To appear in *Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, January 28-30, 1996.

Asymptotic Experimental Analysis for the Held-Karp Traveling Salesman Bound

D. S. Johnson * L. A. McGeoch † E. E. Rothberg ‡

Abstract

The Held-Karp (HK) lower bound is the solution to the linear programming relaxation of the standard integer programming formulation of the traveling salesman problem (TSP). For numbers of cities N up to 30,000 or more it can be computed exactly using the Simplex method and appropriate separation algorithms, and for N up to a million good approximations can be obtained via iterative Lagrangean relaxation techniques first suggested by Held and Karp. In this paper, we consider three applications of our ability to compute/approximate this bound.

First, we provide empirical evidence in support of using the HK bound as a stand-in for the optimal tour length when evaluating the quality of near-optimal tours. We show that for a wide variety of randomly generated instance types the optimal tour length averages less than 0.8% over the HK bound, and even for the real-world instances in `TSPLIB` the gap is almost always less than 2%. Moreover, our data indicates that the HK bound can provide substantial ''variance reduction'' in experimental studies involving randomly generated instances. Second, we estimate the expected HK bound as a function of N for a variety of random instance types, based on extensive computations. For example, for random Euclidean instances it is known that the ratio of the Held-Karp bound to $\sqrt{N}$ approaches a constant C_{HK}, and we estimate both that constant and the rate of convergence to it. Finally, we combine this information with our earlier results on expected HK gaps to obtain estimates for expected optimal tour lengths. For random Euclidean instances, we conclude that C_{OPT}, the limiting ratio of the optimal tour length to $\sqrt{N}$, is .7124 ± .0002, thus invalidating the commonly cited estimates of .749 and .765 and undermining many claims of good heuristic performance based on those estimates. For random distance matrices, the expected optimal tour length appears to be about 2.042, adding support to a conjecture of Krauth and Mézard.

1. Introduction

In the Traveling Salesman Problem, or ''TSP,'' we are given a set $\{c_1, c_2, \cdots, c_N\}$ of *cities* and for each pair $\{c_i, c_j\}$ of distinct cities a *distance* $d(c_i, c_j)$. Our goal is to find an ordering σ of the cities that minimizes the *tour length*, i.e, the quantity

$$\sum_{i=1}^{N-1} d(c_{\sigma(i)}, c_{\sigma(i+1)}) + d(c_{\sigma(N)}, c_{\sigma(1)}) .$$

We are concerned here with the *symmetric* TSP, in which the distances satisfy $d(c_i, c_j) = d(c_j, c_i)$ for $1 \le i, j \le N$. This problem has a long history and many applications, e.g., see [23]. It can be formulated as an integer program based on the interpretation of the optimal tour as a minimum-length Hamiltonian circuit in the complete graph with cities as vertices. We use the variable x_{ij} to represent the edge between cities c_i and c_j, $1 \le i < j \le N$, taking $x_{ij} = 1$ to mean that edge $\{c_i, c_j\}$ is in the tour. The integer program is then

$$\begin{array}{lll}
\text{Minimize} & \displaystyle\sum_{i=1}^{N} \sum_{j=i+1}^{N} d(c_i, c_j) \cdot x_{ij} & \\
\text{Subject to} & \displaystyle\sum_{i=k \text{ or } j=k} x_{ij} = 2, & 1 \le k \le N, \\
& \displaystyle\sum_{|S \cap \{i,j\}|=1} x_{ij} \ge 2, & \text{for all } S \subset \{1,2,\ldots,N\}, \\
\text{and} & x_{ij} \in \{0,1\}, & 1 \le i < j \le N
\end{array}$$

If we relax this integer program by replacing the last constraint by $0 \le x_{ij} \le 1$, we obtain a linear program (LP) whose solution will be a lower bound on the optimal tour length, called the *Held-Karp* lower bound (*HK bound*) based on its early investigation by Held and Karp in [10,11].

Note that this LP can be solved in polynomial time, despite the fact that it contains an exponential number of ''subtour constraints'' (those involving the subset S). This is because there exists a polynomial-time ''separation oracle'' for the subtour constraints (based on calls to a max-flow algorithm). Thus results of [8,20] apply, and we can solve the LP in polynomial time via the ellipsoid method. In practice such approaches tend to be far too slow, but analogous techniques based on the Simplex method (which forfeit the polynomial bound on worst-case running time) are feasible for instances of substantial size. The current best programs for solving the TSP optimally use the Simplex method together with separation routines for a variety of constraint classes, including the subtour constraints, and so it is in principle possible to compute the HK bound exactly simply by leaving certain routines out of such a program [16,28]. Basing such an approach on the optimization program of Applegate, Bixby, Chvátal, and Cook [1], we

* Room 2D-150, AT&T Bell Laboratories, Murray Hill, NJ 07974
† Department of Mathematics and Computer Science, Amherst College, Amherst, MA 01002.
‡ Silicon Graphics, Inc, Mountain View, CA 94043.

have computed exact HK bounds for instances with as many as 30,000 cities [16].

For instances where computation of the exact bound is not feasible or simply takes too long, Held and Karp proposed an iterative method for approximately computing the bound using Lagrangean relaxation. Certain versions of this method can be proved to converge to the true value if run for enough iterations, but this appears to require exponential time. Held and Karp (and subsequent researchers [9,12, 16,28,32,35,36]) thus proposed various schemes that used only a polynomial number of iterations to obtain what was hoped to be a good lower bound on the HK bound. These approximations typically can involve thousands of minimum spanning tree (MST) computations, but fortunately all but the last of these can be performed in an appropriately constructed sparse subgraph and consequently need only take time $O(N \log N)$ [14,16,19,28,36]. For geometric instances the final MST computation for the full graph can be computed using *k-d* trees [3] in which the dual variables of the iterated approach are embedded in an extra dimension, and so it too can be performed in subquadratic time [14,16]. As a result, the iterative approach is feasible for geometric instances with N as large as 10^6. Several versions of this approach typically get quite close to the true bound, at least for randomly generated instances [16].

