Finding the exact integrality gap for small Traveling Salesman Problems

Geneviève Benoit
S.I.T.E., University of Ottawa, Ottawa, Canada K1N 6N5
Sylvia Boyd
S.I.T.E., University of Ottawa, Ottawa, Canada K1N 6N5
email: sylvia@site.uottawa.ca `http://www.site.uottawa.ca/~sylvia/`

The Symmetric Traveling Salesman Problem (STSP) is to find a minimum weight Hamiltonian cycle in a weighted complete graph on n nodes. One direction which seems promising for finding improved solutions for the STSP is the study of a linear relaxation of this problem called the Subtour Elimination Problem (SEP). A well-known conjecture in combinatorial optimization says that the integrality gap of the SEP is 4/3 in the metric case. Currently the best upper bound known for this integrality gap is 3/2.

Finding the exact value for the integrality gap for the SEP is difficult even for small values of n due to the exponential size of the data involved. In this paper we describe how we were able to overcome such difficulties and obtain the exact integrality gap for all values of n up to 10 and a tight lower bound for this gap for $11 \leq n \leq 14$. Our results give a verification of the 4/3 conjecture for small values of n, and also give rise to a new stronger form of the conjecture which is dependent on n.

Key words: symmetric traveling salesman problem; relaxation ; integrality gap; polyhedron.

MSC2000 Subject Classification: Primary: 90C27; Secondary: 90C57.

OR/MS subject classification: Primary: Networks/graphs, traveling salesman.

---

1. Introduction. Given the complete graph $K_n = (V, E)$ on n nodes with non-negative edge costs $c \in \mathbf{R}^E$, $c \neq 0$, the *Symmetric Traveling Salesman Problem* (henceforth STSP) is to find a Hamiltonian cycle (or *tour*) in K_n of minimum cost. When the costs satisfy the triangle inequality, i.e. when $c_{ij} + c_{jk} \geq c_{ik}$ for all $i, j, k \in V$, we call the problem the *metric* STSP.

For any edge set $F \subseteq E$ and $x \in \mathbf{R}^E$, let $x(F)$ denote the sum $\sum_{e \in F} x_e$. For any node set $W \subset V$, let $\delta(W)$ denote $\{uv \in E : u \in W, v \notin W\}$. Let $\mathbb{S} = \{S \subset V, 3 \leq |S| \leq n-3\}$. An integer linear programming (ILP) formulation for the STSP is as follows:

$$\begin{aligned} & \text{minimize } cx & & (1)\\ \text{subject to: } & x(\delta(v)) = 2 & \text{for all } v \in V, & \quad (2)\\ & x(\delta(S)) \geq 2 & \text{for all } S \in \mathbb{S}, & \quad (3)\\ & 0 \leq x_e \leq 1 & \text{for all } e \in E, & \quad (4)\\ & x & \text{integer.} & \quad (5) \end{aligned}$$

We use $TOUR$ to denote the optimal solution value for ILP (1). If we drop the integer requirement (5) from the above ILP, we obtain a linear programming (LP) relaxation of the STSP called the *Subtour Elimination Problem* (SEP). We use $SUBT$ to denote the optimal solution value for the SEP. The associated polytope, which is denoted by S^n, is the set of all vectors x satisfying the constraints of the SEP, i.e.

$$S^n = \{x \in \mathbf{R}^E : x \text{ satisfies } (2), (3), (4)\}.$$

Note that despite the fact that there are an exponential number of constraints (3), the SEP can be solved in polynomial time since there is an exact polynomial-time separation algorithm for each of its constraints (see Grötschel et. al. [11]).

The STSP is known to be NP-hard, even in the metric case (see Lawler et. al. [12]). One approach taken for finding reasonably good solutions is to look for a k-approximation algorithm for the problem, i.e. try to find a heuristic for the STSP which finds a tour which is guaranteed to be of cost at most $k*(TOUR)$ for some constant $k \geq 1$. Currently the best k-approximation algorithm known for the metric STSP is the algorithm due to Christofides [10] for which $k = 3/2$. Surprisingly, no one has been able to improve upon this algorithm in well over two decades. Note that for general costs there does not exist a k-approximation algorithm unless $P = NP$ (see Lawler et. al. [12]).

A related approach for finding improved STSP solutions is to study the *integrality gap* α for the SEP, which is the worst-case ratio between $TOUR$ and $SUBT$, i.e.

$$\alpha = \max_{c\geq 0}\left(\frac{TOUR}{SUBT}\right).$$

Value α gives one measure of the quality of the lower bound provided by the SEP for the STSP. Moreover, a constructive proof for value α would provide an α-approximation algorithm for the STSP.

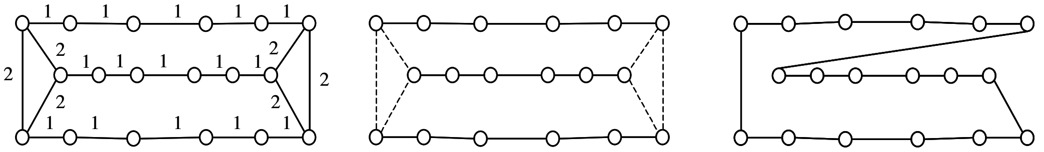
It is known that for the metric STSP, α is at most 3/2 (see Shmoys and Williamson [14], Wolsey [15]), however no example for which this ratio comes close to 3/2 has yet been found. In fact, a well-known conjecture states the following:

Conjecture 1.1 *For the metric STSP, the integrality gap for the SEP is* 4/3.

If Conjecture 1.1 is true, then it is best possible, for consider the metric STSP example with costs c shown in Figure 1(a). In the figure, imagine the three horizontal paths each have length k, and let the cost c_{uv} for edges $uv \in E$ not shown be the cost of a cheapest u to v path in the figure. The optimal solution x^* for the SEP for this set of costs is shown in Figure 1(b), where $x^*_e = 1/2$ for the dashed edges e, 1 for the solid edges e, and 0 for all other edges. The optimal tour is shown in Figure 1(c). Thus this set of costs gives the ratio

$$\frac{TOUR}{SUBT} = \frac{4\text{k} + 6}{3\text{k} + 6}$$

which tends to 4/3 as $k \rightarrow \infty$.



