On the Held-Karp Relaxation for the Asymmetric and Symmetric Traveling Salesman Problems

Robert Carr* Santosh Vempala†

April 14, 2003

Abstract

A long-standing conjecture in combinatorial optimization says that the integrality gap of the famous Held-Karp relaxation of the metric STSP (Symmetric Traveling Salesman Problem) is precisely 4/3. In this paper, we show that a slight strengthening of this conjecture implies a tight 4/3 integrality gap for a linear programming relaxation of the metric ATSP (Asymmetric Traveling Salesman Problem). This is surprising since no constant-factor approximation is known for the latter problem. Our main tools are a new characterization of the integrality gap for linear objective functions over polyhedra, and the isolation of "hard-to-round" solutions of the relaxations.

1 Introduction

The Traveling Salesman Problem has guided and challenged the field of combinatorial optimization for the past several decades. Let $G = (V, E)$ be a directed (undirected) graph. A *Hamilton cycle* (or *tour*) in G is a directed (undirected) simple cycle that spans V. The *Asymmetric Traveling Salesman Problem* or ATSP (*Symmetric Traveling Salesman Problem* or STSP) is given G and a cost function on the edges, find a minimum cost directed (undirected)

*Sandia National Labs, Albuquerque, NM 87185. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. `bobcarr@cs.sandia.gov`.

†Department of Mathematics, M.I.T. Supported by a Miller Fellowship at U.C. Berkeley and an NSF CAREER Award. `vempala@math.mit.edu`

Hamilton cycle. It is well-known that there can be no efficient approximation algorithm unless $P = NP$ [12]. However, for the *metric* ATSP and *metric STSP*, where we restrict the distances to satisfy the triangle inequality, there are algorithms that achieve a $\log n$ approximation [4] in the asymmetric case, and a 3/2 approximation in the symmetric case [3]. Improving on the approximation ratios for these problems has been an outstanding open problem for nearly two decades. The best known lower bound on the approximability assuming a metric cost function (for both problems) is $(1 + \epsilon)$ for some fixed ϵ much smaller than 1 (i.e., it is NP-hard to find a $(1 + \epsilon)$ approximation) [11].

A promising direction for an improved approximation for the STSP is a linear programming relaxation of the problem, proposed by Held and Karp [8]. The *integrality gap* of the relaxation, i.e., the worst possible ratio of the true optimum to the optimum over the relaxation, is conjectured to be 4/3 for the metric STSP (see Section 1.1). A proof of this conjecture would lead to a 4/3 approximation for the metric STSP. (An *algorithmic* proof that yields a TSP tour within 4/3 of this relaxation would lead to a 4/3 approximation algorithm.)

In this paper, we study a linear programming relaxation of the ATSP. Our main result is a reduction from the ATSP to the STSP which implies that a slight strengthening (described below) of the 4/3 conjecture for the metric STSP *yields a 4/3 approximation for the metric ATSP also!*

Our proof relies on two main tools. The first is a simple characterization of the integrality gap that perhaps was known, but never explicitly stated (Goemans, [6], exploits the main idea of this characterization). This is described in Section 2. Applying it to the Held-Karp metric STSP relaxation implies that the integrality gap of the relaxation is 4/3 if and only if for any extreme point x^* of the relaxation of a given instance G, $\frac{4}{3}x^*$ can be expressed as a convex combination of *Eulerian subgraphs* of G. Throughout this paper, we assume that in addition to the even degree requirements, Eulerian subgraphs are also spanning, connected, and that multiple copies of an edge are permitted.

The strengthening of the STSP conjecture needs a simple definition. A graph is called *leafless* if every vertex of degree 2 of the graph has at least two distinct neighbors. Then the stronger conjecture is that for any extreme point x^* of the Held-Karp STSP relaxation, $\frac{4}{3}x^*$ can be expressed as a convex combination of *leafless* Eulerian subgraphs (and thus using Eulerian subgraphs with leaves is no longer allowed here).

Our second important tool is the isolation of a subset of extreme points of the relaxation with the following property: if the true optimum is always within 4/3 of the cost (objective function value) of extreme points in the subset, then it is within 4/3 of *any* extreme point of the relaxation (and hence the integrality gap is 4/3). There is a striking analogy here with the theory of NP-completeness (see Section 3). We refer to extreme points in such a "complete" subset as *fundamental* extreme points. This technique has been used in the context of the STSP in earlier work [2]; here we develop it further and apply it to the ATSP.

To prove the main result, we identify the fundamental extreme points of the ATSP relaxation (Section 3.3), describe a reduction to STSP instances (Section 4), and show that if the stronger conjecture holds, then the integrality gap of the metric ATSP relaxation is 4/3, i.e., the metric ATSP is approximable to within 4/3 in polynomial time.

Both characterizations developed here, for the integrality gap and for the "hardest-to-round" extreme points, are quite general in nature and seem well-suited to the study of approximation algorithms.

1.1 Relaxations of the TSP

Given a complete edge-weighted graph $G = (V, E)$, the symmetric TSP can be formulated as the following integer program (IP) with a variable x_{ij} for each edge $ij \in E$.

$$\min \sum_{ij \in E} c_{ij} x_{ij}$$

$$\sum_{j \in V \setminus \{i\}} x_{ij} = 2 \quad \forall i \in V \tag{1}$$

$$\sum_{ij: i \in S, j \notin S} x_{ij} \geq 2 \quad \forall S \subset V, S \neq \emptyset \tag{2}$$

$$x_{ij} \in \{0, 1\} \quad \forall ij \in E \tag{3}$$

The Held-Karp relaxation is the linear programming relaxation of this integer program obtained by modifying constraints (3) to $x_{ij} \geq 0$ for each $ij \in E$. Let $Q^=$ be the polyhedron for the Held-Karp relaxation, where the "=" superscript indicates that constraints (1) are equations.

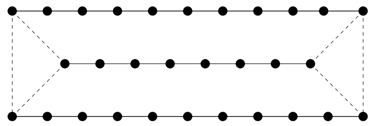

Figure 1: An example showing a 4/3 integrality gap.

