MATHEMATICS OF OPERATIONS RESEARCH
Vol. 15, No. 4, November 1990
Printed in U.S.A.

PROBABILISTIC AND WORST CASE ANALYSES OF CLASSICAL PROBLEMS OF COMBINATORIAL OPTIMIZATION IN EUCLIDEAN SPACE*†

J. MICHAEL STEELE

The classical problems reviewed are the traveling salesman problem, minimal spanning tree, minimal matching, greedy matching, minimal triangulation, and others. Each optimization problem is considered for finite sets of points in $\mathbb{R}^d$, and the feature of principal interest is the value of the associated objective function. Special attention is given to the asymptotic behavior of this value under probabilistic assumptions, but both probabilistic and worst case analyses are surveyed.

1. Introduction. The origin and foundation of the results reviewed here is the 1959 paper by Beardwood, Halton, and Hammersley, "The Shortest Path through Many Points." The importance of that work was made evident in 1976 with the appearance of R. Karp's paper, "The Probabilistic Analysis of Some Combinatorial Search Algorithms." Using the Beardwood-Halton-Hammersley theorem, Karp showed that, in a probabilistic sense, the computationally difficult traveling salesman problem could be solved efficiently. Karp's work illuminated the power inherent in understanding the asymptotic behavior of combinatorial problems under probabilistic models, and, as a consequence, several research initiatives were set in motion. In particular, motivation presented itself for the investigation of the following questions:

(a) Which aspects of Karp's algorithmic paradigm call for further refinement of the associated probability theory?

(b) Which problems of combinatorial optimization have an asymptotic theory like that revealed by Beardwood, Halton, and Hammersley for the traveling salesman problem (TSP)?

(c) Are there features which are more refined than those revealed by the Beardwood, Halton, and Hammersley (BHH) theorem?

(d) Can the proof of the basic BHH theorem be usefully simplified?

The main goal of this review is to survey the work which addresses these questions while keeping a close eye on the techniques which seem likely to lead to further progress. In the course of the review, some new proofs and new results are given, but they are pursued only as far as they serve to illustrate unifying methods.

This review evolves as follows. §2 introduces and motivates the Beardwood, Halton, and Hammersley theorem, then §3 gives a proof of the most basic form of the BHH theorem and a review of alternative approaches. The fourth section surveys the

*Received January 7, 1988; revised June 26, 1989.
AMS 1980 subject classification. Primary: 68C25. Secondary: 60D05, 60C05, 60K30.
IAOR 1973 subject classification. Main: Combinatorial Analysis. Cross references: Programming: Network
OR/MS Index 1978 subject classification. Primary: 011 Analysis of algorithms. Secondary: 490 Networks/graphs/traveling salesman.
Key words. Probabilistic algorithms, Combinatorial optimization, Traveling salesman problem, Minimal Spanning Tree, Subadditive Euclidean functionals, Minimal matching, Minimal triangulation, K-median problem.
†Supported in part by National Science Foundation Grant #DMS-8414069 and #DMS-8812868.

0364-765X/90/1504/0749/$01.25

refinements which can be made in the basic BHH theorem, and the fifth section reviews a generalization of the BHH with enough power to bring the Steiner tree problem, the TSP, and their rectilinear versions under the same umbrella.

The next four sections focus on problems which have some of the structure of the TSP functional but which fail to be covered by the general results of §5. The problems treated are the minimum spanning tree problem with power weighted edges, the directed traveling salesman problem, the K-median problem, the problem of minimal triangulations, and the remarkable two sample matching theorems of Ajtai, Komlós, and Tusnády.

§10 takes a turn in perspective and analyzes a nonoptimal process, the greedy matching algorithm. The tools of subadditivity still apply, but in the absence of optimality, subadditivity requires more digging. The whole subject of probabilistic analysis is then put on hold in §11, and the TSP, MST, and matching problems are examined for their worst case performanced. One intriguing aspect of those results is their formal similarity to the probabilistic theorems, despite the absence of any probability model or probabilistic reasoning.

§12 stretches the review's title to embrace the fact that many interesting problems in combinatorial optimization are not Euclidean. There has been some extraordinary recent work on linear programs with random cost, and this work casts new light on minimal spanning trees and the assignment problem which is too good to miss.

