MATHEMATICS OF OPERATIONS RESEARCH
Vol. 31, No. 2, May 2006, pp. 245–252
ISSN 0364-765X | EISSN 1526-5471 | 06 | 3102 | 0245
DOI 10.1287/moor.1060.0191

On the Integrality Ratio for the Asymmetric Traveling Salesman Problem

Moses Charikar
Department of Computer Science, Princeton University, 35 Olden Street, Princeton, New Jersey 08540, USA, moses@cs.princeton.edu

Michel X. Goemans
Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA, goemans@math.mit.edu

Howard Karloff
AT&T Labs—Research, 180 Park Avenue, Florham Park, New Jersey 07932, USA, howard@research.att.com

We improve the lower bound on the integrality ratio of the Held-Karp bound for asymmetric TSP with triangle inequality from 4/3 to 2.

Key words: asymmetric traveling salesman problem; Held-Karp relaxation; integrality ratio; approximation algorithm; ATSP
MSC2000 subject classification: Primary: 05C85, 68W25, 90C27, 90C59; secondary: 05C45, 68R05, 68R10, 68W40, 90C05, 90C10
OR/MS subject classification: Primary: networks/graphs—traveling salesman; secondary: analysis of algorithms—suboptimal algorithms; integer programming—theory
History: Received January 4, 2005; revised November 15, 2005.

1. Introduction. The traveling salesman problem (TSP)—the problem of finding a minimum cost tour through a set of cities—is the most celebrated combinatorial optimization problem. It is often used as a testbed for novel ideas, as was the case for Adleman's molecular computing (Adleman [1]), memetic algorithms (Moscato [23]), or ant colony optimization (Dorigo et al. [10]), to cite just a few examples. The TSP book (Lawler et al. [20]) provides a tour d'horizon of combinatorial optimization, illustrating all the concepts and techniques on the TSP.

The traveling salesman problem comes in two variants. The symmetric version (STSP) assumes that the cost c_{ij} of going from city i to city j is equal to c_{ji} while the more general asymmetric version (ATSP) does not make this assumption. In both cases, it is usually assumed—and we make this assumption in the rest of this paper—that we are in the metric case, i.e., the costs satisfy the triangle inequality: $c_{ij}+c_{jk}\geq c_{ik}$ for all i, j, k. Even though the TSP is the most studied combinatorial optimization problem, little progress has been made on its approximability in the general metric case in the last quarter of a century. Christofides in 1976 (Christofides [9]) discovered a $\frac{3}{2}$-approximation algorithm[1] for STSP. No better approximation algorithm has since been found for the general symmetric metric case. (For the special case of Euclidean instances, Arora [3] and Mitchell [21] found polynomial-time approximation schemes.) For the asymmetric case, no constant approximation algorithm is known. Frieze et al. [11] gave a simple $\log_2(n)$-approximation algorithm for ATSP in 1982, where n is the number of vertices. In the last two years, this was slightly improved to a guarantee of $0.999\log_2(n)$ by Bläser [4] and subsequently to $\frac{4}{3}\log_3(n)\approx 0.8412\log_2(n)$ by Kaplan et al. [19]. This is in sharp contrast to the best inapproximability result of Papadimitriou and Vempala [24] which shows the nonexistence of an α-approximation algorithm for ATSP for $\alpha=117/116-\epsilon$ and for STSP for $\alpha=220/219-\epsilon$, unless $P=NP$. Whether ATSP can be approximated within a constant factor is a major open question, and so is whether an α-approximation algorithm for STSP can be obtained for a constant $\alpha<\frac{3}{2}$.

Mathematical programming relaxations and especially linear programming relaxations have played a central role both in solving combinatorial optimization problems in practice (see, e.g., Applegate et al.'s [2] record exact solution of the STSP instance having all 24,978 cities in Sweden) and in the design and analysis of approximation algorithms. For the traveling salesman problem, Held and Karp [13, 14] introduced a linear programming relaxation for both the symmetric and asymmetric versions and gave several equivalent formulations for them. The bound they gave is often referred to as the *Held-Karp lower bound*. It can be computed in polynomial

[1] An α-approximation algorithm is a polynomial-time algorithm guaranteed to deliver a solution whose cost is within a factor of α of the optimum value.