Let Q be the polyhedron where these degree constraints (1) are relaxed to greater than or equal to constraints. The two relaxations $Q^=$ and Q are closely related, as is expressed in the lemma below which was proved first by Cunningham [10] and then by Bertsimas and Goemans [5]. This lemma is a consequence of Lovász's edge-splitting Lemma [9].

Lemma 1 *Assume that the costs in a cost vector $c \geq 0$ satisfy the triangle inequality. Then,*

$$\min_{x \in Q^=} c \cdot x = \min_{x \in Q} c \cdot x.$$

For linear functions $c \cdot x$, with $c > 0$ and $x \geq 0$, the *integrality gap* is a measure of the quality of the relaxation P of an integer polyhedron Z in the positive orthant, and can be defined as

$$\max_{c>0} \left(\frac{\min_{x \in Z} c \cdot x}{\min_{x \in P} c \cdot x} \right).$$

One can also consider integrality gaps over more restrictive sets of cost functions, such as those satisfying the triangle inequality. We will in fact impose this metric restriction on our TSP cost functions from now on, since the integrality gap for the TSP would otherwise be infinite. What is the integrality gap of the Held-Karp relaxation? A simple example, illustrated in Figure 1 shows that it is at least 4/3. In this figure, the solid edges have value 1 in the Held-Karp relaxation and the dashed edges have value 1/2. The cost of the solid edges is 1, the cost of the dashed edges is 2, and the costs of the other edges are determined by the triangle inequality. The conjecture that this is *tight*, i.e. the gap is also at most 4/3, has been a subject of intensive study for the past decade.

Conjecture 1 *The integrality gap of the Held-Karp STSP relaxation is 4/3.*

Here we study a relaxation of the ATSP on the complete digraph $G = (V, E)$, which can be viewed as the asymmetric generalization of the Held-Karp relaxation. It should be noted that this relaxation as well as the symmetric one, can both be solved in polynomial-time [7].

$$\min \sum_{ij \in E} c_{ij} x_{ij}$$

$$\sum_{j \in V \setminus \{i\}} x_{ij} = 1 \quad \forall i \in V \tag{4}$$

$$\sum_{i \in V \setminus \{j\}} x_{ij} = 1 \quad \forall j \in V \tag{5}$$

$$\sum_{ij : i \in S, j \notin S} x_{ij} \geq 1 \quad \forall S \subset V, S \neq \emptyset \tag{6}$$

$$x_{ij} \geq 0 \quad \forall ij \in E \tag{7}$$

From the algorithm of Frieze et al. [4], it follows that the integrality gap of this relaxation is at most $\log n$. The largest known lower bound on the integrality gap, however, is just 4/3 (obtained by bi-directing every edge in the undirected 4/3 example). Nevertheless, the following conjecture seems extremely rash.

Conjecture 2 *The integrality gap of the ATSP relaxation is 4/3.*

In what follows, we show that Conjecture 2 is only slightly stronger than Conjecture 1.

2 A characterization of the integrality gap

In this section we develop a simple characterization of the integrality gap in terms of an associated polyhedron which we call the *dominant*.

Let P denote the polyhedron in $\mathbf{R}^n$ defined by $Ax \geq b, \quad x \geq 0$.

Definition 1 *The dominant $\mathcal{D}(P)$ of a polyhedron P is the set of points $y \in \mathbf{R}^n$ which* dominate *some point $x \in P$. That is,*

$$\mathcal{D}(P) = \{y \in \mathbf{R}^n \mid \exists x \in P : \forall i \in \{1, \ldots, n\} \; y_i \geq x_i\}.$$

The following facts are elementary but insightful.

Lemma 2 *The dominant of a polyhedron is a polyhedron.*

Lemma 3 *For any linear function $c \cdot x$ with $c \geq 0$,*

$$\min_{x \in P} c \cdot x = \min_{x \in \mathcal{D}(P)} c \cdot x.$$

We are now ready for the main theorem of this section.

Theorem 1 *The integrality gap of a relaxation P of an integral polyhedron Z in the positive orthant is the smallest real number r such that for any extreme point x^* of P, $rx^* \in \mathcal{D}(Z)$.*

Proof: First we note that it suffices to show that the integrality gap of P with respect to Z is *at most* r iff $rx^* \in \mathcal{D}(Z)$ for every extreme point x^* of P. Suppose that for any extreme point x^* of P, $rx^* \in \mathcal{D}(Z)$. Let a cost vector $c > 0$ be given. Choose an extreme point $x^* \in P$ having minimum cost. Since $rx^* \in \mathcal{D}(Z)$, we can express rx^* as a convex combination $rx^* = \sum_i \lambda_i x_i$, where each x_i dominates an extreme point of Z. Denote the extreme point in Z that is dominated by the minimum cost such x_i by $y \in Z$. By an averaging argument, we have that $c \cdot (rx^*) \geq c \cdot y$, i.e. $rc \cdot x^* \geq c \cdot y$. Hence, the integrality gap for P is at most r.

Conversely, suppose that the integrality gap for P is at most r. Assume for the purpose of a contradiction that there is an extreme point x^* of P such that rx^* *cannot* be expressed as a convex combination of points from $\mathcal{D}(Z)$. Since $\mathcal{D}(Z)$ is a polyhedron, there must be a hyperplane $w \cdot x = t$ which separates rx^* from $\mathcal{D}(Z)$. That is, $w \cdot (rx^*) < t \leq w \cdot y$ for every $y \in \mathcal{D}(Z)$. For such a vector w, it must be the case that every component is non-negative. Otherwise, if some component w_e is negative, then by using a y with y_e equal to a large enough value, we would violate the above inequality. Hence there exists a vector $w > 0$ such that

$$w \cdot (rx^*) < \min_{y \in \mathcal{D}(Z)} w \cdot y.$$

That is,

$$r < \frac{\min_{y \in \mathcal{D}(Z)} w \cdot y}{w \cdot x^*} \leq \frac{\min_{y \in \mathcal{D}(Z)} w \cdot y}{\min_{x \in P} w \cdot x} = \frac{\min_{y \in Z} w \cdot y}{\min_{x \in P} w \cdot x}.$$

But this means the integrality gap of P with respect to Z is greater than r, a contradiction. □

For a polyhedron P, let $\mathcal{Z}(P)$ denote the convex hull of the integer points within P. Thus P is a relaxation of $\mathcal{Z}(P)$. In particular, $\mathcal{Z}(Q^=)$ is the convex hull of all (undirected) Hamilton cycles and $Q^=$ its relaxation (since adding integrality constraints to $Q^=$ yields an IP formulation for STSP). If we let Z_E denote the convex hull of all (undirected) Eulerian graphs, then the following lemma establishes a useful relationship.