Of the new results embedded in this review, it may be useful to note that the proof of the Beardwood, Halton, Hammersley theorem given in §3 seems to be substantially more direct and more elementary than earlier proofs. Also, there is a technical sounding result that has a good story behind it: the sum T_n of the dth power of the edges in any minimal spanning tree of $\{x_1, x_2, \dots, x_n\} \subset [0, 1]^d$ is bounded uniformly in n. This bound generalizes and simplifies the result which was obtained for $d = 2$ by Gilbert and Pollak (1968), and it provides a base from which one might expect progress on a probabilistic conjecture of R. Bland discussed in §6. The rest of the story comes from the use of smooth spacefilling curves and of their potential applications in other problems.

2. The BHH theorem. The results considered here always concern finite subsets $V = \{x_1, x_2, \dots, x_n\}$ of points in $\mathbb{R}^d$. These points are viewed as vertices of a graph, any pair of distinct elements of V will be called an edge, and if $e = \{x_i, x_j\}$ then $|e| = |x_i - x_j|$ will denote the usual Euclidean length of the line from x_i to x_j. For the most part, we will be concerned with special classes of graphs such as the class of all *tours* which can be described (too succinctly) as the set of connected graphs such that each vertex has degree two. The BHH theorem explains the probabilistic behavior of the functional

$$L(x_1, x_2, \dots, x_n) = \min_T \sum_{e \in T} |e|$$

where the minimum is over all tours T with vertex set V. In its most general form one can state the Beardwood, Halton, Hammersley theorem as follows:

THEOREM 1. *If X_i are i.i.d. random variables with compact support, then with probability one*

$$\lim_{n \to \infty} L(X_1, X_2, \dots, X_n)/n^{(d-1)/d} = c_d \int_{\mathbb{R}^d} f(x)^{(d-1)/d}\, dx \tag{2.1}$$

where $c_d > 0$ is a constant not depending on the distribution of the X_i and where f is the density of the absolutely continuous part of the distribution of the X_i.

We should collect some observations which illustrate the meaning of Theorem 1. First, because the X_i have compact support, the right-hand side of (2.1) is always finite. Also, in the case that f is the indicator of a bounded set A, the integral in (2.1) reduces to $m(A)^{1/d}$, where $m(A)$ is the Lebesgue measure of A. Finally, part of the content of (2.1) is that if the X_i have compact support which is singular, then $L(X_1, X_2, \ldots, X_n)$ is almost surely $o(n^{(d-1)/d})$.

There was a great deal of work on the TSP function L which preceded the BHH theorem. For example, Verblunsky (1951) showed

$$L(x_1, x_2, \ldots, x_n) \leqslant (2.8n)^{1/2} + 3.15 \tag{2.2}$$

for any $\{x_1, x_2, \ldots, x_n\} \subset [0,1]^2$, and Few (1955) showed

$$L(x_1, x_2, \ldots, x_n) \leqslant d\{2(d-1)\}^{(1-d)/2d} n^{(d-1)/d} + 0(n^{1-2/d})$$

for any $\{x_1, x_2, \ldots, x_n\} \subset [0,1]^d$. There were also earlier results by Ghosh (1949), Mahalanobis (1940), Jessen (1942), and Marks (1948). The most precise recent result is found in Karloff (1989). More will be said about these results when worst case asymptotics are considered in §11.

In order to appreciate the real thrust of the BHH theorem—that it is an exact asymptotic result—one should not overlook how easy it is to guess that the appropriate order growth rate of L_n is $n^{(d-1)/d}$. For example, the $n^{(d-1)/d}$ lower bound can be guessed by imagining the X_i to be periodically spaced on the rectangular lattice, and, although such imaginings are purely heuristic, their heart is in the right place. For an upper bound on L_n of order $n^{(d-1)/d}$, one can even proceed rigorously and apply the following elementary lemma:

LEMMA 1. *There is a constant $c = c(d) > 0$ such that for n points $\{x_1, x_2, \ldots, x_n\} \subset [0,1]^d$, there is a pair of points x_i and x_j such that*

$$|x_i - x_j| \leqslant cn^{-1/d}. \tag{2.3}$$

No harm is done in omitting the easy proof of Lemma 1 since a substantial generalization will be proved later. Still, the humble bound (2.3) serves us well in many arguments involving sets of points in $[0,1]^d$, and many of the results surveyed here depend on Lemma 1 in one way or another.

3. Looking for the essence of the BHH.

What makes the BHH theorem possible? From Lemma 1 and a naive lower bound based on the expected value of nearest neighbor distances, one is easily lead to conclude that with probability one,

$$0 < \alpha = \liminf L_n/n^{(d-1)/d} \leqslant \limsup L_n/n^{(d-1)/d} = \beta < \infty. \tag{3.1}$$

Here, one should note that Hewitt-Savage Zero-One Law can be used to guarantee that α and β are indeed constants. The central insight of Beardwood, Halton, Hammersley (1959) was that tools of subadditivity and self-similarity could be brought to bear on L_n with sufficient force to boost elementary bounds like (3.1) to the level of a genuine limit theorem.