(a) (b) (c)

Figure 1: An example for which $\alpha = 4/3$.

Note that not much progress has been made on Conjecture 1.1 in the past decade. It has been verified for a special class of cost functions by Boyd and Carr [2], and recent results which are implied by the correctness of the conjecture can be found in Boyd and Carr [1] and Carr and Ravi [6]. A similar conjecture of 4/3 for the integrality gap of the LP relaxation of the asymmetric Traveling Salesman problem was made by Carr and Vempala [7], however it was shown in Charikar et. al. [8] and later also in Boyd and Elliott-Magwood [3] that this gap is at least 2.

In this paper we examine the problem of finding the exact integrality gap for the metric SEP for fixed values of n where n is small. Note that many models for this problem are too complex and too large to be practical to solve. In Section 2 we describe how we were able to overcome such difficulties by instead solving a series of reasonably-sized LPs based on a related problem for a subset of the vertices of the subtour elimination polytope S^n. In Section 3 we describe how we found the necessary set of vertices for our method, and also describe some new structural properties for the vertices of S^n which we needed to use in order to make our ideas work. In Section 4 we report our results, i.e. we give the exact integrality gap for all values of n up to 10 found using our method. These results not only provide a verification of Conjecture 1.1 for $n \leq 10$, but also give rise to a new stronger form of the conjecture we propose which is dependent on n. Such a formulation of the conjecture could be useful in an inductive proof approach for Conjecture 1.1. Finally, in Section 5, we go beyond $n = 10$ and find the integrality gap for $11 \leq n \leq 14$ over all cost functions which are optimized at a half-integer vertex of S^n. The results we find once again support the new stronger conjecture proposed.

2. Finding the integrality gap for fixed n. For the metric STSP, we wish to solve the problem of finding the integrality gap for the SEP when n, the number of nodes in the complete graph $K_n = (V, E)$, is fixed in size. We denote this integrality gap by α_n, i.e.

$$\alpha_n = \max_{c \geq 0}\left(\frac{TOUR}{SUBT}\right)$$

for all metric STSPs on n nodes. Note that it is known that the SEP and the STSP are equivalent problems for $n \leq 5$, and thus $\alpha_n = 1$ for $n \leq 5$ (see Lawler et. al. [12]).

We first note that for a particular metric STSP, if we divide all the edge costs $c \in \mathbf{R}^E$ by the optimum tour value $TOUR$, then the new costs also satisfy the triangle inequality, and the ratio $TOUR/SUBT$ remains unchanged. Note the new value of $TOUR$ using this new cost function will be 1. Thus to solve our problem, it is sufficient to only consider metric cost functions c for which the minimum tour value $TOUR$ is 1, and we have

$$\alpha_n = \max_{\substack{c \geq 0 \\ TOUR=1\ for\ c}} \left(\frac{1}{SUBT}\right),$$

or equivalently,

$$\frac{1}{\alpha_n} = \min_{\substack{c \geq 0 \\ TOUR=1\ for\ c}} (SUBT). \tag{6}$$

This leads to the following quadratic programming model of the problem:

$$\begin{aligned}
&\text{minimize } cx && && (7)\\
\text{subject to: } &x \text{ satisfies constraints (2)-(4)} && && (8)\\
&c(T) \geq 1 && \text{for all 0-1 incidence vectors T of tours of } K_n, && (9)\\
&c_{ij} + c_{jk} \geq c_{ik} && \text{for all } i, j, k \in V, && (10)\\
&c_e \geq 0 && \text{for all } e \in E. && (11)
\end{aligned}$$

Constraints (8) simply ensure that vector x is in the subtour elimination polytope S^n, while constraints (9) and (10) ensure that the minimum tour value is 1 for c and our cost vector c is metric. Note that if the optimal solution for (7) is k, then α_n has value $1/k$. Also note that this model has an exponential number of constraints of types (8) and (9).

As the model above proved impractical to solve even for small n using the tools available to us, we tried several other models in an attempt to get rid of the quadratic objective function. One such model was a binary integer programming model with an explosive number of both constraints and variables. We were able to use this model to find α_n for $n = 6, 7$ and 8, but only through a great deal of exhaustive work used to reduce the problem size by fixing different sets of variables. It was clear that this method was impractical for $n > 8$.

We then found a way of modelling our problem as a series of LPs of a more reasonable size (although still exponential) as follows. Let $\mathbf{X} = \{x_{(1)}, x_{(2)}, \ldots, x_{(p)}\}$ be a complete list of all the vertices of the subtour elimination polytope S^n. We know from polyhedral theory that for every cost function $c \in \mathbf{R}^E$, there exists at least one vertex x^* in $\mathbf{X}$ such that c is minimized over S^n at x^*, i.e. such that $SUBT = cx^*$. This provides a way of breaking our set of cost functions c into different sets (although these sets will not be disjoint). For each vertex $x_{(i)}$, let

$$C_i = \{c \in \mathbf{R}^E : c \text{ is minimized over } S^n \text{ at } x_{(i)}\}. \tag{12}$$

So now we need to solve the following problem for each C_i, where $x_{(i)}$ is fixed and is no longer a variable vector:

$$(\min cx_{(i)} : c \text{ satisfies constraints (9), (10) and (11) }, c \in C_i). \tag{13}$$

If we let OPT_i represent the optimal value of (13), then from (6) we have

$$\frac{1}{\alpha_n} = \min_{1 \le i \le p} (OPT_i). \tag{14}$$

Note that everything in model (13) is linear except the condition that $c \in C_i$. This condition can also be represented by a set of linear constraints using duality theory, as we explain below.