Lemma 4 *Let the costs c satisfy the triangle inequality. Then, we have*

$$\min_{x \in Z_E} c \cdot x = \min_{x \in \mathcal{Z}(Q^=)} c \cdot x.$$

Proof: Since $\mathcal{Z}(Q^=) \subset \mathcal{Z}_{\mathcal{E}}$, the minimum over $\mathcal{Z}_{\mathcal{E}}$ is at most that over $\mathcal{Z}(Q^=)$. Let a minimizing Eulerian graph T be given. Take an Euler tour of T and shortcut it by bypassing any node v after leaving it for a second time as follows. Add the edge which completes a triangle with the last two edges (an edge into v and an edge out of v), and remove these last two edges. This results in a Hamilton cycle that has equal or smaller cost since c is metric. □

The next lemma will be useful in applying the characterization to the STSP.

Lemma 5

$$\mathcal{D}(Z_E) = Z_E.$$

Proof: It is clear the $Z_E \subseteq \mathcal{D}(Z_E)$. For the opposite inequality, consider some $y \in \mathcal{D}(Z_E)$, and let $x \in Z_E$ be a point such that $y \geq x$ component-wise, and further x differs from y in the fewest possible components (edges). Let e be a component such that $y_e > x_e$. Consider the point x' obtained by setting $x'_e = x_e + t$, where t is an even integer greater than $y_e - x_e$, and $x'_f = x_f$ for all other components f. Then $x' \in Z_E$. Consider the convex combination $x'' = (1-\alpha)x + \alpha x'$ where $\alpha = \frac{y_e - x_e}{t}$. Then $x'' \in Z_E$ also. Further, x'' is dominated by y and has one less component different from y than does x. This contradicts our assumption that x differs from y in the fewest possible components. □

Lemma 6 *The integrality gap between Z_E and its relaxation Q for $c > 0$ equals the integrality gap between $\mathcal{Z}(Q^=)$ and $Q^=$ for metric c.*

Proof: From Lemmas 4 and 1, for any cost vector $c > 0$ satisfying the triangle inequality, we have that (i) the optimal value over Z_E is equal to the optimal value over $\mathcal{Z}(Q^=)$, and (ii) the optimal value over the Held-Karp relaxation $Q^=$ is equal to the optimal value for c over the relaxation Q (note: Q is the relaxation obtained by dropping the degree constraints from $Q^=$). Hence, when c is metric in both cases, these gaps are the same.

We will now show that the gap of Q with respect to Z_E over cost vectors satisfying the triangle inequality is the same as the integrality gap of Q with respect to Z_E (i.e., over all cost vectors $c > 0$). Let $c > 0$. Define d_{ij} to be the length of the shortest path from i to j, where the edge lengths are given by c. Let $\min_{x \in Z_E} d \cdot x$ occur at $x^* \in Z_E$. For each edge $ij \in E$ such that $c_{ij} > d_{ij}$ and $x^*_{ij} \geq 1$, replace ij in x^* by a path of c-length equal to d_{ij}. The resulting solution x^{**} then satisfies $d \cdot x^* = c \cdot x^{**}$, and so $\min_{x \in Z_E} d \cdot x = \min_{x \in Z_E} c \cdot x$. We similarly have $\min_{x \in Q} d \cdot x = \min_{x \in Q} c \cdot x$. Hence, the integrality gap is unchanged. $\square$

Theorem 1 can be applied to the STSP to yield the following characterization.

Theorem 2 *The 4/3 Conjecture for the Held-Karp relaxation is equivalent to the following statement: For any extreme point x^* of the relaxation $Q^=$, $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs.*

Proof: We have the following:

(i) The 4/3 Conjecture is equivalent to this Held-Karp integrality gap being at most 4/3, since it is already known to be at least 4/3.

(ii) The Held-Karp integrality being at most 4/3 is equivalent to the gap of Q with respect to Z_E being at most 4/3, by Lemma 6.

(iii) The gap of Q with respect to Z_E being at most 4/3 is equivalent to saying that for any extreme point x^* of Q, $\frac{4}{3}x^*$ can be expressed as a convex combination of points from $\mathcal{D}(Z_E)$, by Theorem 1.

(iv) But $\mathcal{D}(Z_E) = Z_E$, by Lemma 5.

Hence the assertion equivalent to the 4/3 Conjecture is that $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs for all $x^* \in Q$. Certainly then, the 4/3 Conjecture implies that $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs for all $x^* \in Q^=$. Now suppose that $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs for all $x^* \in Q^=$. We could proceed to show that any $\overline{x} \in Q$ can also be so decomposed into Eulerian graphs, from which the 4/3 Conjecture follows. But, an easier approach

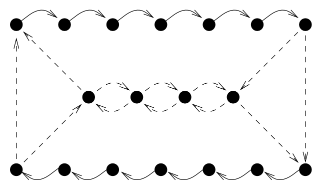

Figure 2: An ATSP extreme point.

is to take the minimum cost Eulerian graph in the decomposition of x^* and shortcut it as in Lemma 4. □

Theorem 1 can also be applied to the ATSP relaxation, producing a result that is completely analogous with our STSP result, but a new idea is needed since Lemma 5 does not hold in the directed case.

Theorem 3 *The integrality gap of the ATSP relaxation is at most r iff for any extreme point x^* of the relaxation, rx^* can be expressed* **exactly** *as a convex combination of directed Eulerian graphs.*

Proof: If for any extreme point x^* of the relaxation, rx^* is a convex combination of directed Eulerian graphs, then clearly our integrality gap is at most r.