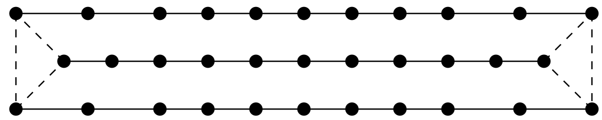

FIGURE 1. A 3-path configuration instance. The cost between two vertices i and j is equal to the length of the shortest path between i and j.

time and very efficiently in practice. As a linear program, this Held-Karp relaxation can be formulated for the symmetric case by writing the cut constraints and the degree constraints:[2]

$$\begin{aligned} \min \quad & \sum_{e\in E} c_e x_e \\ \text{subject to} \quad & x(\delta(S)) \geq 2, \quad \text{for all } \varnothing \neq S \neq V, \\ & x(\delta(v)) = 2, \quad \text{for all } v \in V, \\ & 0 \leq x_e \leq 1, \quad \text{for all } e \in E, \end{aligned} \tag{1}$$

and in the asymmetric case by writing the cut constraints and the in-degree and out-degree constraints:[3]

$$\begin{aligned} \min \quad & \sum_{e\in E} c_e x_e \\ \text{subject to} \quad & x(\delta^+(S)) \geq 1, \quad \text{for all } \varnothing \neq S \neq V, \\ & x(\delta^+(v)) = 1, \quad \text{for all } v \in V, \\ & x(\delta^-(v)) = 1, \quad \text{for all } v \in V, \\ & 0 \leq x_e \leq 1, \quad \text{for all } e \in E. \end{aligned}$$

One can easily see that the asymmetric formulation is equivalent to the symmetric one for the special case of symmetric costs and, therefore, it makes sense to denote by HK the Held-Karp bound both in the symmetric and asymmetric cases.

The quality of the Held-Karp lower bound has puzzled researchers for over 25 years. Wolsey [27] and later Shmoys and Williamson [25] showed that for symmetric (metric) instances we have $STSP \leq \frac{3}{2}HK$, where $STSP$ denotes the optimum value of an STSP instance. This was shown by proving that the tour returned by Christofides' [9] heuristic has cost bounded above by $\frac{3}{2}HK$. For ATSP, Williamson [26] showed that the Frieze et al. [11] heuristic returns a tour of cost at most $\log_2(n)HK$, thereby showing that $ATSP/HK \leq \log_2(n)$, where $ATSP$ is the optimum value of an ATSP instance. On the other hand, the worst family that was known for both integrality ratios $STSP/HK$ and $ATSP/HK$ was symmetric and achieved a ratio arbitrarily close to $\frac{4}{3}$; see Figure 1, whose solid edges e have $x_e = 1$, whose alternating dark and light edges e have $x_e = \frac{1}{2}$, and whose missing edges e have $x_e = 0$. However, instances with large integrality ratios for STSP or ATSP are not easy to encounter "in practice." Based on Johnson and McGeoch [16] and Johnson et al. [17, 18], D. S. Johnson [15] states that the largest percentage integrality gap $(OPT - HK)/HK$ he has seen for an *asymmetric* testbed instance is 2.8%. In contrast, the largest integrality gap he has seen for a *symmetric* testbed instance is 9.55%, achieved on the 225-node tsplib instance `ts225` which was specifically designed to foil TSP software. Except for that one instance, Johnson is not aware of any testbed instance, symmetric or not, with integrality gap exceeding 3%.

In a systematic search for large ratios, Boyd and Labonté [6] and Boyd and Elliott [5] showed that the integrality ratio is at most $\frac{4}{3}$ for STSP instances with at most ten vertices and for ATSP instances with at most seven vertices. They were not able to carry out their experiments beyond these values. We provide a recursive construction leading to ATSP instances whose integrality ratio $ATSP/HK$ approaches 2 arbitrarily closely. The first three levels of our construction are illustrated in Figure 2.

A longstanding conjecture formulated at least twenty years ago states that the integrality ratio $STSP/HK$ for STSP instances is at most $\frac{4}{3}$. There have been many unsuccessful attempts to prove it. A constructive proof of this conjecture would likely lead to a $\frac{4}{3}$-approximation algorithm for STSP. One classical reformulation of the $\frac{4}{3}$

[2] The latter are not necessary for metric instances (Goemans and Bertsimas [12]).

[3] For ATSP metric instances, we can replace the two degree constraints by $x(\delta^-(i)) = x(\delta^+(i))$.