The *dual* LP of the SEP has variables $y \in \mathbf{R}^V$, $u \in \mathbf{R}^E$, and $d \in \mathbf{R}^{\mathbb{S}}$ and is defined as follows:

$$\text{maximize } 2 \cdot y - 1 \cdot u + 2 \cdot d \tag{15}$$

$$\text{subject to: } y_i + y_j - u_{ij} + \sum (d_S : S \in \mathbb{S}, ij \in \delta(S)) \le c_{ij} \qquad \text{for all } ij \in E, \tag{16}$$

$$u_e \ge 0 \qquad \text{for all } e \in E, \tag{17}$$

$$d_S \ge 0 \qquad \text{for all } S \in \mathbb{S}. \tag{18}$$

Given a feasible solution x^* for the SEP, a feasible solution (y, u, d) for the dual LP (15) is said to satisfy the complementary slackness conditions for x^* if it satifies the following:

$$y_i + y_j - u_{ij} + \sum (d_S : S \in \mathbb{S}, ij \in \delta(S)) = c_{ij} \qquad \text{for all } ij \in E \text{ such that } x^*_{ij} > 0, \tag{19}$$

$$u_e = 0 \qquad \text{for all } e \in E \text{ such that } x^*_e < 1, \tag{20}$$

$$d_S = 0 \qquad \text{for all } S \in \mathbb{S} \text{ such that } x^*(\delta(S)) > 2. \tag{21}$$

It follows from the theory of duality and complementary slackness that cost function $c \in C_i$ if and only if there exists a feasible solution (y, u, d) for the dual LP (15) which satisfies the complementary slackness conditions (19), (20) and (21) for $x_{(i)}$. Thus for each C_i we can find OPT_i by solving the following LP, which has variables c and (y, u, d):

$$OPT_i = (\min cx_{(i)} : c \text{ satisfies (9) - (11) , and } (y, u, d) \text{ satisfy (16) - (21) for } x_{(i)}). \tag{22}$$

Note that (22) has an exponential number of variables and constraints. Nevertheless, for small values of n we were able to solve (22) in a reasonable amount of time using the software tool CPLEX for solving LPs.

3. Finding the necessary vertices for S^n.

We need to avoid solving the LP (22) for OPT_i for every vertex $x_{(i)} \in \mathbf{X}$ as the size of $\mathbf{X}$ can be very large even for small n. Luckily we were able to find a much smaller subset of vertices which was sufficient for finding α_n.

Given a vertex $x \in S^n$, the *support graph* $G_x = (V_x, E_x)$ of x is the subgraph of K_n induced by the edge set $E_x = \{e \in E : x_e > 0\}$. We say G_x is the *weighted support graph of* x if we include edge weights x_e for $e \in \mathbf{R}^{E_x}$. Observe that many of the vertices of S^n will have isomorphic weighted support graphs (such as all the vertices representing tours), and thus will have the same objective value OPT_i for LP (22). This observation led to the following steps for reducing the number of vertices we needed to consider:

STEP 1. Generate the set of all the vertices of S^n. To do this we used a software package called PORTA (POlyhedron Representation Transformation Algorithm, see Christof et. al. [9]), a program which, given an LP, generates all the vertices for the corresponding polyhedron.

STEP 2. Reduce the set of vertices from Step 1 by finding the vertices within our set with non-isomorphic weighted support graphs. To do this we used a software package called nauty (see McKay [13]). Note that this package deals with unweighted graphs, but we were able to transform our weighted graphs such that this package could be used for our purposes.

The above method worked for finding α_n for $n = 6, 7$, however for $n = 8$, PORTA was unable to generate all of the vertices for S^n even after running for days. Thus for $n \geq 8$ we needed to exploit some properties of the vertices of S^n in order to aid PORTA in Step 1. These properties are described in Theorems (3.1)-(3.3) below. The proofs of these theorems use ideas similar to those found in Boyd and Pulleyblank [4].

THEOREM 3.1 *Let x be a vertex of S^n. Then the number of edges in the support graph G_x of x is at most $2n - 3$.*

PROOF. This result follows from a stronger result in Boyd and Pulleyblank [4] which says that the number of edges in G_x is at most $2n - k - 3$, where k is the number of nodes in G_x which have degree 3 and for which none of the corresponding incident edges e have value $x_e = 1$. □

THEOREM 3.2 *Consider $x \in \mathbf{R}^E$ such that for some node v we have $x_{uv} = x_{vw} = 1$. Let $\hat{x}$ be the vector indexed by the edge set of $K_n \setminus \{v\}$ defined by*

$$\hat{x}_e = \begin{cases} 1 & \text{if } e = uw, \\ x_e & \text{for all other edges } e. \end{cases} \tag{23}$$

Then for $n \geq 4$, $\hat{x}$ is a vertex of S^{n-1} if and only if x is a vertex of S^n.

PROOF. We begin by showing that $x \in S^n$ if and only if $\hat{x} \in S^{n-1}$.

Suppose that $x \in S^n$. The fact that x satisfies $x(\delta(\{u, v, w\})) \geq 2$ implies that $x_{uw} = 0$. Therefore, since $\hat{x}_{uw} = 1$, the sum of the x_e values for edges incident with u and edges incident with w did not change when $\hat{x}$ was created. Thus the constraints (2) are satisfied by $\hat{x}$ for nodes u and w. For all other constraints (2), (3) and (4) for the SEP, it is clear that if x satisfies them, then so does $\hat{x}$. Hence $\hat{x} \in S^{n-1}$.

Suppose that $\hat{x} \in S^{n-1}$. Vector x is essentially formed from $\hat{x}$ by taking edge uw of value 1 in the weighted support graph of $\hat{x}$ and adding a node v to split it into two edges uv and vw, each with value $x_e = 1$. Thus $x(\delta(v)) = 2$ for vector x. For any $S \in \mathbb{S}$ such that $v \in S$ and $u, w \notin S$, we have $x(\delta(S)) \geq x_{uv} + x_{vw} = 2$, and thus all such cut constraints are satisfied by x. For all other constraints (2), (3) and (4) for the SEP, it is clear that if $\hat{x}$ satisfies them, then so does x. Hence $x \in S^n$.