Now suppose the integrality gap of the ATSP relaxation is at most r. Let an extreme point x^* of that relaxation be given. Choose a convex combination of Eulerian graphs by Theorem 1 (and an analysis similar to that in our STSP result) that is dominated by rx^*. Removing these weighted Eulerian graphs from rx^* results in a weighted graph where the indegree equals the outdegree at each vertex. Such a graph can be decomposed into weighted cycles, which can be arbitrarily added to the Eulerian graphs of our convex combination. □

We now give an example of such a convex combination. Consider the ATSP extreme point x^* in Figure 2 analogous to the STSP point shown in Figure 1.

Weighting the three directed Eulerian graphs in Figure 3 by 1/3 each yields a decomposition of $\frac{4}{3}x^*$ into a convex combination of directed Eulerian graphs.

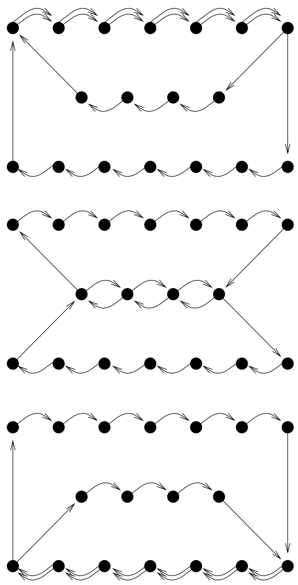

Figure 3: Tours in convex decomposition.

3 Fundamental Extreme Points

The notion of *fundamental extreme points* of the STSP relaxation was first developed in [2]. Intuitively, these are a subset of extreme points that are the most difficult to round to true (integer) solutions. This is shown by taking any extreme point and reducing it to a fundamental extreme point while preserving the integrality gap. In this sense the set of fundamental extreme points forms a *complete* subset for the set of extreme points.

3.1 Fundamental STSP Extreme Points

Suppose we have a subclass S of extreme points $\overline{x}$ such that being able to express $\frac{4}{3}\overline{x}$ as a convex combination of Eulerian graphs for all extreme points $\overline{x}$ belonging to this particular subclass is sufficient to prove Conjecture 1. Then we say that S is *sufficient* to prove Conjecture 1. In [2, 1], two different subclasses (with two variations of each, making four in all) of extreme points which are sufficient to prove Conjecture 1 were found. We call the extreme points in one particular such a subclass *fundamental (STSP) extreme points*, which we define as follows.

Consider any STSP extreme point x^*. Pick the smallest integer k such that x^*_e is a multiple of $\frac{1}{k}$ for every edge $e \in E$. Then form a $2k$-regular $2k$-edge connected multigraph $G_k = (V, E_k)$ as follows. For every edge $e = uv \in E$, put l edges between u and v, where $l := kx^*_e$. Then $x^* = \frac{1}{k}\chi^{E_k}$, and thus showing that $\frac{4}{3k}E_k$ can be expressed as a convex combination of Eulerian graphs is equivalent to showing that $\frac{4}{3}x^*$ can be so expressed. We will construct a larger $2k$-regular $2k$-edge connected multigraph $\overline{G}_k = (\overline{V}, \overline{E}_k)$ from G_k. We then use the above process in reverse on $\overline{G}_k$ to construct an extreme point $\overline{x}^*$. We will show that $\overline{x}^*$ has more special structure than x^* (structure that defines when an extreme point is a fundamental extreme point), but that for any $r \geq 1$, $r\overline{x}^*$ is as hard to express as a convex combination of Eulerian graphs as rx^* is.

We construct $\overline{G}_k = (\overline{V}, \overline{E}_k)$ and $\overline{x}^*$ as follows.

(i) Replace each $v \in V$ with a *circle* of nodes

$$v_1, v_2, \ldots, v_{2k},$$

and define $v_{2k+1} := v_1$. Now, $\overline{V} := \cup_{v \in V}\{v_i \mid i \in \{1, \ldots, 2k\}\}$.

(ii) For each $v \in V$, order the edges in $\delta(v) \subset E_k$ by

$$e^v_1, e^v_2, \ldots, e^v_{2k}.$$

(iii) For each edge $e \in E_k$ with endpoints u and v, place an edge $\overline{e} \in \overline{E}_k$ as follows.

(a) Find i and j such that $e = e^u_i$ and $e = e^v_j$.

(b) Place $\overline{e}$ so its endpoints are u_i and v_j.

(iv) Make an Euler tour on G_k, numbering the edges

$$e_1, e_2, \ldots e_{k|V|}, e_{k|V|+1} := e_1.$$

(v) Do the following for each $v \in V$.

(a) Find k disjoint *consecutive* index pairs $i, i+1$ with $e_i, e_{i+1} \in \delta(v)$.

(b) Let $F_v := \{(e_{i_r}, e_{i_r+1}) \mid r \in \{1, \dots, k\}\}$ be all these edge pairs, with the i_r's in increasing order. Have $i_{k+1} := i_1$, and note $\cup_{(e,f)\in F_v}\{e, f\} = \delta(v)$.

(c) For each $r \in \{1, \dots, k+1\}$, find v_{l_r} and v_{j_r} that are endpoints of $\overline{e}_{i_r}$ and $\overline{e}_{i_r+1}$ respectively. Do the following for $r = 1, \dots, k$.

- Place $k-1$ edges between v_{l_r} and v_{j_r}.
- Place k edges between v_{j_r} and $v_{l_{r+1}}$.

(d) Relabel the v indices so that the cycle C_v resulting from the edges placed in (c) visits v_1 through v_{2k} in order.

(vi) $\overline{x}^* := \frac{1}{k}\chi^{\overline{E}_k}$.

Theorem 4 *The support graph of $\overline{x}^*$ is 3-regular, with its fractional edges forming a Hamilton cycle. Also, for each $v \in V$, there is a cycle of alternating 1-edges and $\frac{k-1}{k}$-edges spanning the vertices of $C_v := \{v_1, v_2, \dots, v_{2k}\}$.*

Theorem 5 *$\overline{x}^*$ in Theorem 4 is a subtour extreme point.*

Proof: We first show that $\overline{x}^*$ is a feasible subtour point. If it were not, there would have to be a cut in the graph $\overline{G}_k = (\overline{V}, \overline{E}_k)$ of value less than $2k$. Clearly, such a cut $\mathcal{C}$ would have to go through some cycle C_v since G_k is $2k$-edge connected. But the contribution of the edges from the cycle C_v to any cut crossing it is at least $2k-2$. Clearly, the contribution from the non-circle edges in the cut $\mathcal{C}$ is at least 2 because of the fractional Hamilton cycle in $\overline{x}^*$. Hence, $\overline{x}^*$ is a feasible subtour point.