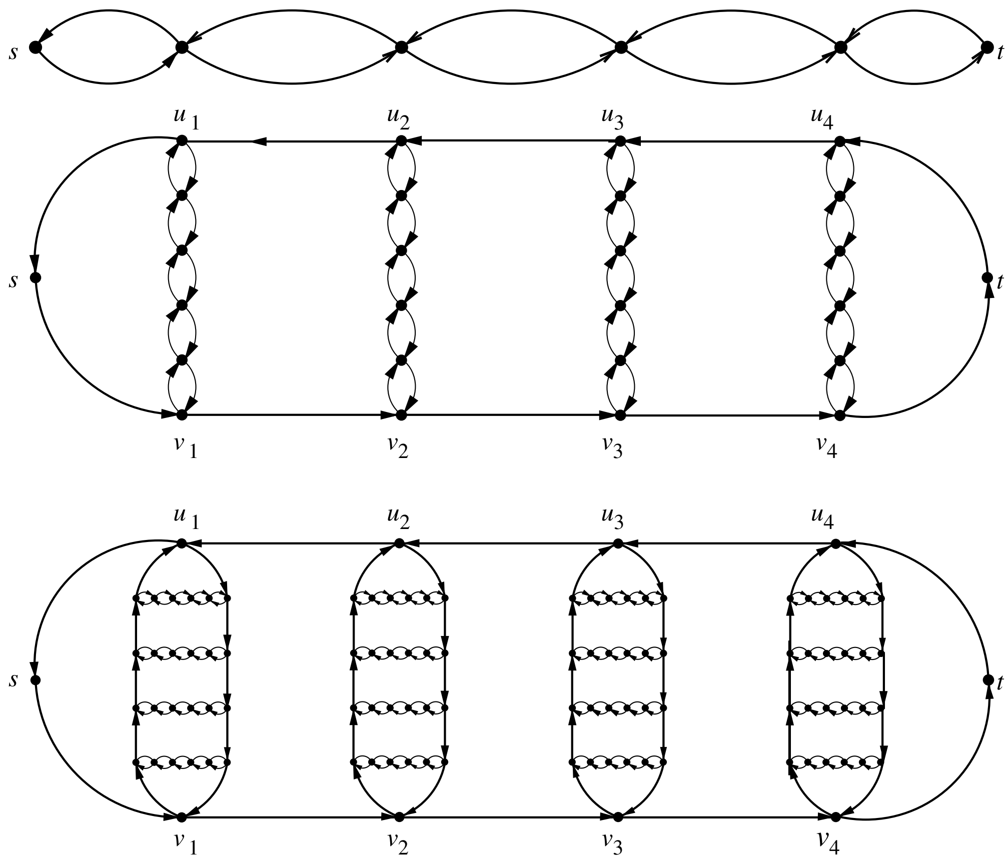

FIGURE 2. An illustration of three levels of our construction: digraphs G_1, G_2, and G_3 for $r=4$, with the arc costs not shown.

conjecture is that any extreme point of the subtour polytope (1) after being multiplied by $\frac{4}{3}$ can be decomposed into a convex combination of Eulerian subgraphs. Carr and Vempala [8] considered what may appear as a slight strengthening of this $\frac{4}{3}$ conjecture—that the decomposition is of *leafless* Eulerian subgraphs, i.e., subgraphs in which the degree-2 vertices are adjacent to two *different* neighbors. They showed that this slight strengthening would imply their conjecture that the integrality ratio for ATSP is also bounded by $\frac{4}{3}$. Our construction refutes their ATSP $\frac{4}{3}$ conjecture and, therefore, also their strengthened the STSP $\frac{4}{3}$ conjecture. It, however, leaves open whether the integrality ratio for STSP instances is bounded by $\frac{4}{3}$.

Our construction of instances for ATSP proves that the worst-case integrality ratio for ATSP (at least 2) is provably larger than that for STSP instances (at most 1.5).

Our construction has a few other interesting features. First, it requires a number of vertices exponential in $1/\epsilon$ in order to achieve a ratio of $2-\epsilon$ while for STSP the worst-case integrality ratio, if equal to $\frac{4}{3}$, would require a number of vertices linear in $1/\epsilon$ to achieve $\frac{4}{3}-\epsilon$ (Monma et al. [22]). Second, our feasible (but not necessarily optimal) solution of the linear programming relaxation for our instances is half-integral (i.e., $x_e \in \{0, \frac{1}{2}, 1\}$). Although the LP relaxation (both in the symmetric and asymmetric versions) contains extreme points which are not half-integral (Boyd and Pulleyblank [7]), proving that the worst-case behavior is attained for instances for which the optimum LP solution is half-integral would allow one to easily prove an integrality ratio of 2 for all ATSP instances: Simply multiply the extreme point by 2 and interpret the resulting vector as the incidence vector of an Eulerian subgraph. Third, our construction is obtained by defining costs on the arcs of a digraph and letting the cost between any two vertices i and j be the shortest-path length from i to j. For our construction to work we need the arc costs to vary from arc to arc.

In the next section, we present our recursive construction of an instance with integrality ratio approaching 2 for ATSP. Even one level of the recursion (the middle part of Figure 2) leads to instances with integrality ratio approaching $\frac{3}{2}$ ($k=2$ in Theorem 1.1 below), thereby already refuting the $\frac{4}{3}$ conjectures. The difficulty in the analysis lies not in computing the Held-Karp bound (or an upper bound on it), but in computing a lower bound on the cost of the optimum tour. Our main result is the following.

THEOREM 1.1. *For $r \geq 3$ and $k \geq 2$, we construct an ATSP instance with $\Theta(r^k)$ vertices for which*

$$\frac{ATSP}{HK} \geq \frac{r-1}{r+1} \cdot \frac{2k-1}{k},$$

which tends to 2 *as r and k tend to infinity.*

2. Definitions.

DEFINITION 2.1. An *Eulerian subdigraph H* of a digraph $G = (V, E)$ is V together with a collection of arcs of *G with multiplicities* such that (i) the indegree of every vertex equals its outdegree, and (ii) H is weakly connected (i.e., the underlying undirected graph is connected).