Now suppose that $x \in S^n$ and x is not a vertex of S^n. This implies that x can be expressed as a convex combination of k distinct points of S^n, $k \geq 2$. For each of these k points x', we must have $x'_{uv} = x'_{vw} = 1$, since $x_{uv} = x_{vw} = 1$. So for each point x' we can form a corresponding point $\hat{x}'$ which will necessarily be in S^{n-1} by the discussion above. By doing this we obtain $\hat{x}$ as a convex combination of k distinct points of S^{n-1}, $k \geq 2$. This shows that $\hat{x}$ is also not a vertex of S^{n-1}.

In a similar manner, we can show that if $\hat{x} \in S^{n-1}$ is not a vertex of S^{n-1} then x is not a vertex of S^n. This completes the proof. □

THEOREM 3.3 *Let x be a vertex of S^n, $n \geq 5$. Then the maximum degree in the support graph G_x of x is $n - 3$.*

PROOF. Suppose the theorem is not true, and for some vertex x of S^n there is a node of degree greater than $n - 3$ in its support graph. Consider a counter example G_x with the number of nodes n as small as possible. We know that the theorem is true for $n = 5$ since all vertices of S^5 correspond to tours. Hence we can assume that $n \geq 6$.

Suppose that G_x has a node v of degree 2, which implies that we have $x_{uv} = x_{vw} = 1$. Thus by Theorem 3.2 we could shrink this vertex x to a vertex $\hat{x}$ of S^{n-1} whose support graph would also contain a node of degree greater than $n-3$. This contradicts the fact that G_x was a smallest counter example.

If G_x has no nodes of degree 2, then

$$2|E(G_x)| = \sum(d_x(v) : v \in V) \geq 3(n-1) + (n-2) = 4n-5,$$

where $d_x(v)$ denotes the degree of node v in G_x. However, by Theorem 3.1 we have that $2|E(G_x)| \leq 4n-6$, so again we have a contradiction. □

For completion we include one more theorem here which is related to Theorem (3.2) and is used in Section 5. Given a vertex x of S^n which does not represent a tour and given its weighted support graph G_x, we call a maximal path of edges of value 1 in G_x a *1-path* of x.

Theorem 3.4 *Let x be a vertex of S^n which does not represent a tour. Then x has at least* 3 *disjoint* 1*-paths.*

Proof. Since x does not represent a tour, we know $n \geq 6$. Suppose the theorem is not true, and let x' be a vertex of S^n which has less than 3 disjoint 1-paths and for which the number of edges in $G_{x'}$ is as small as possible. It is shown in Boyd and Pulleyblank [4] that for any vertex x of S^n, $x_e = 1$ for at least 3 edges $e \in E$. Since x' has at most 2 disjoint 1-paths, this implies that one of these 1-paths must contain two consecutive 1-edges. By Theorem 3.2 we can shrink these two edges into one edge to obtain a vertex $\hat{x}$ of S^{n-1}. Note that this vertex also has at most 2 disjoint 1-paths, contradicting the edge minimality of G_x. □

We used Theorems 3.1 and 3.3 to reduce the work for PORTA by first generating all the non-isomorphic graphs on n nodes that contain $2n-3$ edges and have a maximum degree of $n-3$ (i.e. all the possible non-isomorphic support graphs for vertices of S^n), and then for each of these graphs G' we ran PORTA on the SEP problem constraints with the added constraints $x_e = 0$ for each of the edges e not appearing in G'.

Theorem 3.2 implies that we can obtain any vertex of S^n which has a node of degree 2 in its support graph by subdividing a 1-edge in the support graph of a vertex of S^{n-1}. Thus we obtained all the vertices with a node of degree 2 in the support graph directly from our list of non-isomorphic vertices for S^{n-1}, and only used PORTA to search for vertices for which all nodes of the corresponding support graph have degree at least 3.

To summarize, we found α_n as follows:

STEP 1′. Using nauty, find all the non-isomorphic connected graphs on n nodes that contain $2n-3$ edges, and for which all nodes have a maximum degree of $n-3$ and a minimum degree of 3.

STEP 2′. For each graph G' from Step 1′, use PORTA to find all vertices of the polytope defined by the constraints of the SEP plus the constraints $x_e = 0$ for all edges not appearing in G'. (Note that this polytope is a face of S^n.) Let **A** represent the union of all of these vertices found.

STEP 3′. Find all the vertices of S^n which have a node of degree 2 in the support graph by directly generating them from the list of non-isomorphic vertices created for S^{n-1}. Let **B** represent this set of vertices.

STEP 4′. Use the nauty software package to find the vertices in $\mathbf{A} \cup \mathbf{B}$ with non-isomorphic weighted support graphs. Note that what remains is a complete list of all non-isomorphic weighted support graphs of vertices of S^n.[1,2]

---

[1] The files containing the complete set of non-isomorphic weighted support graphs of vertices for S^n, $6 \leq n \leq 10$ are available at `http://www.site.uottawa.ca/~sylvia/subtourvertices/index.htm`.

[2] Recently Boyd and Xiong [5] also generated all the vertices for $n \leq 10$ using a different direct method not involving PORTA.

STEP 5′. For each vertex $x_{(i)}$ in the set of vertices from Step 4′, solve the LP (22) to find OPT_i. Let the minimum value for OPT_i found be k. Then by (14), α_n is $1/k$.

4. Results and a new stronger conjecture.

Using our method, we were able to find the exact value of α_n for all values of n up to 10. The results are shown in Table 1. In column 2 of the table we list the number of non-isomorphic vertices of S^n we found (this number includes the integer tour vertex). Note that this also represents the number of LPs (22) that we needed to solve for each value of n. Column 3 contains the value of α_n, which is clearly less than 4/3 for all $n \leq 10$.