In this paper, we consider three applications of our ability to compute/approximate the HK bound. In Section 2 we study the merits of using the HK bound as a stand-in for the optimal tour length when evaluating the quality of near-optimal tours. Our evaluation is based on extensive experimental comparisons between the two values for a wide variety of instance types. The test beds include (1) points uniform in the 2-dimensional unit square under Euclidean, rectilinear, and supremum norms and under both standard (planar) and toroidal topologies, (2) points uniform in the 3- and 4-dimensional unit cubes under the Euclidean metric and both standard and toroidal topologies, (3) random distance matrices, and (4) Release 1.2 of `TSPLIB`, Gerd Reinelt's database of ''real-world'' TSP instances, available via anonymous FTP from `softlib.cs.rice.edu`.

In Section 3, we empirically estimate the expected values of the HK bound as a function of N for the various random instance classes mentioned above, based on extensive computations using a near-exact approximation algorithm. For each geometric instance class, the ratio of the expected HK bound to $N^{(d-1)/d}$, where d is the dimension, approaches a constant [7], and we obtain good estimates both for this constant and for the rate of convergence to it. For random distance matrices, the expected HK bound itself approaches a constant, which we again estimate.

Finally, in Section 4 we put the results of Sections 2 and 3 together to estimate the expected optima for the various instance classes, in particular deriving tight bounds on the corresponding asymptotic constants. This requires extrapolations about the values of the expected HK gaps, since we typically were unable to compute them exactly for $N > 1,000$. Our data is fortunately such that plausible extrapolations can be made. Moreover, for the geometric classes the same constants hold whether one uses the planar or the toroidal topology, and we can exploit the fact that the latter provides a much faster rate of convergence.

Space limitations prevent us from including a detailed discussion of the statistical techniques we used in making our asymptotic estimates and deriving confidence intervals for them. The summary data on which the main estimates are based *is* included, however, and so readers are free to evaluate our claims on their own. Readers interested in more details about the algorithms used in our HK-bound computations are referred to the forthcoming paper [16], which will cover implementations and running times for the algorithms used here and a variety of alternative approaches. Finally, we should note that this paper has had a long gestation period, with less-precise versions of our main conclusions (based on less-complete data and covering fewer instance classes) having been presented orally as early as the 1990 ORSA/TIMS conference [14].

2. The Held-Karp Gap

A key application of the HK bound is to serve as a stand-in for the optimal tour length when evaluating the quality of near-optimal tours, an approach taken for example in [15,17,18]. For large instances, some such substitute is necessary. Although many TSP applications exist with as many as 30,000 or 100,000 cities, to date the largest nontrivial TSP instance solved to optimality involves only 7,397 cities, and the computation took several CPU-years using state-of-the-art computers [1]. Moreover, although many optimization codes are now capable of handling 1000-city instances with a high probability of success, running times for such instances are unpredictable and can vary from hours to weeks.

To cope with this difficulty, researchers have typically taken one of three approaches to the empirical evaluation of heuristics. First, one might simply restrict one's experiments to instances for which the optimal solution is already known, the most famous of which are all included in `TSPLIB`. This limits the studies to a small collection of instances, none of more than moderate size, but it has served as the basis for several interesting studies, such as [28]. Second, one might concentrate on randomly generated instances from some class for which estimates are known on the expected optimal solution value. The effectiveness of this approach depends on the quality of the estimates used, and unfortunately the commonly used estimates for the most popular instance class are highly inaccurate, as we shall see in the next section. Furthermore, even with accurate estimates, this approach may require very large sample sizes in order to provide high-confidence results when N is small. The third approach is to compare the solution found for a given instance to a *bound* on the optimal tour length for that instance. The effectiveness of this approach depends both on the quality of the bound and on the repeatability of its computation. The HK bound has a particular advantage with respect to the latter consideration, since it is a well-defined number for every instance, and as mentioned above can typically be computed exactly or with high accuracy in reasonable amounts of time [16]. Some papers have used ''the best known lower bound'' as their reference point. This can be a more accurate estimate, but it is not a well-defined number and is typically a moving target. Moreover, the HK bound by itself provides a reasonably close and consistent estimate of the optimal tour length. Worst-case results tell us that, assuming the triangle inequality, the HK bound is at least $2/3$ of the optimal tour length [31,37]. In practice, it appears to be much better, even in cases where the triangle inequality is violated. Let us consider the instance classes mentioned in the introduction, one by one.

Random Euclidean Instances. This class, in which cities correspond to points uniformly distributed in the unit square under the Euclidean metric, is by far the most-studied instance class for the TSP. Table 1 summarizes our measurements of the average gaps for such instances, which we express as the percent by which the optimal length exceeds the HK bound. Instance sizes range from 100 to 100,000 cities, going up by factors of $\sqrt{10}$, although the entries for $N > 1,000$ should be viewed as at best anecdotal. (These simply cover the instances of these sizes in the random Euclidean testbed of [17,18], and the entries in the *Min* and *Max* columns are simply the minimum and maximum gaps for the relevant instances.)

For $N \le 1,000$, the results are based on large enough samples that variances and confidence intervals for the mean gap could be estimated. Note that, as might be expected from the probabilistic results of [30,33] about optimal tour length, the standard deviations decline substantially as N increases. Thus we need fewer samples for larger values of N, partially offsetting the increase in the time needed to generate a single data point. The entries in the table are given with sufficient precision that rounding effects won't obscure the size of the confidence intervals. Based on the latter, however, one should only view the first two of the four included digits as having any statistical significance. For all the instances in these samples, precise values of the HK bound were used, and for all of the instances with $N \le 316$ the precise optimal value was used, as computed by the Applegate et al. [1] optimization code. For the 100 instances with $N = 1,000$, we approximated the optimal tour length by the best tour found in twenty 10,000-iteration runs of the Iterated Lin-Kernighan algorithm (ILK), implemented as described in [17,18]. To estimate the quality of these approximations, we ran the Applegate et al. optimization code on the first 36 of the 100 instances. The code identified optimal solutions for 33 of these within the time bounds we imposed and obtained reasonably tight lower bounds in the 3 remaining cases. The best ILK-produced tours were optimal for 32 of the 33 solved instances. Overall, they averaged .0025% above the best lower bounds produced by the optimization code. Thus any overestimate of the average HK gap resulting from our use of ILK should only affect the (not statistically significant) 3rd and 4th digits of the table entries for $N = 1,000$.

100 · (Opt - HK)/HK						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
100	2098	.7512	.4756	.7308	.7715	.0407
316	600	.7822	.2395	.7630	.8013	.0383
1,000	100	.7611	.1354	.7345	.7876	.0531
				Min	Max	
3,162	5	.745	—	.672	.800	—
10,000	4	.740	—	.692	.794	—
31,623	2	.731	—	.705	.757	—
100,000	2	.740	—	.732	.747	—

TABLE 1. Excess of optimal over the HK bound for 2-dimensional random Euclidean instances.