(Strictly speaking, an "Eulerian subdigraph" H of a digraph G is not a subdigraph in the traditional sense, because the multiplicity of an arc in the subdigraph may exceed the multiplicity of the same arc in G (i.e., 1) and, hence, H is really a "multidigraph.")

DEFINITION 2.2. The *deficiency* of a vertex v in a digraph G is equal to its outdegree minus its indegree and is denoted $\mathrm{def}_G(v)$.

An Eulerian subdigraph is therefore precisely a weakly connected subdigraph with no deficiency at any vertex. An Eulerian subdigraph can be traversed by a walk that starts and ends at the same vertex and visits every vertex at least once.

Given a digraph $G = (V, A)$ with arc costs w_a, the ATSP asks for an Eulerian subdigraph of minimum total cost, i.e., one that minimizes $\sum_{a \in A} m_a w_a$, where m_a denotes the multiplicity of arc a in the Eulerian subdigraph. This is equivalent to finding the minimum cost tour for the shortest-path metric corresponding to the arc costs w_a.

DEFINITION 2.3. An *(s, t)-Eulerian subdigraph F* of a digraph G on V having vertices s and t, $s \neq t$, is V together with a collection of arcs of G with multiplicities such that F becomes an Eulerian subdigraph of G if the vertices s and t are identified (or contracted or shrunk). The (s, t)-Eulerian subdigraph is said to be *closed* if the deficiency of s (and t) is 0 and *open* otherwise.

In other words, a closed (s, t)-Eulerian subdigraph has no deficiency at any vertex but, if not Eulerian, consists of two weakly connected components, every vertex being weakly connected to either s or t. On the other hand, an open (s, t)-Eulerian subdigraph *is* always weakly connected but the deficiency of both s and t is nonzero, the deficiency of any other vertex being zero.

3. Bad instances and their analysis. We now define a family (G_k) of digraphs with costs; our family (L_k) of instances of the ATSP will be derived from the family (G_k). Fix any positive integer r for the remainder of the paper; large r's will give integrality ratios approaching 2 as r approaches infinity. G_1 consists of a bidirected path of $r+2$ vertices, starting at a distinguished "source" s and ending at a distinguished "sink" t whose $2(r+1)$ arcs all have cost 1. For each $k \geq 2$, we build G_k as follows (see Figure 2). Start with $2+2r$ distinct vertices s; $u_1, u_2, \ldots, u_r$; $v_r, v_{r-1}, \ldots, v_1$; t, with s and t being the source and sink, respectively. For simplicity, let $u_0 = v_0 = s$ and $u_{r+1} = v_{r+1} = t$. Add arcs (u_i, u_{i+1}) and arcs (v_{i+1}, v_i) for $i = 0, 1, \ldots, r$, all of cost r^{k-1}. Then, add r vertex-disjoint isomorphic copies of G_{k-1} using new vertices, except that the ith copy of G_{k-1}, $1 \leq i \leq r$, uses u_i as its source and v_i as its sink. Hence, $|V(G_k)| = 2 + r|V(G_{k-1})|$, i.e., $|V(G_k)| = r^k + 2\sum_{i=0}^{k-1} r^i$, which is $\Theta(r^k)$.

Let t_k be the sum of the costs of the arcs in G_k. Clearly, $t_1 = 2(r+1)$ and $t_k = 2(r+1)r^{k-1} + rt_{k-1}$ for $k \geq 2$, whose solution is

$$t_k = (2k)r^{k-1}(r+1)$$

for all $k \geq 1$.

Let a_k denote the minimum cost of an open (s, t)-Eulerian subdigraph of G_k, and b_k the minimum cost of a closed (s, t)-Eulerian subdigraph. The next proposition provides a recursive formula to bound these quantities from below.

PROPOSITION 3.1. $a_1 = r + 1$, $b_1 = 2r$, *and for* $k \geq 2$,

$$a_k \geq \min_{l \in \{0,1,2,\ldots,r\}} \{(l+r+1)r^{k-1} + la_{k-1} + (r-l)b_{k-1}\}$$

and

$$b_k \geq \min_{l \in \{0,1,2,\ldots,r\}} \{(l+2r-2)r^{k-1} + la_{k-1} + (r-l)b_{k-1}\}.$$

Proof. That $a_1 = r+1$ and $b_1 = 2r$ are obvious.

