}(F) \le \frac{2n(G) + 2n_2(G)}{7} \ .$$

Proof of Lemma 12. We first apply the algorithm from Lemma 11 to get a collection of $m \le n/2 + 2$ spanning Eulerian subgraphs $F_1, \dots, F_m$ and probabilities $p_1, \dots, p_m$. We show that

$$\mathbb{E}\left[\mathrm{exc}(F)\right] \le \frac{2n(G) + 2n_2(G)}{7} \ , \tag{2}$$

which implies the statement of the lemma since the number of the subgraphs $F_1, \ldots, F_m$ is linear in n and the excess of each them can be computed in linear time. In particular, the algorithm can output the subgraph F_i with the smallest $\mathrm{exc}(F_i)$.

We now show that (2) holds. We apply a double counting argument, which we phrase as a discharging argument. At the beginning, we assign each vertex of degree three charge of $2/7$ and to each vertex of degree two charge of $4/7$. Let $c_1(v)$ be the initial charge of a vertex v. Note that the sum of the initial charges of the vertices is the right side of the inequality (2).

We next choose a random spanning Eulerian subgraphs F among the subgraphs $F_1, \ldots, F_m$ with probabilities given by $p_1, \ldots, p_m$. The charge of each vertex that is isolated in F is decreased by one unit, and the charge of each vertex contained in a cycle of length k by $2/k$ units. Let $c_2(v)$ be the new charge of a vertex v. Observe that the total decrease of charge of the vertices is equal to $\mathrm{exc}(F)$, i.e.,

$$\mathrm{exc}(F) = \sum_{v \in V(G)} c_1(v) - c_2(v) \; .$$

Hence, it is enough to prove that

$$\mathbb{E}\left[\sum_{v \in V(G)} c_2(v)\right] \geq 0. \tag{3}$$

To prove (3), we consider the expectation of $c_2(v)$ for individual vertices v of G.

If v is a vertex of G of degree two, then every cycle of G that contains v has length at least eight by (CT1). With probability $1/3$, the vertex v is isolated and looses one unit charge; with probability $2/3$, it is contained in a cycle and looses at most $2/8 = 1/4$ units of charge. We conclude that

$$\mathbb{E}\left[c_2(v)\right] \geq \frac{4}{7} - \frac{1}{3} - \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{14} > 0 \; .$$

If v is a vertex of G of degree three with type (ℓ_1, ℓ_2, ℓ_3), we proceed as follows. Since each edge incident with v is contained in F with probability $2/3$, v is contained in a cycle of F with a particular pair of its neighbors with probability $1/3$. It follows that the expected value of $c_1(v)$ is at least

$$\mathbb{E}\left[c_2(v)\right] \geq \frac{2}{7} - \frac{1}{3}\left(\frac{2}{\ell_1} + \frac{2}{\ell_2} + \frac{2}{\ell_3}\right) \; .$$

Since G is clean, the type of v dominates $(6, 6, 6)$. If the type of v dominates $(7, 7, 7)$, then $\mathbb{E}\left[c_2(v)\right] \geq 0$. Hence, we focus on vertices contained in cycles of length six in G in the rest of the proof.

Let $K = v_1 \ldots v_6$ be a cycle of length six in G. Since G is clean, each vertex of K has degree three. Suppose that K is not a θ-cycle. By (CT2), K is disjoint from all other cycles of length six in G. Observe that

- if the type of v_i dominates $(6, 8, 8)$, then $\mathbb{E}\left[c_2(v_i)\right] \geq \frac{1}{126}$,
- if the type of v_i dominates $(6, 7, 7)$, then $\mathbb{E}\left[c_2(v_i)\right] \geq -\frac{1}{63}$,
- if the type of v_i dominates $(6, 7, 8)$, then $\mathbb{E}\left[c_2(v_i)\right] \geq -\frac{1}{252}$, and
- if the type of v_i dominates $(6, 7, 9)$, then $\mathbb{E}\left[c_2(v_i)\right] \geq \frac{1}{189}$.