For the instances with $N > 1,000$, we again used multiple runs of ILK to obtain upper bounds on the optimal length, but we can no longer calibrate how good those bounds are, and we suspect that the figures we present may well over-estimate the true gaps, especially for the largest values of N in the table. Entries in Table 1 only go up to $N = 100,000$, as these are the largest instances for which long runs of ILK were feasible. A final technical note: because the codes we use assume that distances have integral values, our ''unit square'' was actually a 10^6 by 10^6 grid and distances were taken to be the ceiling of the standard Euclidean distance. (Rounding to the nearest integer can lead to violations of the triangle inequality.) Limited experiments with higher precision and other rounding schemes suggest that our choices in these matters had little effect on the results.

Note that the average gaps in Table 1 do not appear to vary greatly, lying between .73 and .79% for all values of N in the table. The values for $N \le 1,000$ are consistent with a fixed expected gap size of between .76 and .77%, although they suggest that the expected gap increases slightly from $N = 100$ to $N = 316$ and thereafter declines, a decline that may continue as N increases beyond 1,000, given the anecdotal evidence for $N > 1,000$. Another important point about the gaps is that they are much less variable than the corresponding tour lengths, both as a function of N and for fixed N, as illustrated in Figure 1. This figure combines two plots covering the instances in the random Euclidean testbed of [17,18]. The first plot gives the ratio to $\sqrt{N}$ of the optimal (or at least shortest known) tour lengths (the +'s in the figure). We normalize by $\sqrt{N}$ since Beardwood, Halton, and Hammersley [2] have shown that for random Euclidean instances, the expected ratio of the optimal tour length to $\sqrt{N}$ approaches a limiting constant C_{OPT} as $N \to \infty$. The

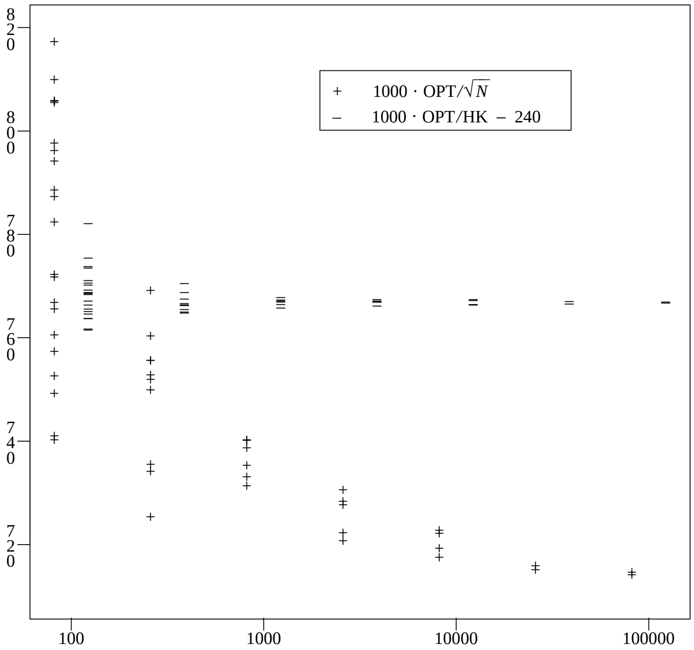

FIGURE 1. Two methods for normalizing tour length, with HK results shifted downward to fit the same frame.

second plot gives the same tour lengths when divided by the corresponding HK bounds (the –'s), shifted by a constant amount so that they fit in the same frame. As is apparent, the rate of convergence to C_{OPT} is slow, and so the average ratio of optimal tour length to $\sqrt{N}$ changes substantially as N increases, whereas the ratio to the HK bound is much more stable. The latter ratios also have a much smaller range for each N, even though their denominators are 25-30% smaller (as we shall see in the next section). The obvious implication is that good performance estimates can be obtained with far fewer data points (and values of N) if one compares to the HK bounds for the instances in question rather than to the expected optimal.

Random Euclidean Instances, Toroidal Topology. Table 2 again presents results for instances consisting of points uniformly distributed in the unit square under the Euclidean metric, but now we compute distances under the *toroidal* rather than the planar topology, i.e., under the assumption that the opposite sides of the unit square are identified. Once again the entries for $N \le 316$ are based on exact computations of the optimal tour length, whereas those for $N = 1{,}000$ are based on the best of twenty 10,000-iteration runs of ILK. For these instances, ILK is not as effective as it was under the planar topology, where the possibilities for tours were constrained by the presence of a boundary. We ran the Applegate et al. optimization code on 49 of the 824 instances in the sample. The code identified optimal solutions for all 49, but ILK only found the optimal on 39% of these, compared to 89% for the planar case. More importantly, ILK's average overestimate was .0128%, some 5 times larger than it was in the planar case. This error is too large to be ignored, so we have adjusted the mean in Table 2 downward to reflect it. The entry for the standard deviation is simply that for the ILK results, but the confidence interval is based on the assumption that the errors in the ILK/HK and ILK/OPT ratios compound rather than cancel. (For these two ratios, the 95% confidence intervals on the percentage errors are [.6184,.6323] and [.0084,.0173], respectively.)

100 · (Opt - HK)/HK						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
100	13957	.5542	.3301	.5488	.5597	.0109
316	1154	.6048	.1839	.5942	.6154	.0212
1,000	824	.6124	.1014	.6010	.6238	.0228

TABLE 2. Percentage gap for toroidal 2-dimensional random Euclidean instances.

Note that the average HK gaps are significantly smaller than those for the planar topology, although the difference appears to be shrinking as N increases. Asymptotically, the difference should go to 0. If we let C_{HK} be the limiting value of the ratio of the expected HK bound to $\sqrt{N}$, proved to exist by Goemans and Bertsimas in [7], then the limiting percentage gap is $100 * (C_{OPT} / C_{HK} - 1)$ under both the planar and toroidal topologies, since the values of C_{OPT} and C_{HK} are the themselves the same under both topologies [6,13]. In Section 3 we shall have more to say about where (between .61% and .78%) this limiting gap might lie.