We show that $\overline{x}^*$ is an extreme point by showing that it cannot be expressed as $\frac{1}{2}x^1 + \frac{1}{2}x^2$, where x^1 and x^2 are distinct subtour points. Suppose $\overline{x}^*$ could be so expressed. Then the support graphs of x^1 and x^2 would coincide with or be subgraphs of the support graph $\overline{E}_k$ of $\overline{x}^*$. Because of the structure of the support graph, setting the value of just one fractional edge determines the entire solution due to the degree constraints. Hence, all the edges $e \in \overline{E}_k$ such that $x_e = \frac{1}{k}$ would have to say be smaller than $\frac{1}{k}$ in x^1 and larger than $\frac{1}{k}$ in x^2. But, then a cut separating any cycle C_v from the rest of the vertices in x^1 would have a value less than 2, which contradicts x^1 being a subtour point. $\square$

Next we establish the main property of $\overline{x}^*$.

Theorem 6 *If $\frac{4}{3}\overline{x}^*$ can be expressed as a convex combination of Eulerian graphs spanning $\overline{V}$, then $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs spanning V.*

Proof: Suppose $\frac{4}{3}\overline{x}^*$ can be expressed as a convex combination $\frac{4}{3}\overline{x}^* = \sum_i \lambda_i \chi^{\overline{H}_i}$, where the $\overline{H}_i$'s are Eulerian graphs spanning $\overline{V}$. For each i, contract each node set of C_v back to the vertex $v \in V$ in $\overline{H}_i$. Call the resulting graph H_i. Since contraction preserves the Eulerian property, H_i is an Eulerian graph spanning V. When one performs this contraction on $\overline{x}^*$, one gets x^*. As a result, we obtain that $\frac{4}{3}x^* = \sum_i \lambda_i \chi^{H_i}$. □

We can now define the subclass of *fundamental STSP extreme points* and state their main property.

Definition 2 *A fundamental STSP extreme point is an extreme point of the STSP relaxation satisfying the following conditions.*

(i) The support graph is 3-regular,

(ii) There is a 1-edge incident to each vertex,

(iii) The fractional edges form a Hamilton cycle.

Theorem 7 *The subclass of fundamental STSP extreme points is sufficient to prove Conjecture 1.*

Proof: If there is an extreme point x^* such that $\frac{4}{3}x^*$ cannot be expressed as a convex combination of Eulerian graphs, then by Theorem 6, the fundamental extreme point $\overline{x}^*$ is such that $\frac{4}{3}\overline{x}^*$ cannot be expressed as a convex combination of Eulerian graphs either. The theorem follows. □

3.2 Ultra-Fundamental STSP Extreme Points

To further restrict the class of extreme points that are sufficient to prove Conjecture 1 we replace each 1-edge uv in a fundamental extreme point x^* by the (first) construction shown in Figure 4, where the new vertices u', v', w, x, y, and z are created. In Figure 4, there is a long 1-path from u to u', a 1/2-cycle between u', w, and x, long 1-paths from w to y and from x to z, a 1/2-cycle between v', y, and z, and a long 1-path from v' to v. It is not

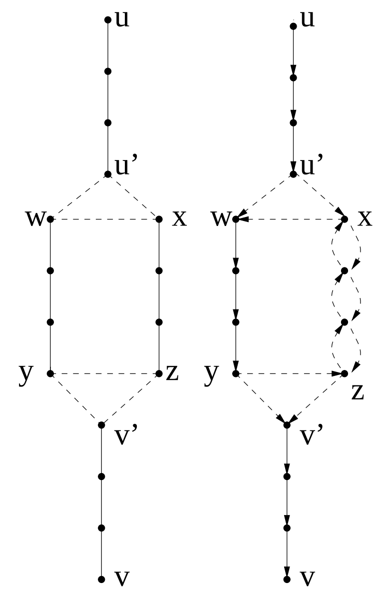

Figure 4: Gadgets for ultra-fundamental extreme points.

hard to see that the resulting solution continues to be an extreme point. We call these ultra-fundamental extreme points.

The idea behind these extreme points is that every Eulerian decomposition of $\frac{4}{3}\overline{x}^*$ for an ultra-fundamental extreme point $\overline{x}^*$ can be easily converted into an Eulerian decomposition of $\frac{4}{3}x^*$ of the corresponding fundamental extreme point x^* by contracting the nodes u', v' and the nodes in between them in each gadget and each Eulerian graph. On the other hand, one can imagine Eulerian decompositions of $\frac{4}{3}x^*$ that cannot be extended to an Eulerian decomposition of $\frac{4}{3}\overline{x}^*$. So, in this particular sense, the ultra-fundamental extreme points are harder than the fundamental extreme points to decompose into Eulerian graphs.

The following lemma, which involves the special case of *leafless* Eulerian graphs, will be very useful in the proof of the main reduction.

Lemma 7 *Let x^* be an ultra-fundamental STSP extreme point, and assume that $\frac{4}{3}x^*$ can be expressed as a convex combination of leafless Eulerian graphs,*

$$\frac{4}{3}x^* = \sum_i \lambda_i \chi^{K_i}.$$

Then each 1-path P of x^* must occur singly in Eulerian graphs that contribute 2/3 to the convex combination and must occur doubly in the remaining. That is,

$$\sum_{i:P \text{ occurs singly}} \lambda_i = \frac{2}{3}, \qquad \sum_{i:P \text{ occurs doubly}} \lambda_i = \frac{1}{3}.$$

Further, the path from u to u' in Figure 4 must occur with the same cardinality as the path from v' to v.