Table 1: Integrality gap results.

n	**Number of vertices**	α_n
6	2	10/9
7	3	9/8
8	13	8/7
9	56	7/6
10	462	20/17

It surprised us that for each value of n, there was a unique vertex of S^n that gave the maximum ratio α_n. Even more surprising was the definite pattern that these vertices formed. The unique vertex x^* of S^n which gave the maximum ratio for each value of n is shown in Figure 2, where $x^*_e = 1$ for solid edges e, $x^*_e = 1/2$ for dashed edges e, and $x^*_e = 0$ for all the other edges. Note that the costs c which gave α_n are shown on the edges, and for edges not shown the cost c_{ij} was the cost of a minimum cost i to j path using the costs shown.

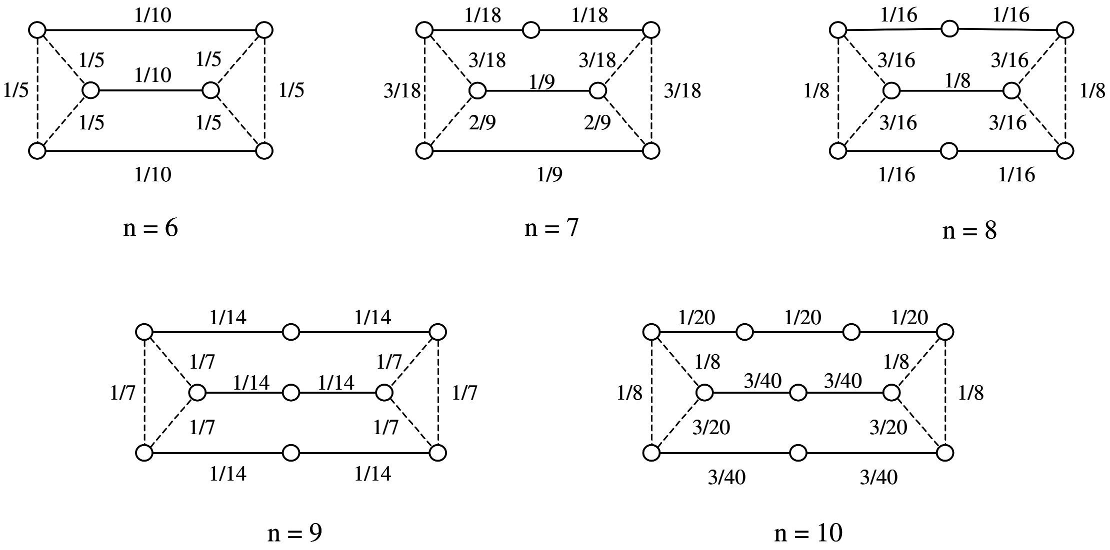


Figure 2: Vertices and costs which give α_n.

By extrapolating the pattern of the vertices shown in Figure 2, we developed the following conjecture for the value of α_n.

Conjecture 4.1 *For all integers $n \geq 3$, $\alpha_n = \beta_n$, where β_n is defined as follows:*

$$\beta_n = \begin{cases} \dfrac{4n+6}{3n+9} & \textit{if } n = 0 \ (\textit{mod } 3), \\ \dfrac{4\lfloor \frac{n}{3} \rfloor^2 + 2\lfloor \frac{n}{3} \rfloor - 2}{3\lfloor \frac{n}{3} \rfloor^2 + 3\lfloor \frac{n}{3} \rfloor - 2} & \textit{if } n = 1 \ (\textit{mod } 3), \\ \dfrac{4\lfloor \frac{n}{3} \rfloor^2 + 2\lfloor \frac{n}{3} \rfloor - 4}{3\lfloor \frac{n}{3} \rfloor^2 + 3\lfloor \frac{n}{3} \rfloor - 4} & \textit{if } n = 2 \ (\textit{mod } 3). \end{cases} \tag{24}$$

Note that in the formulas in this new conjecture, β_n approaches 4/3 as $n \to \infty$. Thus the correctness of Conjecture 4.1 would imply that Conjecture 1.1 is true. Also note that for each n we have found a set of costs for which the value in Conjecture 4.1 is tight, i.e. for which $TOUR/SUBT = \beta_n$, showing that $\alpha_n \geq \beta_n$. We describe this result next.

Consider the family $\mathfrak{F}$ of vertices x^* of S^n, $n = a+b+c+3$, $a, b, c \geq 1$, with weighted support graphs G_{x^*} as shown in Figure 3. In the figure, the solid edges represent edges for which $x^*_e = 1$, the dashed edges represent edges for which $x^*_e = 1/2$, in the 1-path between nodes 1 and 4 there are a edges, in the 1-path between nodes 2 and 5 there are b edges, and in the 1-path between nodes 3 and 6 there are c edges.

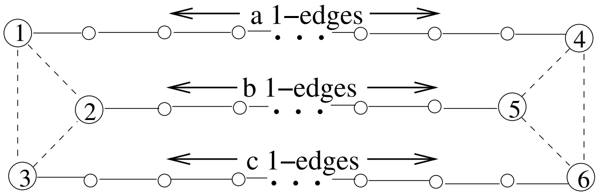


Figure 3: Weighted support graphs of the family $\mathfrak{F}$ of vertices x^* of S^n.

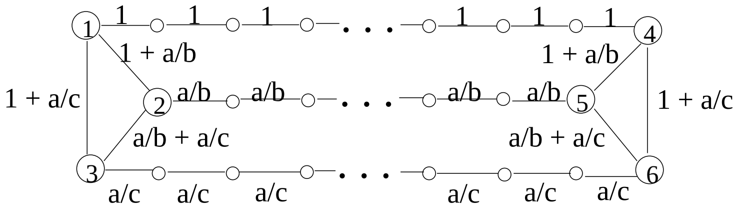


Figure 4: Cost vector w for family $\mathfrak{F}$ of vertices of S^n.