Two-Dimensional Instances under other Metrics. Table 3 presents results for points uniform in the unit square

100 · (Opt - HK)/HK						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
Supnorm Metric				2d Unit Square		
100	600	.7355	.4667	.6981	.7728	.0847
316	200	.7408	.2341	.7084	.7733	.0649
Supnorm Metric				2d Unit Torus		
100	600	.5158	.3300	.4894	.5422	.0528
316	200	.5649	.1757	.5406	.5893	.0487
Rectilinear Metric				2d Unit Square		
100	600	.7442	.4537	.7079	.7805	.0726
316	200	.7460	.2260	.7146	.7773	.0627
Rectilinear Metric				2d Unit Torus		
100	600	.4962	.3037	.4719	.5205	.0486
316	200	.5524	.1635	.5297	.5750	.0453

TABLE 3. Average gaps for other 2-d metrics.

100 · (Opt - HK)/HK						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
Random Euclidean Instances				3d Unit Cube		
100	632	.2942	.1943	.2790	.3093	.0303
316	200	.2760	.0976	.2625	.2895	.0270
Random Euclidean Instances				3d Unit Torus		
100	600	.1833	.1249	.1733	.1933	.0200
316	200	.1844	.0656	.1754	.1935	.0181
Random Euclidean Instances				4d Unit Cube		
100	600	.1773	.1235	.1674	.1872	.0198
316	200	.1421	.0537	.1347	.1496	.0149
Random Euclidean Instances				4d Unit Torus		
100	600	.0977	.0766	.0915	.1038	.0123
316	200	.0873	.0358	.0823	.0922	.0099

TABLE 4. Average gaps for other 3- and 4-d instances.

and torus under the supnorm (L^{∞}) and the rectilinear norm (L^1), restricted to $N \le 316$ because of the expense of computing results for $N \ge 1,000$. Although the tour lengths and HK bounds under these metrics are substantially different from those in the Euclidean case, note that the percentage gaps are reasonably close to those in the corresponding Euclidean cases (both planar and toroidal).

Higher-Dimensional Euclidean Instances. A different story arises when we go to higher dimensions. Table 4 presents results for the Euclidean metric for points uniform in the 3- and 4-dimensional unit cubes and tori. (In k dimensions, the toroidal topology identifies all k pairs of opposing faces of the cube.) Here the average gaps are much smaller than in the 2-dimensional case, with the size of the gap declining further as the number of dimensions increases. Once again, the toroidal topologies yield smaller gaps, although now these are not so obviously increasing with N. The gaps for the standard topologies seem to be decreasing with N, however, suggesting that the asymptotic expected gaps are closer to the toroidal values. (Note that because we were less interested in the precise values of the expected gaps reported in Tables 3 and 4, we used fewer sample points than in the Euclidean case, and we hence have wider confidence intervals.)

Random Distance Matrices. Table 5 gives results for our final main class of randomly generated instances, ones in which each inter-city distance is chosen uniformly from the interval (0,1], independently of all other distances. Note that instances generated this way typically do not satisfy the triangle inequality. This class is the second most popular for testing TSP heuristics. The expected optimal is bounded, independent of N, so no normalization based on N is needed. Moreover, this sort of instance tends to be easier for optimization codes to handle, and so we can include exact sample points for instances with N as large as 3,162. As with our geometric instances, we were forced by our codes to use integer edge lengths, and so the actual edge lengths were integers chosen uniformly from the range $\{1,\ldots,10^6\}$, rather than samples from (0,1]. This limited precision does not appear to have had a substantial impact on our major conclusion, however, which is that the average gap declines rapidly with increasing N. It may even be approaching 0 as a limit, although to provide evidence for that we *would* need higher precision data, as well as results for larger N. This is contrast to the behavior for geometric instances, where the average gaps change slowly with N and appear to have asymptotic limits significantly larger than 0.

Instances from TSPLIB. The behavior of the HK gap is a bit less consistent for real-world instances, but not overly so. Figure 2 displays the HK gaps for all `TSPLIB` instances for which the optimum tour length is currently known. With the exception of two instances, all the gaps are less than 1.76%, with the average being roughly .8%. The major outlier is the 225-city instance `ts225`, in which cities are points on a regular 15 by 15 grid. This instance was specifically designed to be hard for optimization codes

100 · (Opt - HK)/HK						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
Random Distance Matrices						
100	1261	.1237	.1408	.1160	.1315	.0165
316	200	.0468	.0490	.0400	.0536	.0136
1,000	30	.0083	.0096	.0049	.0118	.0069
3,162	15	.0036	.0036	.0017	.0054	.0037

TABLE 5. Average gaps for random distance matrices.

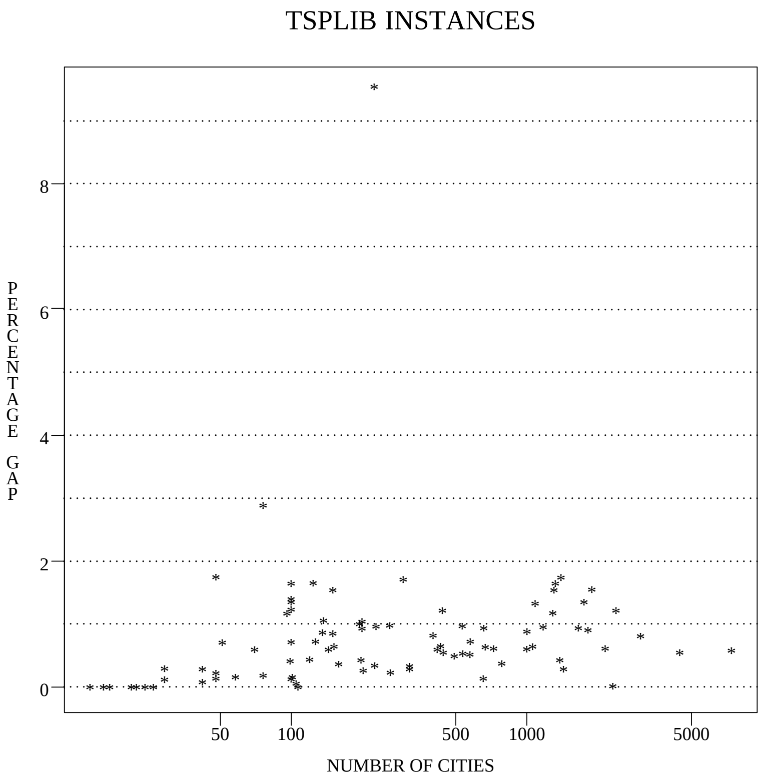

FIGURE 2. Percent excess of optimal tour length over Held-Karp bound for TSPLIB instances.