Since the type of the cycle K dominates one of the three multisets listed in Lemma 9, it holds that

$$\mathbb{E}\left[c_2(v_1) + \cdots + c_2(v_6)\right] \geq 0 \; .$$

It remains to analyze the case that K is a θ-cycle. By symmetry, we can assume that the vertices v_1 and v_4 are its poles. Let x_i be the neighbor of v_i outside of K, $i = 1, \ldots, 6$. Further, let $P = x_6v_6v_1v_2x_2$, $P_1 = x_6v_6v_1$ and $P_2 = x_2v_2v_1$. Since each of the paths P_1 and P_2 is contained in F with probability $1/3$, the subgraph F contains the path P with probability at most $1/3$; let p be this probability. Since G is clean (and so proper), the distance between x_2 and x_3 in $G - V(K)$ is at least three; likewise, the distance between x_5 and x_6 in $G - V(K)$ is at least three. Hence, any cycle containing P_1 or P_2 has length at least 10, and any cycle containing P has length at least 14. Since F contains the path P with probability p, the path P_1 but not P with probability $1/3 - p$, the path P_2 but not P with probability $1/3 - p$, and neither P_1 nor P_2 with probability $1/3 + p$, it follows that

$$\mathbb{E}\left[c_2(v_1)\right] = \frac{2}{7} - p \cdot \frac{1}{7} - 2\left(\frac{1}{3} - p\right) \cdot \frac{1}{5} - \left(\frac{1}{3} + p\right) \cdot \frac{1}{3} = \frac{13}{315} - \frac{8}{105}p \geq \frac{1}{63} \; .$$

The symmetric argument yields that $\mathbb{E}\left[c_2(v_4)\right] \geq \frac{1}{63}$. Since every cycle in G containing the path P_2 has length at least 10, the type of v_2 dominates $(6, 6, 10)$ and thus $\mathbb{E}\left[c_2(v_2)\right] \geq -\frac{1}{315}$. The same holds for vertices v_3, v_5 and v_6.

Let Q_1 be the set of all poles of θ-cycles in G, and let Q_2 be the set of vertices contained in θ-cycle that are not a pole of a (possibly different) θ-cycle. Since each vertex of Q_2 has a neighbor in Q_1, it follows $|Q_2| \leq 3|Q_1|$. The previous analysis yields that

$$\mathbb{E}\left[\sum_{v \in Q_1 \cup Q_2} c_2(v)\right] \geq |Q_1| \left(\frac{1}{63} - 3 \cdot \frac{1}{315}\right) = \frac{2}{315}|Q_1| \geq 0.$$

Since the set $Q_1 \cup Q_2$ and the vertex set of cycles of length six that are not θ-cycles are disjoint, the inequality (3) follows. ◀

We are ready to prove Theorems 1 and 2.

Proof of Theorem 1. By Observation 6, it is enough to construct a spanning Eulerian subgraph F of G with

$$\mathrm{exc}(F) \leq \frac{2(n(G) + n_2(G))}{7} + 1.$$

If G is basic, such a subgraph F exists by Observation 5, and can easily be constructed in polynomial time. If G is not basic, we can find a reduction G' of G that is either basic or clean in polynomial time by Theorem 8.

If G' is basic, then we find a spanning Eulerian subgraph with

$$\mathrm{exc}(F') \leq \frac{2(n(G') + n_2(G'))}{7} + 1$$

as in the case when G itself is basic. If G' is clean, then Lemma 12 yields that we can construct in polynomial time a spanning Eulerian subgraph F' of G' such that

$$\mathrm{exc}(F') \leq \frac{2(n(G') + n_2(G'))}{7}.$$

Since G' is a reduction of G, we can find in polynomial time a spanning Eulerian subgraph F of G such that

$$\mathrm{exc}(F) \leq \mathrm{exc}(F') + \frac{\delta(G, G')}{4} \leq \mathrm{exc}(F') + \frac{2\delta(G, G')}{7} \leq \frac{2(n(G') + n_2(G'))}{7} + 1 \; ,$$

which finishes the proof of the theorem. ◀

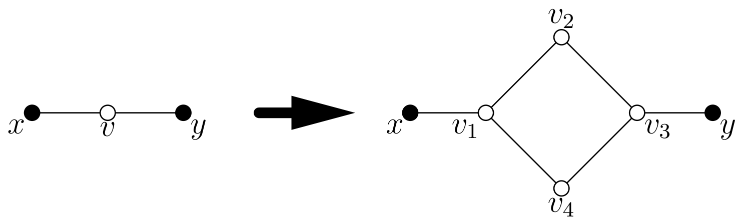

Figure 2 Replacing a vertex of degree two with a cycle of length four in Lemma 13.

Proof of Theorem 2. Let G be an input cubic graph and let n be the number of its vertices. We assume that G is connected since G would not have a TSP walk otherwise. Let F be the set of bridges of G, which can be found in linear time using the standard algorithm based on DFS. Further, let G' be the graph obtained from G by removing the edges of F, and let n_0 and n_2 be the number of its vertices of degree zero and two, respectively. Note that G' has no vertices of degree one since if two edges incident with a vertex v in a cubic graph are bridges, then the third edge incident with v is also a bridge. Finally, let k be the number of non-trivial components of G', i.e., the components of G' that are not formed by a single vertex. Observe that the number of vertices of degree two in G' is at most $2k-2$, i.e., $n_2 \leq 2k-2$.