Let $k \geq 2$. Let F be an (s,t)-Eulerian subdigraph of G_k. Let H_i denote the copy of G_{k-1} for which u_i and v_i are the source and sink, respectively. Since H_i has only u_i and v_i in common with the rest of G_k, F restricted to H_i (denoted by F_i) induces a (u_i, v_i)-Eulerian subdigraph of H_i (since identifying u_i and v_i in F_i is equivalent to shrinking $V(G_k)\backslash(V(H_i)\backslash\{u_i, v_i\})$ in G_k). Each F_i can be open or closed (regardless of whether F is open or closed). Let l denote the number of open (u_i, v_i)-Eulerian subdigraphs F_i. The number of closed ones is therefore $r-l$. The total cost of the arcs in the union of the F_i's is therefore at least $la_{k-1} + (r-l)b_{k-1}$ and this accounts for the second and third terms in the statement of the lemma. We now need to evaluate the total cost of the remaining arcs in F.

For $i = 0, \ldots, r$, let $p_i \geq 0$ denote the multiplicity of arc (u_i, u_{i+1}) in F and $q_i \geq 0$ the multiplicity of arc (v_{i+1}, v_i). The total cost of the arcs in $F\backslash\bigcup_{i=1}^r F_i$ is thus $r^{k-1}\sum_{i=0}^r (p_i + q_i)$. Our goal is to show that

$$\sum_{i=0}^r (p_i + q_i) \geq \begin{cases} l + r + 1 & \text{if } F \text{ is open,} \\ l + 2r - 2 & \text{if } F \text{ is closed.} \end{cases} \tag{2}$$

In order to derive this, we need two basic observations. First, for any digraph G and for any subset S of vertices of $V(G)$, the sum of the deficiencies of the vertices in S is equal to the number of arcs leaving S minus the number of arcs entering S, i.e.,

$$\sum_{v\in S} \mathrm{def}_G(v) = |\delta^+(S)| - |\delta^-(S)|.$$

Applying this to F with $S = \{s\} \cup V(H_1) \cup V(H_2) \cup \cdots \cup V(H_i)$ where $i \in \{0, \ldots, r\}$, we obtain that $p_i - q_i = \mathrm{def}_F(s)$ for $i = 0, 1, \ldots, r$, because F has deficiency 0 at all vertices other than s and t.

The second observation is that

$$\mathrm{def}_{F_i}(u_i) = p_{i-1} - p_i = q_{i-1} - q_i$$

(where $i \in \{1, \ldots, r\}$). This follows from the facts that $\mathrm{outdeg}_F(u_i) = p_i + \mathrm{outdeg}_{F_i}(u_i)$, that $\mathrm{indeg}_F(u_i) = p_{i-1} + \mathrm{indeg}_{F_i}(u_i)$, and that $0 = \mathrm{def}_F(u_i) = \mathrm{outdeg}_F(u_i) - \mathrm{indeg}_F(u_i) = p_i - p_{i-1} + \mathrm{def}_{F_i}(u_i)$, implying that $\mathrm{def}_{F_i}(u_i) = p_{i-1} - p_i$; now $p_i - q_i = \mathrm{def}_F(s)$ for all i implies that $p_i - q_i = p_{i-1} - q_{i-1}$ and that $p_{i-1} - p_i = q_{i-1} - q_i$.

From this second observation we conclude that F_i is closed iff $p_{i-1} = p_i$ iff $q_{i-1} = q_i$.

We now prove (2). Let us first assume that F is open. Without loss of generality we can assume that $\mathrm{def}_F(s) > 0$, and thus we have $p_i - q_i \geq 1$ for all i. Thus, $\sum_{i=0}^r (p_i + q_i) = \sum_{i=0}^r (p_i - q_i + 2q_i) \geq r + 1 + 2\sum_{i=0}^r q_i$. We claim that $\sum_{i=0}^r q_i \geq l/2$ where l is the number of open (u_i, v_i)-Eulerian subdigraphs F_i. Indeed, if q_i were 0 for all i, l would be 0 by our second observation. Whenever we increase some q_i by 1, l can increase by at most two units since only F_i and F_{i+1} could now be open. This proves our claim and thus $\sum_{i=0}^r (p_i + q_i) \geq r + 1 + 2\sum_{i=0}^r q_i \geq r + 1 + l$, proving Equation (2) if F is open.

Now assume that F is closed. We have $p_i = q_i$ for $i = 0, 1, \ldots, r$, implying that $\sum_{i=0}^r (p_i + q_i) = 2\sum_{i=0}^r p_i$. Observe that there can be at most one index $x \in \{0, 1, \ldots, r\}$ with $p_x = 0$. Otherwise, if $p_x = p_y = 0$ for $x < y$, then $V(H_{x+1}) \cup \cdots \cup V(H_y)$ would be disconnected from the rest of the graph and F would induce an undirected graph with at least three components (as a closed (s,t)-Eulerian subdigraph, F is allowed to induce at most two). Therefore, the vector p is greater than or equal to p' where p' is a vector of all 1's except for one 0 suitably placed (at x if $p_x = 0$ and anywhere otherwise). In p', the number of consecutive entries that differ is at most 2 and the corresponding l' would therefore satisfy $l' \leq 2$. As we increase p' toward p one entry at a time, l' increases by at most 2 for every increase of p by the second observation, implying that

$$l \leq 2 + 2\left(\left(\sum_{i=0}^r p_i\right) - r\right);$$

that is,

$$2\sum_{i=0}^r p_i \geq l + 2r - 2,$$

proving Equation (2) if F is closed. □

The following lemma gives a solution to the recurrence of Proposition 3.1.