and we now see that it is anomalous in other ways as well. The lesser outlier lies among the instances with fewer than 100 cities, where we might expect special behavior; note that the gaps for the seven smallest instances are all 0. For the ten `TSPLIB` instances whose optima are not yet known, ranging in size from 1,577 to 85,900 cities, the maximum gap is at most 1.66% based on the best current upper bounds available to us for these instances. The journal version of this paper will list the HK bounds for all the instances in `TSPLIB`, with only the value for `pla85900` being an approximation. (In studying instances from `TSPLIB`, we are partially duplicating work by Reinelt, who in [28] reported the exact HK bound for a selection of 24 `TSPLIB` instances from 198 to 5,934 cities. His bounds disagree with ours on two instances, `fl1577` and `d1655`, where our bounds are 21886 and 61550, respectively. The first difference is due to a typographical error in his Table 10.16 [28]. In the second case his bound is .009% lower than ours, and the problem is likely due to rounding differences, given that we used double precision for all of our computations and some of his involved single precision.)

We have also measured the gaps for many of the 3-dimensional X-ray crystallography instances of Bland and Shallcross [4], which typically seem to be less than 0.5%. Another class of structured instances that we have studied in a limited way consists of randomly generated ‘‘clustered’’ instances, where normally distributed clusters of 50 cities are scattered uniformly over the unit square. The instance called `dsj1000` in `TSPLIB` is an example. For a collection of 5 samples each with N = 100, 1,000, and 10,000, the gaps ranged from .16 to .90%, with the biggest gaps tending to occur for the largest instances. (Our reported gaps for these may be overestimates, however, given that they are based on the best tours we were able to find, not necessarily the optimal tours.) We should point out that for these clustered instances and for the instances in `TSPLIB`, the iterative techniques for approximating the Held-Karp bound are not nearly as good or consistent as they are for randomly generated instances. For example, although they often provide estimates that are within .01% of the true bounds, in the case of instance `fl3795` they missed the true bound by over 3% (thus overestimating the gap by a factor of four) [16]. For such instances our ability to compute the exact bounds may be more critical, and when these are not available, we recommend for the sake of future comparisons that authors report the precise values of whatever reference bounds they use.

3. Estimating the Expected HK Bound

As mentioned earlier, not all past studies of TSP heuristics have followed our advice and presented their results in terms of average excess over the HK bound. More typically, papers simply report the average tour lengths for random Euclidean instances of specified sizes and compare

Random Euclidean Instances (Planar)						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
(Approx.HK Bound)/$\sqrt{N}$						
100	102058	.77027	.02240	.77013	.77040	.00027
316	51203	.74088	.01172	.74077	.74098	.00021
1,000	23752	.72578	.00631	.72570	.72586	.00016
3,162	15936	.71769	.00349	.71764	.71775	.00011
10,000	4000	.71334	.00195	.71328	.71340	.00012
31,622	1362	.71101	.00109	.71095	.71106	.00011
100,000	236	.70970	.00060	.70962	.70977	.00015
316,228	39	.70892	.00035	.70881	.70903	.00022
(HK Bound – Approx.HK Bound)/$\sqrt{N}$						
100	2098	.00010	.00031	.00009	.00011	.00002
316	600	.00015	.00022	.00013	.00017	.00004
1,000	100	.00005	.00002	.00004	.00005	.00001
3,162	54	.00004	.00002	.00003	.00004	.00001
10,000	14	.00004	.00001	.00003	.00004	.00001

TABLE 6. Normalized HK-bounds and errors, planar case.

them to estimates of the expected optima. In order to evaluate such results in the context of studies that do follow our advice, it is thus necessary to have good estimates for the expected HK bound as a function of N.

Let us first consider random Euclidean instances. Let $C_{HK,p}(N)$ denote the expected value of the HK bound divided by $\sqrt{N}$ for N-city random Euclidean instances under the planar topology. Table 6 summarizes the results of HK bound computations for a large collections of random Euclidean instances with N increasing from 100 to 316,228 by factors of roughly $\sqrt{10}$. These were computed using a variant of the Held-Karp iterative approach suggested by Helbig-Hansen and Krarup [9]. The implementation incorporated the speed-up tricks mentioned in Section 1, to be described more fully in [16]. For $N \le 10{,}000$, Table 6 also presents values for the expected error of this approximation technique, based on comparisons to the true bound. The larger average error values for the smaller values of N are for the most part due to occasional large outliers. Note that the average normalized error seems to have settled down to about .00004 once $N \ge 1{,}000$, a figure we can assume will continue to hold even for $N > 10{,}000$. On the other hand, the convergence of $C_{HK,p}(N)$ seems still unfinished when $N = 316{,}228$. For the one million-city instance in the testbed of [17,18], the normalized HK bound is .7086, and even this does not appear to be the ultimate value.

One way to estimate the ultimate value is by a weighted least squares curve fit to the adjusted means of Table 6 (i.e., the sums for each N of the average approximate bound and the expected approximation shortfall). Using standard techniques, we obtain the function

$$C_{HK,p}(N) \sim .70805 + \frac{.52229}{N^{.5}} + \frac{1.31572}{N} - \frac{3.07474}{N^{1.5}}$$

The form of this function corresponds to the product of an $(a + b/N^{.5})$ factor times a power series in $1/N$, with the lower order terms dropped. The $b/\sqrt{N}$ term in the function reflects the impact of the boundary of the unit square, which causes some $\Theta(\sqrt{N})$ cities to have fewer near neighbors than is typical. The need for a power series factor is suggested by the analysis of Percus and Martin in [26]. Certainly all these terms are needed if we are to obtain a good fit. This curve as given never strays more than .00004 away from the adjusted means of Table 6 and lies well within the 95% confidence intervals for all values of N. If the $d/N^{1.5}$ term is omitted, a consistent pattern of errors is introduced, and if both it and the c/N term are omitted, the best fits we could find wandered far outside the confidence interval for at least one data point.