We next apply the algorithm from Theorem 1 to each non-trivial component of G', and obtain a collection of k TSP walks such that the sum of their lengths is at most

$$\frac{9}{7}(n-n_0)+\frac{2}{7}n_2-k\ .$$

These k TSP walks can be connected by traversing each of the edges of F twice, which yields a TSP walk in G of total length at most

$$\frac{9}{7}(n-n_0)+\frac{2}{7}n_2-k+2|F|\ \leq\ \frac{9}{7}(n-n_0)+2|F|\ . \tag{4}$$

The inequality in (4) follows from the inequality $n_2 \leq 2k-2$, which we have observed earlier in the proof. Since any TSP walk in G must have length at least $(n-n_0)+2|F|$ (it must contain at least $n-n_0$ edges inside the non-trivial blocks and each bridge must be traversed twice), the upper bound in (4) on the length of the constructed TSP walk is at most the multiple of $9/7$ of the length of the optimal TSP walk in G, which yields the desired approximation factor of the algorithm. ◀

5 Lower bounds

In this section, we provide two constructions of 2-connected subcubic graphs that illustrate that the bound claimed in Conjecture 3 would be the best possible. The constructions are based on two operations that we analyze in Lemmas 13 and 15. The proofs of the lemmas are omitted because of the space constraints.

▶ **Lemma 13.** *Let G be a 2-connected subcubic graph, let v be a vertex of G that has exactly two neighbors, and let x and y be its two neighbors. Further, let G' be the graph obtained from G by removing the vertex v, adding a cycle $v_1v_2v_3v_4$ and edges xv_1 and yv_3 as in Figure 2. The graph G' is a 2-connected subcubic graph and it holds that $n(G')=n(G)+3$, $n_2(G')=n_2(G)+1$ and $\operatorname{minexc}(G')=\operatorname{minexc}(G)+1$.*

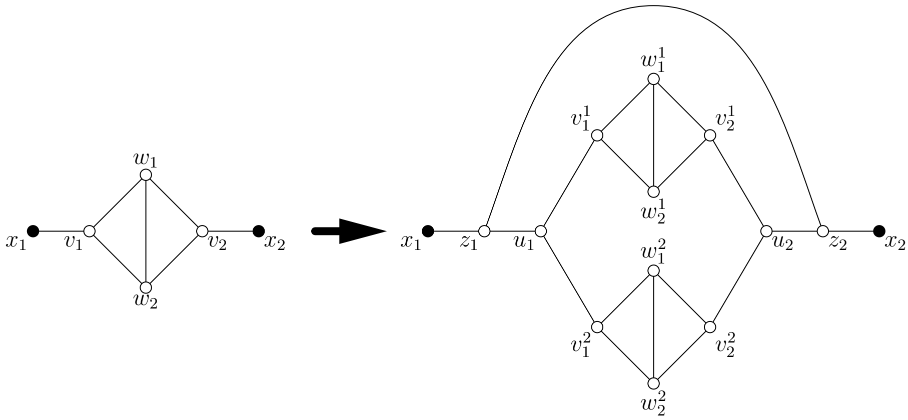

Figure 3 The operation of replacing a diamond analyzed in Lemma 15.

Repeated applications of the operation described in Lemma 13 starting with the graph $K_{2,3}$ yields the following.

▶ **Proposition 14.** *For every integer* $n \geq 5$, $n \equiv 2 \bmod 3$, *there exists a* 2*-connected subcubic* n*-vertex graph* G *such that*

$$\mathrm{minexc}(G) = \frac{n(G) + n_2(G)}{4} + 1.$$

The second operation is more involved. A *diamond* in a graph G is an induced subgraph isomorphic to K_4^-, i.e., the graph K_4 with one edge removed.

▶ **Lemma 15.** *Let* G *be a 2-connected cubic graph containing a diamond* D*. Let* v_1, v_2, w_1 *and* w_2 *be the vertices of the diamond as depicted in Figure 3, and let* x_1 *and* x_2 *be the neighbors of* v_1 *and* v_2 *outside of the diamond* D*. Further, let* G' *be the graph obtained from* G *by removing the vertices of the diamond* D *and inserting the subgraph depicted in Figure 3. The graph* G' *is a 2-connected cubic graph with* $n(G') = n(G)+8$ *and* $\mathrm{minexc}(G') = \mathrm{minexc}(G)+2$*. Moreover, the graph* G' *contains at least two diamonds.*

Consider the cubic graph formed by two diamonds and two edges joining the vertices of degree two in different diamonds, and repeatedly apply the operation described in Lemma 15.