Without loss of generality, assume that $a \geq b \geq c$, and consider the metric cost vector w for K_n, $n \geq 6$, shown in Figure 4, where the cost w_{uv} for any edge $uv \in E$ not shown is the cost of a cheapest u to v path in the figure. Then we have the following results.

Lemma 4.1 *For the cost function w, vertex x^* of $\mathfrak{F}$ is an optimal solution for the SEP.*

Proof. To prove the result, it is sufficient to find a feasible solution (y^*, u^*, d^*) for the dual LP (15) which satisfies the complementary slackness conditions (19), (20) and (21) for x^* with respect to the cost function w. It is easy to verify that the following values provide such a solution: In the 1-path joining nodes 1 and 4 let $y^*_v = 1$ for every node v and let $u^*_{vz} = 1$ for every edge vz, in the 1-path joining nodes 2 and 5 let $y^*_v = a/b$ for every node v and let $u^*_{vz} = a/b$ for every edge vz, in the 1-path joining nodes 3 and 6 let $y^*_v = a/c$ for every node v and let $u^*_{vz} = a/c$ for every edge vz, let $u^*_e = 0$ for all other edges $e \in E$, and let $d^*_S = 0$ for all $S \in \mathfrak{S}$. □

Now consider the tour T^* for K_n, $n = a+b+c+3$, shown in Figure 5, which consists of the three 1-paths in Figure 3 plus the edges 13, 24, and 56.

Lemma 4.2 *The tour T^* of value $4a + 2(1 + a/b + a/c)$ is an optimal solution for the STSP for the cost function w of K_n, $n = a+b+c+3$.*

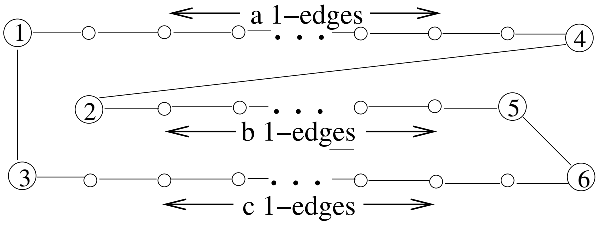


Figure 5: Optimal tour T^* for cost function w.

Proof. Let G^* be the weighted subgraph of K_n whose edges and weights are shown in Figure 4. Since the cost w_{uv} of every edge $uv \in K_n$ not found in G^* is equal to the cost of a cheapest u to v path, for any tour T' of K_n we can obtain a corresponding Eulerian tour T of the same cost as T' whose edge set consists only of multiple copies of edges of G^*. This Eulerian tour T is obtained by replacing any edge $e = uv$ in T' which is not in $E(G^*)$ by the cheapest u to v path in G^*. Thus in order to show that the tour T^* is an optimal tour for cost function w, it suffices to show that every Eulerian tour T consisting of edges in G^* has cost at least $4a + 2(1 + a/b + a/c)$, which is the cost of T^*.

For any Eulerian tour T let $x^T \in \mathbf{R}^{E(G^*)}$ be the vector representing T, i.e. for each edge $e \in E(G^*)$ let x^T_e be the number of copies of e in T. Let A', B' and C' represent the set of nodes in the paths joining nodes 1 and 4, nodes 2 and 5 and nodes 3 and 6, respectively, and let A, B and C be the set of edges in these three paths, respectively. Finally, let H represent the set of six edges in G^* not in the three paths, i.e. let $H = \{12, 23, 31, 45, 56, 61\}$.

First notice that the cost of each of the 3 paths is a, i.e.

$$\sum_{e\in A} w_e = \sum_{e\in B} w_e = \sum_{e\in C} w_e = a. \tag{25}$$

Second notice that we have the following result for the cost of the Eulerian tour T for the edges in H:

$$\begin{aligned}\sum_{e\in H} w_e &= 1(x^T_{12} + x^T_{13} + x^T_{45} + x^T_{46}) + \frac{a}{b}(x^T_{12} + x^T_{23} + x^T_{45} + x^T_{56}) + \frac{a}{c}(x^T_{23} + x^T_{13} + x^T_{46} + x^T_{56})\\ &= 1(x^T(\delta(A')) + \frac{a}{b}(x^T(\delta(B')) + \frac{a}{c}(x^T(\delta(C')).\end{aligned}$$

Thus for any Eulerian tour T, we have that its cost wx^T can be calculated as follows:

$$wx^T = \sum_{e\in A} w_e x^T_e + \sum_{e\in B} w_e x^T_e + \sum_{e\in C} w_e x^T_e + 1(x^T(\delta(A')) + \frac{a}{b}(x^T(\delta(B')) + \frac{a}{c}(x^T(\delta(C')). \tag{26}$$

Finally notice that since T forms an Eulerian graph, the number of edges in any cut in T is even and at least two. Thus if we consider the three sets of edges A, B and C from the three paths in G^*, at most one edge e from each set can have $x^T_e = 0$, and at most two of the three path edge sets can have such an edge with value zero in x^T. Thus we can partition our examination of the Eulerian tours into three cases.

Case 1. None of the three edges sets A, B or C have an edge e such that $x^T_e = 0$. For this case we have that $x^T_e \geq 1$ for every edge $e \in A \cup B \cup C$. Since T is an Eulerian tour, any cut which crosses the three paths of G^* must cut across an even number of edges of T, thus at least 4 edges, which implies $x^T_e \geq 2$ for every edge e in one of the edge sets A, B or C. Using (25), this implies that

$$\sum_{e\in A\cup B\cup C} w_e x^T_e \geq 4a.$$

Since every cut in T has size at least two, it then follows from (26) that

$$wx^T \geq 4a + 2(1 + \frac{a}{b} + \frac{a}{c}),$$

as required.

Case 2. Exactly one of the three edge sets A, B and C has an edge e such that $x^T_e = 0$. Here we consider the case where it is the set A which has an edge for which $x^T_e = 0$. The cases for the sets B and C follow using similar arguments.