On the assumption that we have chosen the correct functional form, and given the standard deviations and numbers of samples reported in Table 6, the above least squares fit yields a 95% confidence interval for C_{HK} of .70805±.00007. Although one might be skeptical about estimating C_{HK} by this approach, the formula does provide a convenient way of summarizing our data and enables us to derive estimates for $C_{HK,p}(N)$ for values of $N \ge 100$ not in the table. (The formula is not accurate for N much less than 100; at the very least a c/N^2 term would need to be included in the function if it were to cover this range.)

An easier way to estimate C_{HK} is by studying random Euclidean instances under the toroidal topology, as defined in the previous section. As noted there, the value of C_{HK} is not affected by this change in topology. Moreover, we eliminate the ‘‘boundary effect’’ of the planar topology, which appears to be responsible for the slow convergence to C_{HK} under that topology. This is confirmed in Table 7, where we summarize our toroidal topology data for random Euclidean instances with N ranging from 100 to 31,662. Note that $C_{HK,t}(N)$ appears to have converged to at least its

Random Euclidean Instances (Toroidal)						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
(Approx.HK Bound)/$\sqrt{N}$						
100	98246	.70894	.01910	.70882	.70906	.00024
316	43736	.70818	.01083	.70808	.70828	.00020
1,000	25000	.70802	.00603	.70794	.70809	.00015
3,162	14317	.70790	.00340	.70785	.70796	.00011
10,000	6039	.70794	.00191	.70789	.70799	.00010
31,622	783	.70791	.00108	.70784	.70799	.00015
(HK Bound – Approx.HK Bound)/$\sqrt{N}$						
100	2208	.00003	.00008	.00003	.00003	.00001
316	200	.00004	.00006	.00003	.00004	.00002
1,000	100	.00004	.00003	.00004	.00005	.00001
3,162	29	.00004	.00002	.00003	.00004	.00001

TABLE 7. Normalized HK-bounds and errors, toroidal case.

first four digits by the time $N = 3,162$. In fitting the adjusted means for this data, the $c/\sqrt{N}$ term can indeed be safely omitted. One can stay within .00004 of all the adjusted means from Table 7 by simply using the function

$$C_{HK,t}(N) = .70794 + \frac{.10081}{N}$$

The resulting 95% confidence interval for C_{HK} is .70794±.00003, a confidence interval that is disjoint from our previous one, although the means differ only by .0001. Given the relative convergence rates, the toroidal estimate seems the more credible, and so we believe it safe to conjecture that $C_{HK} \sim .70795 \pm .00005$, with the shift in the mean and the expansion of the confidence interval representing a mild compromise with the planar data. (Independently of these computations, Valenzuela and Jones [36] estimated a significantly lower value, but their data was only for the planar topology and $N \le 20,000$, and they seem to have had substantially fewer data points per N.)

Because of space limitations, we shall have to omit the analogous data for the rectilinear and supnorm metrics, and for 3- and 4-dimensional Euclidean instances, although our estimates for the associated asymptotic constants are listed in the next section (Table 8), as are our results for random distance matrices (Table 9). Before concluding this section, however, two issues affecting the validity of the data in Tables 6 and 7 should be addressed. First, we note that the averages appear not to depend significantly on the random number generator used. Most of our data was generated using a variant on the ‘‘shift register’’ random number generator described in [21]. However, for the $N = 100,000$ case in the planar topology and the $N = 1,000$ case in the toroidal topology we also generated data using the widely distributed linear congruential generator `drand48`. In the first case, the means of 100+ samples for each generator were .70964 and .70975, respectively, and each mean lay within the 95% confidence interval of the other. The data for $N = 100,000$ in Table 6 is based on the combination of these two samples. In the the second case, the means for 5000 samples were .70801 and .70814, respectively, and again each lay within the 95% confidence interval for the other. As a further check, we also generated 5000 samples using the sum (mod 1) of the two generators, which one would expect to be at least as random as the better of the two generators, and obtained another reasonably compatible value (.70794), whose confidence interval had a substantial intersection with the other two.

For the $N = 1,000$ planar case, we also evaluated a second factor that might have affected our averages: our use of bounded precision coordinates and rounding. Recall from Section 2 that because our codes require integer edge lengths we typically choose city coordinates as integers in the range from 1 to 10^6 and round distances to integral values. Dividing by 10^6 then yields an approximation to the standard unit square model. For the computations in this section, we rounded to the nearest integer, rather than taking the ceiling, as in the previous section. The difference between the two rounding schemes does not seem to affect the gap between the HK bound and the optimal tour length or the quality of our approximate HK computations. It could, however, affect our estimates of the expected HK bounds. Taking the ceiling would introduce a consistent upward bias, whereas rounding to the nearest integer should allow us to simulate higher precision, given the randomness of the coordinates. For example, when $N = 100,000$ the average length of an edge in a good tour is slightly over 2,200. Thus the potential error due to rounding the length of a typical edge is at most .023%, and one would expect the cumulative error to be much less. We tested this hypothesis by performing 5000 trials with $N = 1,000$ in which the coordinates were chosen between 1 and 10^5. This led to an average edge length of 2,200, compared to the normal average of 22,000. The mean result was .70802, once again consistent with all the other obtained values, and coincidentally precisely the same value as obtained when we merged all our samples. Thus it appears that the limited precision of our coordinates does not introduce a significant bias in our results. We expect this to remain true even for the case of $N = 316,228$ in Table 6, which corresponds to an average edge length of about 1200.

4. Estimating the Expected Optimum

The most straightforward way to estimate the expected optimal tour length for a given instance class and number of cities N would be to randomly sample instances of this type and compute the optimal tour length for each. Unfortunately, even for $N = 100$ current optimization codes do not seem to be fast enough to generate the 100,000 or so data points that would be necessary if we wished to get tight confidence bounds on our estimates, and for larger N, the situation is much worse. As we saw in Section 2, however, we can generate enough data points to get a tight estimate on the expected gap between the HK bound and the optimal tour length, because the variance for the latter quantity is so much lower. This suggests the ‘‘variance reduction’’ trick of combining such an estimate with one for the expected HK bound (much easier to obtain because HK bounds can be computed more quickly) to derive an estimate of the expected optimal. In Sections 2 and 3 we provided the data for doing just this in the case of random Euclidean instances under the planar and toroidal metrics for $N \in \{100, 316, 1000\}$. In this section we will undertake the apparently more difficult task of estimating the asymptotic constants for the optimal tour lengths.