Lemma 3.1. *Let a_k and b_k be as defined in Proposition 3.1. Then, $a_k \geq (2r-2)kr^{k-1} - (r-3)r^{k-1}$ and $b_k \geq (2r-2)kr^{k-1}$.*

Proof. Let $a'_1 = r+1$, $b'_1 = 2r-2$, and for $k \geq 2$, let

$$a'_k = \min_{l \in \{0,1,2,\ldots,r\}} \{(l+r+1)r^{k-1} + la'_{k-1} + (r-l)b'_{k-1}\},$$

and let

$$b'_k = \min_{l \in \{0,1,2,\ldots,r\}} \{(l+2r-2)r^{k-1} + la'_{k-1} + (r-l)b'_{k-1}\}.$$

The only differences with the recurrences for a_k and b_k are the value of b'_1 and the fact that inequalities have been replaced by equalities. By induction on k, we have $a_k \geq a'_k$ and $b_k \geq b'_k$ for every $k \geq 1$.

Given these expressions for a'_k and b'_k, we have $b'_k = a'_k + (r-3)r^{k-1}$ for $k \geq 2$ and this is true also for $k=1$ given the values chosen for a'_1 and b'_1. We can therefore rewrite the expression for b'_k as

$$b'_k = \min_{l \in \{0,1,2,\ldots,r\}} \{(l+2r-2)r^{k-1} + rb'_{k-1} - l(r-3)r^{k-2}\}.$$

As the coefficient of l in b'_k is $3r^{k-2}$, the expression is minimized for $l=0$; thus,

$$b'_k = (2r-2)r^{k-1} + rb'_{k-1}.$$

Since $b'_1 = 2r-2$, this implies that $b'_k = (2r-2)kr^{k-1}$ for all k.

The fact that $b_k \geq b'_k$ implies the second part of the lemma while the first follows from $a_k \geq a'_k$ and $b'_k = a'_k + (r-3)r^{k-1}$. □

From G_k we define a new digraph L_k with costs, as follows. Let $V(L_k) = V(G_k) \backslash \{s, t\}$ and replace the arcs (s, u_1), (u_r, t), (t, v_r), and (v_1, s) of G_k with new arcs (u_r, u_1) and (v_1, v_r) in L_k of cost r^{k-1}. L_k inherits all other arcs of G_k. The total cost of L_k is thus equal to the total cost of G_k minus $2r^{k-1}$ and is therefore less than $t_k = 2kr^{k-1}(r+1)$.

We will now derive first a lower bound on the cost of any Eulerian subdigraph of L_k and then an upper bound on the Held-Karp lower bound for this instance.

Lemma 3.2. *For $r \geq 3$ and $k \geq 2$, the cost of any Eulerian subdigraph of L_k is at least* $(2k-1)(r-1)r^{k-1}$.

Proof. We proceed in a similar fashion as for G_k and use the same notation and arguments as in the proof of Proposition 3.1. Consider any Eulerian subdigraph F of L_k. Let l denote the number of open (u_i, v_i)-Eulerian subdigraphs F_i. As for G_k, the total cost within the F_i's is at least $la_{k-1} + (r-l)b_{k-1}$. We claim that the other arcs, those between the u_i's and those between the v_i's, contribute at least $(l+r-1)r^{k-1}$. Assuming this claim, the total cost of F is at least $(l+r-1)r^{k-1} + la_{k-1} + (r-l)b_{k-1}$ and by the proof of Lemma 3.1 this is at least $a'_k - 2r^{k-1} \geq (2r-2)kr^{k-1} - (r-1)r^{k-1} = (2k-1)(r-1)r^{k-1}$, proving the desired bound.

To prove the claim, let $p_i \geq 0$ denote the multiplicity of the arc (u_i, u_{i+1}) in F for $i = 1, \ldots, r$. Here, as in the rest of this proof, all indices have to be interpreted cyclically and thus $r+1$ represents 1 and 0 represents r. Similarly, let $q_i \geq 0$ denote the multiplicity in F of the arc (v_{i+1}, v_i). The arcs between the u_i's and those between the v_i's contribute $\sum_{i=1}^r (p_i + q_i) r^{k-1}$ and, thus, our goal is to prove that $\sum_{i=1}^r (p_i + q_i) \geq l + r - 1$.