We know $x^T_e \geq 1$ for all edges $e \in B \cup C$, and thus it follows from (25) that $\sum(w_e x^T_e : e \in B) \geq a$ and $\sum(w_e x^T_e : e \in C) \geq a$. Let uv be the edge in A which does not appear in T. Since T is an Eulerian tour, we must have $x^T_e \geq 2$ for all other edges $e \in A$. Thus again using (25) it follows that

$$\sum_{e \in A} w_e x^T_e \geq (2a - 2w_{uv}) = 2a - 2,$$

and thus

$$\sum_{e \in A \cup B \cup C} w_e x^T_e \geq (2a - 2) + a + a = 4a - 2.$$

Now consider the cuts $x^T(\delta(S))$ for S equal to each of the sets A', B' and C'. We know $x^T(\delta(B')) \geq 2$ and $x^T(\delta(C')) \geq 2$ since T is an Eulerian tour. Moreover, since $x^T_{uv} = 0$, we must also have that $x^T_{13} + x^T_{12} \geq 2$ and $x^T_{45} + x^T_{46} \geq 2$, and thus $x^T(\delta(A')) \geq 4$. Hence it follows from (26) that

$$wx^T \geq 4a - 2 + 4 + 2(\frac{a}{b} + \frac{a}{c}) = 4a + 2(1 + \frac{a}{b} + \frac{a}{c}),$$

as required.

Case 3. Exactly two of the three edge sets A, B and C have an edge e such that $x^T_e = 0$. Here we consider the case where sets A and B each have an edge not in T. The other two cases follow using similar arguments.

As argued in Case 2, we have $\sum(w_e x^T_e : e \in A) \geq 2a - 2$ and $\sum(w_e x^T_e : e \in B) \geq 2a - 2(a/b)$. Since we also have $x^T_e \geq 1$ for all edges $e \in C$, it follows that

$$\sum_{e \in A \cup B \cup C} w_e x^T_e \geq 2a - 2 + 2a - 2\frac{a}{b} + a = 5a - 2 - 2\frac{a}{b}.$$

Moreover, we have $x^T(\delta(A')) \geq 4$, $x^T(\delta(B')) \geq 4$ and $x^T(\delta(C')) \geq 2$, and thus by (26) it follows that

$$wx^T \geq 5a - 2 - 2\frac{a}{b} + 4 + 4\frac{a}{b} + 2\frac{a}{c} \geq 4a + 2(1 + \frac{a}{b} + \frac{a}{c}),$$

as required. □

Theorem 4.1 *For the cost function w for K_n, $n \geq 6$, there are values of a, b and c such $TOUR/SUBT = \beta_n$, where β_n is as described in Conjecture 4.1.*

Proof. From Lemmas 4.1 and 4.2, we know that for cost function w and general values of a, b and c that $TOUR = 4a + 2(1 + a/b + b/c)$ and $SUBT = 3a + 2(1 + a/b + b/c)$. We consider three cases.

Case 1. $n = 0 \pmod 3$. Letting $a = b = c = (n/3) - 1$, we have

$$\frac{TOUR}{SUBT} = \frac{4(\frac{n}{3} - 1) + 2(1 + 1 + 1)}{3(\frac{n}{3} - 1) + 2(1 + 1 + 1)} = \frac{4\frac{n}{3} + 2}{3\frac{n}{3} + 3} = \frac{4n + 6}{3n + 9},$$

as required.

Case 2. $n = 1 \pmod 3$. Letting $a = \lfloor n/3 \rfloor$, $b = c = \lfloor n/3 \rfloor - 1$, we have

$$\frac{TOUR}{SUBT} = \frac{4a + 2(1 + \frac{a}{a-1} + \frac{a}{a-1})}{3a + 2(1 + \frac{a}{a-1} + \frac{a}{a-1})} = \frac{4a(a-1) + 2(a-1) + 4a}{3a(a-1) + 2(a-1) + 4a} = \frac{4\lfloor \frac{n}{3} \rfloor^2 + 2\lfloor \frac{n}{3} \rfloor - 2}{3\lfloor \frac{n}{3} \rfloor^2 + 3\lfloor \frac{n}{3} \rfloor - 2},$$

as required.

Case 3. $n = 3 \pmod 3$. Letting $a = b = \lfloor n/3 \rfloor$, $c = \lfloor n/3 \rfloor - 1$, we have

$$\frac{TOUR}{SUBT} = \frac{4a + 2(1 + 1 + \frac{a}{a-1})}{3a + 2(1 + 1 + \frac{a}{a-1})} = \frac{4a(a-1) + 4(a-1) + 2a}{3a(a-1) + 4(a-1) + 2a} = \frac{4\lfloor \frac{n}{3} \rfloor^2 + 2\lfloor \frac{n}{3} \rfloor - 4}{3\lfloor \frac{n}{3} \rfloor^2 + 3\lfloor \frac{n}{3} \rfloor - 4},$$

as required. □

Since we have $\beta_n = \alpha_n = 1$ for $n = 3, 4, 5$, the following is a direct consequence of Theorem 4.1.

Corollary 4.1 *For $n \geq 3$, $\alpha_n \geq \beta_n$.*

5. Beyond n=10.

The methods described thus far were sufficient for finding the exact value of α_n for $n \leq 10$, however for $n = 11$ we found that PORTA was unable to generate all of the vertices for S^n using this method even after running for weeks. We therefore decided to restrict ourselves to the *half-integer vertices* of S^n, i.e. the vertices x of S^n for which the values of x_e are all 0, 1 or 1/2.

To find all the half-integer vertices for S^n, $11 \leq n \leq 14$ we used the following method, which did not require the use of PORTA. Since the maximum degree in the support graph G_x of any half-integer vertex x of S^n is 4 and since the half-integer vertices of S^n which have a node of degree 2 can be generated from the list of half-integer vertices of S^{n-1} by Theorem 3.2, we generated all the non-isomporphic connected graphs on n nodes containing at most $2n - 3$ edges which have degree 3 or 4 at every node using the nauty software package (by Theorem 3.1 this gives all the possible non-isomorphic support graphs for half-integer vertices of S^n with no nodes of degree 2). In fact, using Theorem 3.4 we were able to further restict our graph generation specifications for nauty to include the restriction that there had to be at least 6 nodes of degree 3.