Let us begin with the two-dimensional random Euclidean case. Beardwood et al. [2] provided the first estimate of the limiting constant C_{OPT}. Using tours that they constructed by hand for a 202- and a 400-city instance, they estimated C_{OPT} to be roughly $.53 \cdot \sqrt{2} \sim .75$ (rounded with unwarranted precision to .749 in many subsequent citations). In 1977, based on computer experiments with more sophisticated heuristics, Stein [34] estimated the constant to be 0.765, a result that has been widely cited. In recent years, various researchers besides ourselves have realized that both these figures are overestimates. In 1989, Ong and Huang [25] observed that a version of 3-Opt yielded normalized tour lengths converging to 0.74, which therefore would be an upper bound on C_{OPT}. In 1994, Fiechter [5] reported tours from a ''parallel tabu search'' algorithm whose normalized lengths approached .73, and Lee and Choi [24] estimated that their ''multicanonical annealing'' algorithm yields normalized tour lengths converging to 0.721. Mathematical analysis is not yet precise enough to provide much insight, although Krauth and Mézard [22], using a somewhat suspect statistical mechanical argument, estimated the constant to be .7257. It turns out that even the lowest of these estimates is still over 1% too high.

There are several approaches one might take to estimate C_{OPT} from our data. The first would be to extrapolate a value for the asymptotic expected HK gap from the estimates our data provides for $N \le 1,000$. Unfortunately, although we know that the gaps under the planar and toroidal metrics must approach the same limit, it is hard to see what that limit might be from the data in Tables 1 and 2. The planar estimates are all above .75%, perhaps trending downward, while the toroidal estimates are under .62%, definitely trending upwards. How can we tell where these trends might meet? To answer this question let us look in more detail at what is happening with the gap under the toroidal topology. The observed increase has two possible sources: either the expected HK bound is decreasing or the expected optimal is increasing. The data on which Table 2 is based, together with additional experiments covering $N = 32$, imply that the latter is not occurring. Thus the increase in expected gap must be due to a decrease in the expected HK bound. But the data in Table 7 plus our estimate of C_{HK} implies that $C_{HK,t}(1,000)/C_{HK} < 1.00017$. Thus the asymptotic expected gap should be at most .017 percentage points larger than the value for $N = 1,000$, for an ultimate value of about .63%. Using this value and our earlier estimate of C_{HK}, we conclude that $C_{OPT} \sim .7124$. Alternatively, we can use our data to directly estimate $C_{OPT,t}(N)$ for $N \in \{100,316,1000\}$, and attempt to curve fit the data. If one does this, one discovers that $C_{OPT,t}(N)$ appears essentially to have converged by $N = 1,000$, with the limiting value again being about .7124. Taking a conservative approach that simply adds up the lengths of all the confidence intervals involved, we arrive at a final estimate that $C_{OPT} \sim .7124 \pm .0002$.

This should be compared to the independent estimate of .7120 ± .0004 by Percus and Martin in the forthcoming paper [26]. Percus and Martin base their estimate on optimal tour length computations under the toroidal topology for $N \in \{12,13,14,15,16,17,30,100\}$. One might worry that such an extrapolation could be corrupted by factors that are only significant for very small N, but the relatively close agreement between the two estimates suggests that such fears are for the most part unjustified.

Instance Class	C_{HK}		C_{OPT}	
2d Unit Square, Euclidean	.70795	± .00005	.7124	± .0002
2d Unit Square, Recilinear	.8891	± .0003	.8943	± .0007
2d Unit Square, Supnorm	.6286	± .0002	.6323	± .0005
3d Unit Square, Euclidean	.6968	± .0002	.6980	± .0003
4d Unit Square, Euclidean	.7228	± .0002	.7234	± .0003

TABLE 8. Estimated constants for geometric instance classes.

We have derived analogous estimates for the constants C_{HK} and C_{OPT} in two dimensions under the rectilinear and supnorm metrics, and in 3 and 4 dimensions under the Euclidean metric. The details must be postponed to the journal version of this paper, although the values are summarized in Table 8. (Note that the normalizing factor varies with the dimension d, being equal to $N^{(d-1)/d}$.) We have not attempted to obtain estimates with as high precision for these latter cases, since they come up much less frequently (if at all) in the TSP literature. Note that the values for the rectilinear metric should be $\sqrt{2}$ times those for the supnorm metric, and they roughly are. Percus and Martin [26] estimate that for the 3-dimensional Euclidean case, $C_{OPT} = .6978 \pm .0004$, an estimate that is consistent with ours. The results of [2] imply that C_{OPT} should ultimately be increasing with the number of dimensions, so it is interesting to note that it actually decreases in going from $d = 2$ to $d = 3$ before increasing as we go to $d = 4$.

We conclude with a summary (in Table 9) of our estimates for random distance matrices. (The average approximation error was .0002 for all values of N.) Although the data we provide is simply for the HK bound, recall from Table 5 that the expected gap for this class appears to be going to 0, so the estimate of C_{HK} derived from this table should also be a good estimate for C_{OPT}. Krauth and Mézard [22], using a statistical mechanical argument that is much less suspect than the one cited above for the Euclidean case, estimate that $C_{OPT} \sim 2.0415...$ for random distance matrices. Our data is clearly consistent with this conjecture and suggests that convergence is essentially complete by $N = 1,000$.

Approximate HK Bound, Random Distance Matrices						
No. of Cities	No. of Samples	Mean	Std. Dev.	95% Conf. Interval Lower	Upper	Length
100	122290	2.0339	.1882	2.0328	2.0349	.0021
316	41157	2.0393	.1083	2.0382	2.0403	.0021
1,000	13461	2.0416	.0610	2.0406	2.0426	.0020
3,162	3529	2.0407	.0336	2.0396	2.0418	.0022
10,000	1061	2.0420	.0188	2.0409	2.0431	.0022

TABLE 9. Estimates of the constant for random distance matrices.

Acknowledgment

The authors thank David Applegate and Bill Cook for providing us access to the optimization code of [1] and for modifying the code so that it would compute exact HK bounds.