Proof: Since the graphs K_i are leafless, if part of a 1-path is present in K_i then the entire 1-path must be present. Since the 1-paths are all in cuts of exactly 3 edges (at the triangle vertices), in order to acheive a usage of $\frac{4}{3}$ times each 1-path in the convex combination, it is necessary to traverse each 1-path exactly once 2/3'rds of time in the convex combination and exactly twice the other 1/3'rd of the time. That is to say, when λ is the vector of convex multiples, we have

$$\sum_{i:P \text{ occurs once}} \lambda_i = 2/3,$$

where P is one of these newly created 1-paths. If P is from u to u' in Figure 4, it and its partner P' from v' to v form a cut of exactly 2 edges. Hence, these 1-paths are both traversed exactly once or both traversed exactly twice for every Eulerian graph in the convex combination. □

The proof of the next theorem is outlined by the explanation following "the idea behind these extreme points" at the beginning of this section.

Theorem 8 *The subclass of ultra-fundamental STSP extreme points is sufficient to prove Conjecture 1.*

Proof: Suppose Conjecture 1 is false. Then since the class of fundamental extreme points is sufficient to prove Conjecture 1, there must be a fundamental extreme point x^* such that $\frac{4}{3}x^*$ cannot be expressed as a convex combination of Eulerian graphs. Form an ultra-fundamental extreme point $\overline{x}^*$ from x^*. Then $\frac{4}{3}\overline{x}^*$ cannot be expressed as a convex combination of Eulerian graphs either, since otherwise a contraction argument similar to that in Theorem 6 would show that $\frac{4}{3}x^*$ could be so expressed as well. Therefore, the ultra-fundamental extreme points are also sufficient to prove Conjecture 1.
□

3.3 Fundamental ATSP Extreme Points

Here we study fundamental extreme points of the ATSP relaxation, which are analogous to those of the symmetric relaxation. The basic idea is to map any ATSP extreme point onto a fundamental ATSP extreme point by creating a directed multigraph. Consider any extreme point x^*. Pick the smallest integer k such that x^*_e is a multiple of $\frac{1}{k}$ for every arc e. Then form a k-regular k-edge connected directed multigraph $G_k = (V, E_k)$ as follows. For every arc $e = uv \in E$, put l arcs between u and v, where $l := kx^*_e$. Then $x^* = \frac{1}{k}\chi^{E_k}$, and thus showing that $\frac{4}{3k}E_k$ can be expressed as a convex combination of Eulerian graphs is equivalent to showing that $\frac{4}{3}x^*$ can be so expressed.

We construct $\overline{G}_k$ and $\overline{x}^*$ by amending our STSP construction (please refer back to) of these objects as follows.

(iii)(b) If $e = (u, v)$, then $\overline{e} := (u_i, v_j)$.

(iv) We make a *directed* Euler tour.

(v)(a) Choose $i, i+1$ so that $e_i \in \delta^-(v)$, $e_{i+1} \in \delta^+(v)$.

(v)(c) We place $k-1$ arcs (v_{j_r}, v_{l_r}) and k arcs $(v_{l_{r+1}}, v_{j_r})$.

(v)(d) When the v indices are relabeled, C_v is now a directed cycle.

Theorem 9 *The support graph of $\overline{x}^*$ is 3-regular, with its fractional edges forming a Hamilton cycle, but of alternating directions. Also, for each $v \in V$, there is a directed cycle of alternating 1-edges and $\frac{k-1}{k}$-edges spanning the vertices of $C_v := \{v_1, v_2, \ldots, v_{2k}\}$.*

The following Theorems are the directed analogues of Theorems 5 and 6 respectively. Their proofs are very similar.

Theorem 10 *$\overline{x}^*$ in Theorem 9 is an ATSP extreme point.*

Theorem 11 *If $\frac{4}{3}\overline{x}^*$ can be expressed as a convex combination of Eulerian graphs spanning $\overline{V}$, then $\frac{4}{3}x^*$ can be expressed as a convex combination of Eulerian graphs spanning V.*

We can now define the class of *fundamental ATSP extreme points* and state their main property.

Definition 3 *A fundamental ATSP extreme point is an extreme point for the subtour relaxation satisfying the following conditions.*

(i) The support graph is 3-regular,

(ii) There is a 1-arc incident to each vertex,

(iii) The fractional arcs, when considered as edges, form a Hamilton cycle,

(iv) Any two fractional arcs incident to any vertex v are both into v or both out of v.

The proof of the next theorem is analogous to that of Theorem 7.

Theorem 12 *The class of fundamental ATSP extreme points is sufficient to prove Conjecture 2.*

3.4 Ultra-Fundamental ATSP Extreme Points

Suppose now that we replace each 1-arc uv of a fundamental extreme point by the (second) construction in Figure 4. Then the resulting ultra-fundamental extreme points are harder to decompose into Eulerian graphs than fundamental extreme points for the same reasons as their symmetric counterparts, yielding the following theorem.

Theorem 13 *The ultra-fundamental ATSP extreme points are sufficient to prove Conjecture 2.*

4 The Main Reduction

We now will establish a strong connection between Conjecture 1 and Conjecture 2:

Theorem 14 *The integrality gap of the metric ATSP relaxation is at most $\frac{4}{3}$ iff for every ultra-fundamental STSP extreme point x^*, one can express $\frac{4}{3}x^*$ as a convex combination of leafless Eulerian graphs.*

Although the use of ultra-fundamental STSP extreme points is critical in our current proof, this theorem might still remain true if "ultra-fundamental STSP extreme point" is replaced in its statement by just "STSP extreme point". Also, this theorem works only when the ratio used is $\frac{4}{3}$, which we note in its proof and in the next section.

Proof of if part of Theorem 14: Suppose that for any STSP extreme point x^S, $\frac{4}{3}x^S$ can be expressed as a convex combination

$$\frac{4}{3}x^S = \sum_i \lambda^i \chi^{K_i} \tag{8}$$

of leafless Eulerian graphs. Let an ultra-fundamental ATSP extreme point x^A be given. Let x^S denote the corresponding STSP extreme point obtained by ignoring the directions on the edges. Note that x^S is an ultra-fundamental STSP extreme point. Express $\frac{4}{3}x^S$ as a convex combination of leafless Eulerian graphs. Consider an Eulerian graph K_i in the convex combination. We will prove that the edges of K_i can be directed to get a graph H_i which is a directed Eulerian subgraph of x^A, and we can express $\frac{4}{3}x^A$ as a convex combination of these H_i's, which completes the proof.