As in Proposition 3.1, F_i is closed iff $p_{i-1} = p_i$ iff $q_{i-1} = q_i$, and this implies that

$$l \leq 2\left(\sum_{i=1}^r p_i\right) = 2\left(\sum_{i=1}^r q_i\right). \tag{3}$$

Since we must have the same positive number of arcs leaving and entering $V(H_i)$, we have $p_{i-1} + q_i = p_i + q_{i-1} \geq 1$. This implies that $p_i - q_i = p_{i-1} - q_{i-1}$ for all i.

We distinguish two cases: either (i) $p_i - q_i = 0$ for all i, or (ii) $p_i - q_i \geq 1$ for all i (without loss of generality). In case (i), i.e., $p_i = q_i$ for all i, we cannot have $p_j = q_j = 0$ and $p_k = q_k = 0$ for two indices $j \neq k$. Otherwise, $V(H_{j+1}) \cup V(H_{j+2}) \cup \cdots \cup V(H_k)$ would have no arcs coming in or out. Thus, $\sum_{i=1}^r p_i = \sum_{i=1}^r q_i \geq r-1$ with equality only if all p_i's are 1 except one which is 0 (and similarly for the q_i's), implying that at equality $l=2$ because l is equal to the number of i such that $p_{i-1} \neq p_i$. Thus, either $\sum_{i=1}^r p_i \geq r$ and, hence, $\sum_{i=1}^r (p_i + q_i) \geq 2r \geq l + r$, or $\sum_{i=1}^r p_i = r-1$ and, hence, $\sum_{i=1}^r (p_i + q_i) = 2(r-1) \geq r+1 \geq l+r-1$ for $r \geq 3$. This proves the claim in case (i). For case (ii), we have $p_i - q_i \geq 1$ for all i, which together with Equation (3) implies that $\sum_{i=1}^r (p_i + q_i) \geq \sum_{i=1}^r (p_i - q_i) + 2\sum_{i=1}^r q_i \geq r + l$, also proving the claim. □

We will now prove that the asymmetric LP relaxation has a half-integral solution of low cost.

Lemma 3.3. *The vector x with $x_a = \frac{1}{2}$ for every arc a is feasible for the asymmetric subtour polytope for L_k. Hence, $HK \leq t_k/2 = kr^{k-1}(r+1)$.*

Proof. The lemma follows from the claims that (i) L_k has exactly two arcs entering any vertex and exactly two arcs leaving any vertex, and that (ii) L_k has at least two arcs in both directions in any cutset.

The first claim follows easily by induction by proving that all vertices of G_k have indegree and outdegree 2, except s and t which have indegree and outdegree 1.

By Menger's theorem, the second claim can be rephrased by fixing any one vertex of L_k, say, u_1, and saying that for every vertex v, there exist in L_k two arc-disjoint paths from v to u_1 and two arc-disjoint paths from u_1 to v. We will show the existence of two arc-disjoint paths to u_1; this will be enough since L_k is Eulerian, implying that every cutset has the same number of arcs in either direction.

To prove the existence of these two arc-disjoint paths to u_1, we first argue by induction on k that there exist in G_k two arc-disjoint paths from every vertex v, one to s and the other to t. This is clear since if $v \in H_i$, then the two arc-disjoint paths to u_i and v_i in H_i (a copy of G_{k-1}) which exist by our inductive hypothesis can be extended: the first to t (by visiting $u_{i+1}, \ldots, u_r, t$) and the other to s (by visiting $v_{i-1}, \ldots, v_1, s$). Now, to prove our second claim, consider any vertex $v \in H_i$. We know that there exist two arc-disjoint paths within H_i, one from v to u_i (call it P_1), and the other from v to v_i (call it P_2). P_1 can be extended to u_1 by using the arcs $(u_i, u_{i+1}), \ldots, (u_{r-1}, u_r), (u_r, u_1)$ to get a path P_1' from v to u_1. If $i \neq 1$, then we can extend P_2 from v_i to v_1 via the arcs $(v_i, v_{i-1}), \ldots, (v_2, v_1)$ and then from v_1 to u_1 through any directed path within H_1; the resulting path will be arc-disjoint from P_1'. If $i = 1$, then P_1' does not use the arc (u_r, u_1) and, therefore, we can extend P_2 first from v_1 to v_r by adding the arc (v_1, v_r), then to u_r through a directed path in H_r, and finally from u_r to u_1 through the arc (u_r, u_1). □

Combining Lemmas 3.2 and 3.3, we get a restatement of Theorem 1.1.

Theorem 3.1. *For L_k with $k \geq 2$ and $r \geq 3$, the ratio between the minimum cost of any Eulerian subdigraph and the asymmetric Held-Karp lower bound is at least $((r-1)/(r+1)) \cdot (2k-1)/k$, which can be made arbitrarily close to 2 by choosing k and r sufficiently large.*

Acknowledgments. The authors thank the referees for improving the paper with their pithy, relevant, and helpful comments.