References

1. D. APPLEGATE, R. BIXBY, V. CHVÁTAL, AND W. COOK, ''Finding cuts in the TSP (A preliminary report),'' Report No. 95-05, Center for Discrete Mathematics and Theoretical Computer Science (DIMACS), Rutgers University, Piscataway, NJ, 1995.
2. J. BEARDWOOD, J. H. HALTON, AND J. M. HAMMERSLEY, ''The shortest path through many points,'' *Proc. Cambridge Philos. Soc.* **55** (1959), 299-327.
3. J. L. BENTLEY, ''Multidimensional binary search trees used for associative search,'' *J. Assoc. Comput. Mach.* **18** (1975), 309-517.
4. R. G. BLAND AND D. F. SHALLCROSS, ''Large traveling salesman problems arising from experiments in X-ray crystallography: A preliminary report on computation,'' *Operations Res. Lett.* **8** (1989), 125-128.
5. C.-N. FIECHTER, ''A parallel tabu search algorithm for large traveling salesman problems,'' *Disc. Applied Math.* **51** (1994), 243-267.
6. M. X. GOEMANS, private communication (1995).
7. M. X. GOEMANS AND D. BERTSIMAS, ''Probabilistic analysis of the Held-Karp lower bound for the Euclidean traveling salesman problem,'' *Math. Oper. Res.* **16** (1991), 72-89.
8. M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *Geometric Algorithms and Combinatorial Optimization*, Springer, Berlin, 1988.
9. K. H. HELBIG-HANSEN AND J. KRARUP, ''Improvements of the Held-Karp algorithm for the symmetric traveling salesman problem,'' *Math. Programming* **7** (1974), 87-96.
10. M. HELD AND R. M. KARP, ''The traveling-salesman problem and minimum spanning trees,'' *Operations Res.* **18** (1970), 1138-1162.
11. M. HELD AND R. M. KARP, ''The traveling-salesman problem and minimum spanning trees: Part II,'' *Math. Programming* **1** (1971), 6-25.
12. M. HELD, P. WOLFE, AND H. P. CROWDER, ''Validation of subgradient optimization,'' *Math. Programming* **6** (1974), 62-88.
13. P. JAILLET, ''Cube versus torus models and the Euclidean minimum spanning tree constant,'' *Ann. Appl. Probab.* **3** (1992), 582-592.
14. D. S. JOHNSON, ''New computational experiments on upper and lower bounds for the traveling salesman problem,'' talk given at the ORSA/TIMS Joint National Meeting (Philadelphia, October 29, 1990).
15. D. S. JOHNSON, ''Local optimization and the traveling salesman problem,'' in *Proc. 17th Colloq. on Automata, Languages, and Programming*, Lecture Notes in Computer Science **443**, Springer-Verlag, Berlin, 1990, 446-461.
16. D. S. JOHNSON, D. L. APPLEGATE, L. A. MCGEOCH, AND E. E. ROTHBERG, ''On fast methods for computing the Held-Karp traveling salesman bound,'' to appear.
17. D. S. JOHNSON, J. L. BENTLEY, L. A. MCGEOCH, AND E. E. ROTHBERG, ''Near-optimal solutions to very large traveling salesman problems,'' in preparation.
18. D. S. JOHNSON AND L. A. MCGEOCH, ''The traveling salesman problem: A case study in local optimization,'' to appear in *Local Search in Combinatorial Optimization*, E. H. L. Aarts and J. K. Lenstra (eds.), John Wiley & Sons, New York, 1995.
19. D. S. JOHNSON AND E. E. ROTHBERG, implemented program (September 1985).
20. R. M. KARP AND C. H. PAPADIMITRIOU, ''On linear characterizations of combinatorial optimization problems,'' *SIAM J. Comput.* **11** (1982), 620-632.
21. D. E. KNUTH, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, 2nd Edition, Addison-Wesley, Reading, Mass., 1981, 171-172.
22. W. KRAUTH AND M. MÉZARD, ''The cavity method and the travelling-salesman problem,'' *Europhys. Lett.* **8** (1989), 213-218.
23. E. L. LAWLER, J. K. LENSTRA, A. H. G. RINNOOY KAN, AND D. B. SHMOYS, *The Traveling Salesman Problem*, John Wiley & Sons, Chichester, 1985.
24. J. LEE AND M. Y. CHOI, ''Optimization by multicanonical annealing and the traveling salesman problem,'' *Phys. Rev. E* **50** (1994), R651-R654.
25. H. L. ONG AND H. C. HUANG, ''Asymptotic expected performance of some TSP heuristics: An experimental evaluation,'' *Eur. J. Oper. Res* **43** (1989), 231-238.
26. A. G. PERCUS AND O. C. MARTIN, ''Finite size and dimensional dependence in the Euclidean traveling salesman problem,'' *Phys. Rev. Lett.*, to appear.
27. G. REINELT, ''TSPLIB—A traveling salesman problem library,'' *ORSA J. Comput.* **3** (1991), 376-384.
28. G. REINELT, *The Traveling Salesman Problem: Computational Solutions for TSP Applications*, Lecture Notes in Computer Science 840, Springer-Verlag, Berlin, 1994, 172-186.
29. G. REINELT, personal communication.
30. W. T. RHEE AND M. TALAGRAND, ''A sharp deviation inequality for the stochastic traveling salesman problem,'' *Annals of Probability* **17** (1988), 1-8.
31. D. B. SHMOYS AND D. P. WILLIAMSON, ''Analyzing the Held-Karp TSP bound: A monotonicity property with applications,'' *Inform. Process. Lett.* **35** (1990), 281-285.
32. T. H. C. SMITH AND G. L. THOMPSON, ''A LIFO implicit enumeration search algorithm for the symmetric traveling salesman problem using Held & Karp's 1-tree relaxation,'' *Ann. Disc. Math.* **1** (1977), 479-493.
33. J. M. STEELE, ''Complete convergence of short paths and Karp's algorithm for the TSP,'' *Math. Oper. Res.* **6** (1989), 374-378.
34. D. STEIN, *Scheduling Dial-a-Ride Transportation Systems: An Asymptotic Approach*, Ph.D. Dissertation, Harvard University, Cambridge, MA, 1977.
35. T. VOLGENANT AND R. JONKER, ''A branch and bound algorithm for the symmetric traveling salesman problem based on the 1-tree relaxation,'' *Eur. J. Oper. Res* **9** (1982), 83-89.
36. C. L. VALENZUELA AND A. J. JONES, ''Estimating the Held-Karp lower bound for the geometric TSP,'' Manuscript dated 8 June 1995.
37. L. WOLSEY, ''Heuristic analysis, linear programming, and branch and bound,'' *Math. Prog. Study* **13** (1980), 121-134.