Consider a vertex v of K_i. If v is an internal vertex of a 1-path of x^S then, from Lemma 7 we see that this path must occur in each K_i either singly or doubly, and we direct the edges incident to v according to their direction in x^A. It is easy to see that we maintain $\text{indeg}(v) = \text{outdeg}(v)$ in this process.

We now show the fractional triangle edges of a gadget are never doubled. Consider such a fractional triangle of vertices u', w, and x, see Figure 4. Each of the following three mutually exclusive cases occurs exactly 1/3'rd of the time in a leafless Eulerian decomposition.

(a) The 1-edge into u' is doubled.

(b) The 1-edge into w is doubled.

(c) The 1-edge into x is doubled.

In each above case, the total number of triangle edges used in any Euler tour (including multiplicities) is at least 2 so as to avoid leaves. If any triangle edge is used twice in one of the Euler tours, this tour would use at least 3 triangle edges (including multiplicities) in all. But, the sum of the triangle fractional edge values is only 3/2, which when multiplied by 4/3 yields exactly

2. Therefore, 3 triangle edges may never be used in any Euler tour of the decomposition (since at least 2 triangle edges are used in every such Euler tour). This consideration yields the following unique solution for both the symmetric and asymmetric case.

(a) We use $(u', w), (u', x)$, and a 1-path from x to z.

(b) We use $(u', w), (x, w)$, and a 1-path from z to x.

(c) We use $(u', x), (x, w)$, a 1-path from z to x, and a 1-path from x to z.

So, if $v = u', w$, or x, we again have $indeg(v) = outdeg(v)$. Moreover, the 1/2 edges in the paths between x and z in the asymmetric case are being used in the correct amounts by the directed Eulerian tours in our decomposition.

Finally, consider when v is a vertex of degree 3 that is incident to two of the edges of the fractional Hamilton cycle in x^S. It has one 1-edge qv and two fractional edges vw and vz incident to it in x^S. Since K_i is leafless, at least one copy of qv must be present. Also there are at most two copies of qv. Since the fractional edges sum to 1 in x^S, their total usage has to be at most $\frac{4}{3}$. Their minimal usage in any K_i is one fractional edge if there is one copy of qv and two fractional edges if there are two copies of qv. Since qv occurs singly 2/3 of the time and doubly the other 1/3 by Lemma 7, this already leads to a combined usage of 4/3 for the fractional edges. Hence this indeed must be the actual usage. (In particular, when qv occurs singly, we cannot have 3 fractional edges, even though this satisfies the even degree constraint. However, we could have this case if the ratio in this theorem were not $\frac{4}{3}$ since the usage of these fractional edges may then be more than $\frac{4}{3}$, from which this theorem would not follow.) Now there are two possibilities for the edges incident to v in K_i:

1. Only one of the fractional edges occurs. That is either the edge qv occurs singly, and one of the two fractional edges, say vw also occurs singly or qv occurs doubly and vw occurs doubly. In this case if qv is directed into (out of) v in x^A, then vw must be directed out of (into) v and we can make these their directions in H_i.

2. qv occurs doubly and vw and vz occur singly. In this case too we direct the edges according to their direction in x^A.

In both cases we maintain $\text{indeg}(v) = \text{outdeg}(v)$ in H_i. Thus each H_i obtained is a directed Eulerian subgraph of x^A. From this and equation (8), it follows that

$$\frac{4}{3}x^A = \sum_i \lambda_i \chi^{H_i}.$$

This establishes the if implication of the theorem. □

The "only if" part of this proof relies on the following lemma, which has an interesting proof (follows this proof).

Lemma 8 *If the integrality gap of the ATSP relaxation is at most $\frac{4}{3}$, then for every ultra-fundamental ATSP extreme point x^*, we have that $\frac{4}{3}x^*$ can be expressed as a convex combination of directed Eulerian graphs where every 1-arc occurs singly in exactly $\frac{2}{3}$'rds of these graphs.*

Proof of only if part of Theorem 14: Suppose the integrality gap of the ATSP relaxation is at most $\frac{4}{3}$. Let an ultra-fundamental STSP extreme point x^S be given. By directing the edges appropriately, form a corresponding ultra-fundamental ATSP extreme point x^A. By Lemma 8, we have that $\frac{4}{3}x^A$ can be expressed as a convex combination of directed Eulerian graphs where every 1-arc occurs singly in exactly $\frac{2}{3}$'rds of these graphs. One is then free to have each 1-arc occur doubly in the other $\frac{1}{3}$'rd of these graphs. Then clearly, these directed Eulerian graphs can be leafless. By undirecting the arcs in these Eulerian graphs, one obtains that $\frac{4}{3}x^S$ can be expressed as a convex combination of the resulting Eulerian graphs. □

Proof of Lemma 8. Suppose the integrality gap of the ATSP relaxation is at most $\frac{4}{3}$. Let an ultra-fundamental ATSP extreme point x^* be given. Consider the corresponding fundamental ATSP extreme point $\overline{x}$. If $\overline{x}$ satisfies the consequent of this lemma, then it is an easy exercise to show that x^* will also.

We are going to use linear programming methods to form convex combinations of Eulerian graphs. Consider listing the incidence vectors of all the Eulerian graphs on the support graph of $\overline{x}$. Denote the index set by I. Denote the support graph arcs of $\overline{x}$ by E. Denote the set of 1-arcs by E_1. For each arc $e \in E$, let M_e, (resp. N_e) denote the index sets of all the Eulerian graphs where e occurs singly (resp. doubly). Let λ_i be the convex multiplier for the ith Eulerian graph. Consider the following linear program.

$$
\begin{array}{lcll}
\min \epsilon & & & \\
\text{subject to} & & & \\
\sum_{i\in M_e} \lambda_i + \sum_{i\in N_e} 2\lambda_i & \leq & \frac{4}{3}\overline{x}_e & \forall e \in E \\
\sum_{i\in M_e} \lambda_i & \leq & \frac{2}{3} + \epsilon & \forall e \in E_1 \\
\sum_{i\in M_e} \lambda_i & \geq & \frac{2}{3} - \epsilon & \forall e \in E_1 \\
\sum_{i\in I} \lambda_i & = & 1 & \\
\lambda_i & \geq & 0 & \forall i \in I
\end{array}
\tag{9}
$$

If the minimum of this linear program were always 0, then this lemma would follow. Suppose the minimum is $\epsilon^* > 0$. It is clear that each 1-arc $e \in E_1$ is present either singly or doubly in Eulerian graph i for every $i \in I$. Hence, the constraints

$$\sum_{i\in M_e} \lambda_i \geq \frac{2}{3} - \epsilon$$

are unnecessary in our linear program.