References

[1] Adleman, L. M. 1994. Molecular computation of solutions to combinatorial problems. *Science* **266** 1021–1024.
[2] Applegate, D., R. Bixby, V. Chvátal, W. Cook, K. Helsgaun. The traveling salesman problem homepage. http://www.tsp.gatech.edu.
[3] Arora, S. 1998. Polynomial-time approximation schemes for Euclidean TSP and other geometric problems, *J. Assocation Comput. Machinery* **45**(5) 753–782.
[4] Bläser, M. 2003. A new approximation algorithm for the asymmetric TSP with triangle inequality. *Proc. 14th Annual ACM-SIAM Sympos. Discrete Algorithms*, Baltimore, MD, 638–645.
[5] Boyd, S., P. Elliott. 2004. Finding the exact integrality gap for small ATSP's. *Combinatorial Optimisation Conference* (*CO2004*), Lancaster, England.
[6] Boyd, S., G. Labonté. 2002. Finding the exact integrality gap for small traveling salesman problems. *Proc. 9th Conf. Integer Programming Combin. Optim.*, 83–92.
[7] Boyd, S., W. R. Pulleyblank. 1990. Optimizing over the subtour polytope of the traveling salesman problem. *Math. Programming* **49** 163–187.
[8] Carr, R., S. Vempala. 2004. Towards a 4/3 approximation algorithm for the asymmetric TSP. *Math. Programming* **100** 569–587.
[9] Christofides, N. 1976. Worst-case analysis of a new heuristic for the traveling salesman problem. Technical report, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA.
[10] Dorigo, M., V. Maniezzo, A. Colorni. 1996. The ant system: Optimization by a colony of cooperating agents. *IEEE Trans. Systems, Man, Cybernetics, Part B* **26** 29–41.
[11] Frieze, A. M., G. Galbiati, M. Maffioli. 1992. On the worst-case performance of some algorithms for the asymmetric traveling salesman problem. *Networks* **12** 23–39.
[12] Goemans, M. X., D. Bertsimas. 1993. Survivable networks, linear programming relaxations and the parsimonious property. *Math. Programming* **60** 145–166.
[13] Held, M., R. M. Karp. 1970. The traveling salesman problem and minimum spanning trees. *Oper. Res.* **18** 1138–1162.
[14] Held, M., R. M. Karp. 1971. The traveling-salesman problem and minimum spanning trees: Part II. *Math. Programming* **1** 6–25.
[15] Johnson, D. S. 2004. Personal communication.
[16] Johnson, D. S., L. A. McGeoch. 2002. Experimental analysis of heuristics for the STSP. G. Gutin, A. Punnen, eds. *The Traveling Salesman Problem and its Variations.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 369–443.
[17] Johnson, D. S., L. A. McGeoch, E. E. Rothberg. 1996. Asymptotic experimental analysis for the Held-Karp traveling salesman bound. *Proc. 7th ACM-SIAM Sympos. Discrete Algorithms*, 341–350.
[18] Johnson, D. S., G. Gutin, L. A. McGeoch, A. Yeo, W. Zhang, A. Zverovich. 2002. Experimental analysis of heuristics for the ATSP. G. Gutin, A. Punnen, eds. *The Traveling Salesman Problem and its Variations*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 445–487.
[19] Kaplan, H., M. Lewenstein, N. Shafir, M. Sviridenko. 2003. Approximation algorithms for asymmetric TSP by decomposing directed regular multidigraphs. *Proc. 44th Annual IEEE Sympos. Foundations Comput. Sci.*, Cambridge, MA, 56–67.

[20] Lawler, E. L., J. K. Lenstra, A. H. G. Rinnooy Kan, D. B. Shmoys. 1985. *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. John Wiley & Sons Ltd.

[21] Mitchell, J. S. B. 1999. Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, k-MST, and related problems. *SIAM J. Comput.* **28** 1298–1309.

[22] Monma, C., B. Munson, W. Pulleyblank. 1990. Minimum-weight two-connected spanning networks. *Math. Programming* **46** 153–171.

[23] Moscato, P. 1989. On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms. Report 790, Caltech Concurrent Computation Program, Pasadena, CA.

[24] Papadimitriou, C. H., S. Vempala. 2000. On the approximability of the traveling salesman problem. *Proc. 32nd Annual ACM Sympos. Theory Comput.*, Portland, OR, 126–133.

[25] Shmoys, D., D. Williamson. 1990. Analyzing the Held-Karp TSP bound: A monotonicity property with application. *Inform. Processing Lett.* **35** 281–285.

[26] Williamson, D. P. 1990. Analysis of the Held-Karp heuristic for the traveling salesman problem. M.S. thesis, Massachusetts Institute of Technology.

[27] Wolsey, L. A. 1980. Heuristic analysis, linear programming and branch and bound. *Math. Programming Stud.* **13** 121–134.