▶ **Proposition 16.** *For every integer* $n \geq 8$, $n \equiv 0 \bmod 8$, *there exists a* 2*-connected cubic* n*-vertex graph* G *with* $\mathrm{minexc}(G) = n/4$.

Propositions 14 and 16, and Observation 6 yield that neither the coefficient 5/4 nor the coefficient 1/4 in Conjecture 3 can be improved. Indeed, for every $\alpha < 5/4$, there exist infinitely many 2-connected cubic graphs G with $\mathrm{tsp}(G) > \alpha n(G) + o(n(G))$ by Proposition 16. Likewise, for every $\beta < 1/4$, there exist infinitely many 2-connected subcubic graphs G with $\mathrm{tsp}(G) > \frac{5}{4}n(G) + \beta n_2(G) + o(n(G))$. While neither of the two coefficients in Conjecture 3 can be improved, it may be possible to prove a stronger bound that is not linear in both $n(G)$ and $n_2(G)$.

Acknowledgements. The authors would like to thank the anonymous reviewers for their valuable and insightful comments on the submission.

References

1 Nishita Aggarwal, Naveen Garg, and Swati Gupta. A 4/3-approximation for TSP on cubic 3-edge-connected graphs. *arXiv e-prints*, 1101.5586, 2011.

2 Sylvia Boyd, René Sitters, Suzanne van der Ster, and Leen Stougie. The traveling salesman problem on cubic and subcubic graphs. *Mathematical Programming*, 144(1-2):227–245, 2014.

3 Barbora Candráková and Robert Lukoťka. Cubic TSP – a 1.3-approximation. *arXiv e-prints*, 1506.06369, 2015.

4 Nicos Christofides. Worst-case analysis of a new heuristic for the travelling salesman problem. Technical report, DTIC Document, 1976.

5 José Correa, Omar Larré, and José A Soto. TSP tours in cubic graphs: beyond 4/3. *SIAM Journal on Discrete Mathematics*, 29(2):915–939, 2015.

6 Jack Edmonds. Maximum matching and a polyhedron with 0, l-vertices. *J. Res. Nat. Bur. Standards B*, 69(1965):125–130, 1965.

7 David Gamarnik, Moshe Lewenstein, and Maxim Sviridenko. An improved upper bound for the TSP in cubic 3-edge-connected graph. *Operations Research Letters*, 33:467–474, 2005.

8 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

9 Jeremy Karp and R Ravi. A 9/7-approximation algorithm for graphic TSP in cubic bipartite graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*, volume 28 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 284–296. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, 2014. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.284.

10 Marek Karpinski, Michael Lampis, and Richard Schmied. New inapproximability bounds for TSP. *Journal of Computer and System Sciences*, 81:1665–1677, 2015.

11 Marek Karpinski and Richard Schmied. Approximation hardness of graphic TSP on cubic graphs. *RAIRO Operations Research*, 49:651–668, 2015.

12 M. Lampis. Improved inapproximability for TSP. *Lecture Notes in Computer Science*, 7408:243–253, 2013.

13 Jan Mazák and Robert Lukoťka. Simple cubic graphs with no short travelling salesman tour, 2016. Manuscript.

14 Tobias Mömke and Ola Svensson. Approximating graphic TSP by matchings. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 560–569. IEEE, 2011.

15 Marcin Mucha. 13/9-approximation for graphic TSP. *Theory of Computing Systems*, 55:640–657, 2014.

16 Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. A randomized rounding approach to the traveling salesman problem. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 550–559. IEEE, 2011.

17 Manfred W. Padberg and M. R. Rao. Odd minimum cut-sets and b-matchings. *Mathematics of Operations Research*, 7(1):67–80, 1982.

18 Christos H. Papadimitriou and Santosh Vempala. On the approximability of the traveling salesman problem. *Combinatorica*, 26:101–120, 2006.

19 András Sebö 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*, 34:597–629, 2014.

20 Anke van Zuylen. Improved approximations for cubic and cubic bipartite TSP. In *Integer Programming and Combinatorial Optimization (IPCO)*, volume 9682 of *Lecture Notes in Computer Science*, pages 250–261. Springer, 2016.