Next for each of these graphs G' generated by nauty we fixed every edge in G' to either value 1/2 or 1 in such a way as to generate every possible half-integer vertex with support graph G' satisfying the degree constraints $x(\delta(v)) = 2$ for every node v (graphs for which this was impossible were removed). We then went through the list of weighted graphs generated and removed the ones which did not satisfy all of the subtour elimination constraints (3). We did so by removing all the half-integer graphs which had a minimum cut of value strictly less than 2. After doing this, we were guaranteed that every vector x corresponding to the graphs in our current list belonged to the subtour polytope S^n. It therefore only remained to verify that such vectors were in fact vertices of S^n. Since a vertex is the unique solution of its tight constraints, it follows that x' is a vertex of S^n if and only if the rank of the following constraint matrix is $|E| = n(n-1)/2$: $x(\delta(v)) = 2$ for all v, $x(\delta(S)) = 2$ for every set S such that $x'(\delta(S)) = 2$, and $x_e = 0$ for every edge e such that $x'_e = 0$. Thus finding the rank associated with every vector x corresponding to graph G' in our list allowed us to find the complete list of half-integer vertices of S^n for $11 \leq n \leq 14$. Isomorphisms were removed using the nauty package.

Note that the LP's (22) that needed to be solved for $n \geq 11$ had over one million tour constraints (9), and were too large to be solved directly using CPLEX. This problem was easily overcome by initially only including the tour constraints for all tours that had at most one edge not in the support graph of the vertex, then repeatedly checking to see if the resulting LP solution satisfied all tour constraints (9), and if not, adding the violated ones to the LP, until all tour constraints were satisfied by the LP solution.

Let α_n^* represent the value of α_n when further restricted to cost functions c which are optimized at half-integer vertices of S^n. Using our method on the half-integer vertices of S^n, we were able to find the exact value of α_n^* for all $11 \leq n \leq 14$. The results are shown in Table 2. In column 2 of the table we list the number of non-isomorphic half-integer vertices of S^n we found (this number includes the integer tour vertex). Note that this also represents the number of LPs (22) that we needed to solve for each value of n. Column 3 contains the value of α_n^*.

Table 2: Integrality gap results for half-integer vertices.

n	**Number of half-integer vertices**	α_n^*
11	1,022	19/16
12	5,637	6/5
13	31,686	35/29
14	185,625	17/14

Note that for all values of n considered here we have $\alpha_n^* = \beta_n$ from Conjecture 4.1, and there was again a unique vertex which gave the optimal which follows the same pattern discussed in Section 4. This lends more support to Conjecture 4.1, and also indicates that if that conjecture is not true for these values of n, then it must be that the exact integrality gap α_n comes from a cost function which is not optimized at a half-integer vertex of S^n.

Acknowledgments. We would like to thank Lucia Moura and Marc Raaphorst for their helpful suggestions on creating non-isomorphic subgraphs with certain degree properties, and Kevin Cheung for some useful discussions. Note that this research was partially supported by grants from the Natural Sciences and Engineering Research Council of Canada, and we are thankful for that support.

References

[1] S. Boyd, R. Carr, *A new bound for the ratio between the 2-matching problem and its linear programming relaxation*, Math. Programming Series A **86** (1999), 499-514.

[2] S. Boyd, R. Carr, *Finding low cost TSP and 2-matching solutions using certain half-integer subtour vertices*, Technical Report, University of Ottawa, Ottawa, 2000.

[3] S. Boyd, P. J. Elliott-Magwood, *Computing the integrality gap of the asymmetric travelling salesman problem*, Electronic Notes in Discrete Mathematics Vol. **19** (2005), 241-247.

[4] S. Boyd, W. R. Pulleyblank, *Optimizing over the subtour polytope of the travelling salesman problem*, Math. Programming **49** (1991), 163-187.

[5] S. Boyd, S. Xiong, *Generating all non-isomporphic vertices for the subtour elimination polytope*, in preperation.

[6] R. Carr, R. Ravi, *A new bound for the 2-edge connected subgraph problem*, Proceedings of Integer Programming and Combinatorial Optimization (IPCO), Lecture Notes in Computer Science, Springer (1998), 112-125.

[7] R. Carr, S. Vempala, *Towards a 4/3 approximation for the Asymmetric Traveling Salesman Problem*, Proceedings of the eleventh annual ACM-SIAM symposium on discrete algorithms (SODA) (2000), 116-125.

[8] M. Charikar, M. X. Goemans, H. Karloff, *On the integrality ratio for asymmetric TSP*, Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, Rome, Italy (2004), 101–107.

[9] T. Christof, A. Löbel, M. Stoer, *PORTA, A POlyhedron Representation Transformation Algorithm*, `http://www.zib.de/Optimization/Software/Porta/index.html`, 1997.

[10] N. Christofides, *Worst case analysis of a new heuristic for the traveling salesman problem*, Report 388, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, 1976.

[11] M. Grötschel, L. Lovasz, A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.

[12] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy-Kan, D. B. Shmoys (eds.), *The Travelling Salesman Problem: A Guided Tour of Combinatorial Optimization.* Wiley, New York, 1985.

[13] B. McKay, *Nauty User's Guide (Version 1.5)*, Technical Report TR-CS-90-02, Department of Computer Science, Australia National University, 1990.

[14] D. B. Shmoys,D. P. Williamson, *Analyzing the Held-Karp TSP bound: A monotonicity property with application*, Inf. Process. Lett. **35** (1990), 281-285.

[15] L. A. Wolsey, *Heuristic analysis, linear programming and branch and bound*, Math. Programming Study **13** (1980), 121-134.