From the fundamental ATSP extreme point $\overline{x}$, form the ultra-fundamental ATSP extreme point x^k with k internal vertices in every 1-path. Refer back to figure 4. There are $k+1$ 1/2-arcs from z to x, which we call *up arcs*, and $k+1$ 1/2-arcs from x to z, which we call *down arcs*. Suppose every up arc down arc pair are either both present or both absent in the Eulerian graphs in a given convex combination for $\frac{4}{3}x^k$ exactly z^d of the time. Since, in every such Eulerian graph, at most one up arc down arc pair is absent, the most used up arc down arc pair is used at least $\frac{k}{k+1}z^d$ of the time. The remaining $1-z^d$ of the time, either all of the up arcs are present or all of the down arcs are present. Since the total usage in our convex combination of edges in a 1-path is at most $\frac{4}{3}$, we have

$$1 - z^d + 2\frac{k}{k+1}z^d \leq \frac{4}{3}.$$

Simplifying this yields

$$z^d \leq \frac{1}{3} \cdot \frac{k+1}{k-1}.$$

Denote by y^d the fraction of the time that the path from w to y is doubled. Denote by x^d the fraction of the time that the paths from u to u' and v' to v

are doubled, and by $x^s = 1 - x^d$ the fraction of the time that they are single. We clearly have

$$y^d \leq \frac{1}{3}.$$

We also have

$$x^d + y^d + z^d \geq 1,$$

from which we derive that

$$x^d \geq \frac{2}{3} - \frac{1}{3} \cdot \frac{k+1}{k-1}.$$

But this means that

$$x^s \leq \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{k-1}.$$

So, we choose k so that

$$\frac{1}{3} \cdot \frac{2}{k-1} < \epsilon^*.$$

Our given convex combination of Eulerian graphs trivially yields a convex combination of Eulerian graphs for $\frac{4}{3}\overline{x}$ by removing the extra vertices of x^k from each Eulerian graph. But this new convex combination contradicts the idea that the optimal solution to (9) is greater than or equal to ϵ^*, from which this lemma follows. $\square$

5 Remarks

The results of this paper lead us to formulate the following conjecture:

Conjecture 3 *For any ultra-fundamental extreme point x^* of the Held-Karp relaxation, $\frac{4}{3}x^*$ can be expressed as a convex combination of leafless Eulerian graphs.*

An affirmative resolution of the above conjecture would result in a nice advance in approximability. On the other hand, a negative answer (i.e., there is an extreme point x^* such that expressing $\frac{4}{3}x^*$ as a sum of Eulerian graphs requires that at least one of them has a leaf), would indicate the crucial role played by Eulerian graphs with leaves and thus provide a valuable insight towards resolving the integrality gap of the Held-Karp relaxation.

Finally, we would like to highlight a rather mysterious aspect of our reduction — 4/3 is the *only* constant for which it seems to work! To see this, refer back to Theorem 14, but replace 4/3 by an $r \in \mathbf{R}$ such that $r \geq 4/3$. In our theorem's proof, we first produce a convex decomposition of rx^S into leafless Eulerian graphs, where x^S is an arbitrary ultra-fundamental STSP extreme point. Then we consider the ultra-fundamental ATSP extreme point x^A that corresponds to x^S. We show that our convex decomposition of rx^S implies a convex decomposition of rx^A into directed Eulerian graphs by taking the leafless Eulerian graphs in the convex decomposition and directing each edge from x^S according to the direction of the arc in x^A that corresponds to it. However, this technique fails when the resulting indegree and outdegree of a vertex v differ.

But if we could guarantee that 1-edges always occurred singly or doubly in our convex combination, then the indegree and outdegree of v could not differ for the following reason. The usage of the fractional edges incident to v could never be less than the usage of the 1-edges incident to v in any of these Eulerian graphs because of the leafless condition, but the average usage of these fractional edges is the same as the average usage of the 1-edges. We can ensure (by Lemma 7) that 1-edges always occur either singly or doubly only for $r = 4/3$.

References

[1] S. Boyd and R. Carr, "Finding Low Cost TSP and 2-matching Solutions Using Certain Half-Integer Subtour Vertices," Working Paper, 1999.

[2] R. Carr and R. Ravi, "A New Bound for the 2-Edge Connected Subgraph Problem," IPCO Proceedings, 1998, 112-125.

[3] N. Christofides, "Worst-case analysis of a new heuristic for the traveling salesman problem," Tech. Rep., GSIA, Carnegie Mellon University, 1976.

[4] A. Frieze, G. Galbiati, and F. Maffioli, "On the worst-case performance of some algorithms for the asymmetric traveling salesman problem, " Networks 12, 23-39, 1982.

[5] M. Goemans and D. Bertsimas, "Survivable networks, linear programming relaxations and the parsimonious property," Mathematical Programming 60 (1993) 145-166.

[6] M. Goemans, "Worst Case Comparison of Valid Inequalities for the TSP," Mathematical Programming 69 (1995) 335-349.

[7] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, 1988.

[8] M. Held and R. M. Karp, "The traveling salesman problem and minimum spanning trees," Operations Research, 18, 1970, 1138-1162.

[9] L. Lovász, "On some connectivity properties of Eulerian graphs," Acta Mathematica Academiae Scientiarum Hungaricae 28 (1976) 129-138.

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

[11] C. Papadimitriou and M. Yannakakis, "The traveling salesman problem with distances one and two," Math. of Operations Research, 18(1), 1993, 1-11.

[12] S. Sahni and T. Gonzalez, "P-complete approximation problems," JACM, 23, 1976, 555-565.