There are at least three ways to continue the development of those basic insights. One can refine the original methods in order to give simpler or more intuitive proofs of the BHH theorem. One can examine the ways in which the BHH theorem can be sharpened. Or, one can look to abstractions and generalizations of the BHH theorem which might prove fruitful.

To make life easy and to make the method crystal clear, consider X_i which are i.i.d. uniform on $[0,1]^2$. This case is already nontrivial, and it is surely the most often applied. Moreover, the tools for extending from this special case to the case of general distribution with compact support are somewhat removed from the basic geometry of the TSP.

Our first step will be to determine the asymptotics of the expectations $\phi(n) = EL(X_1, X_2, \dots, X_n)$ and, in particular, to show

$$\phi(n) \sim c\sqrt{n}. \tag{3.2}$$

We first divide the unit square $[0,1]^2$ into m^2 subsquares of side m^{-1}, and we let $Z_i = \sum_{k=1}^{n} 1_{Q_i}(X_k)$ denote the number of the X_k, $1 \leqslant k \leqslant n$, which occupy the ith subsquare Q_i.

Since Q_i has side length m^{-1}, we see by scaling that the expected length of the shortest tour through k points in Q_i is equal to $m^{-1}\phi(k)$. Also, by sewing the m^2 subcubes together by passing through them row by row, we can tie any set of subtours of the m^2 subsquares of Q_i into a grand tour with a total incremental cost bounded by $3m$. In terms of expectations this says

$$\begin{aligned} \phi(n) &\leqslant \sum_{i=1}^{m^2} \sum_{k=0}^{n} m^{-1}\phi(k) P(Z_i = k) + 3m \\ &\leqslant m \sum_{k=0}^{n} \phi(k) \binom{n}{k} m^{-2k} (1 - m^{-2})^{n-k} + 3m. \end{aligned} \tag{3.3}$$

Before taking limits in (3.3) we should verify that ϕ is decently smooth and does not grow too rapidly. In particular, using Lemma 1 we find constants c_1 and c_2 such that

$$\phi(k) \leqslant c_1 k^{1/2} \quad \text{and} \quad \phi(k+m) \leqslant \phi(k) + c_2 m k^{-1/2}.$$

If we now let n and m go to infinity in such a way that (a) $m^{-2}n$ converges to λ and (b) $\phi(n)/n^{1/2}$ tends to its limit superior, the preceding bounds and elementary analysis give us that for each $\lambda > 0$ we have

$$\limsup \phi(n)/n^{1/2} \leqslant \lambda^{-1/2} E\phi(W_\lambda) + 3\lambda^{-1/2}, \tag{3.4}$$

where W_λ is a Poisson random variable with mean λ. We now choose a subsequence n_k such that

$$\lim_{k\to\infty} \phi(n_k)/n_k^{1/2} = \liminf_{n\to\infty} \phi(n)/n^{1/2}.$$

The shooting is almost over. We just let $\lambda = n_k$ in (3.4) and let k go to infinity. Since the probability mass of W_λ is highly concentrated in the interval $\lambda \pm \varepsilon\lambda$ and since ϕ is well behaved we obtain

$$\lim_{k\to\infty} n^{-1/2} E\phi(W_{n_k}) + 3n_k^{-1/2} = \liminf_{n\to\infty} \phi(n)/n^{1/2}.$$

In view of the main inequality (3.4), we have proved that the limit of $\phi(n)/n^{1/2}$ exists. To give more details would spoil the fun, so we count the first step (3.2) as done.

The second step is to show that with probability one,

$$\lim_{n \to \infty} (L_n - EL_n)/n^{1/2} = 0, \tag{3.5}$$

and we begin with some recent observations of Rhee and Talagrand (1987).

For each $1 \leqslant i \leqslant n$, we let A_i denote the sigma field generated by X_j, $1 \leqslant j \leqslant i$. Also, we let T^i denote the length of the shortest tour through the n points $\{X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n+1}\}$. Two nice facts about the T^i is that $E(T^i|A_i) = E(T^i|A_{i-1})$, and, of course, $E(T^i|A_{i-1}) = E(L_n|A_{i-1})$. So if we let $\Delta_i = L_n - T^i$, we see

$$\begin{aligned} d_i &= E(L_n|A_i) - E(L_n|A_{i-1}) \\ &= E(L_n - T^i|A_i) - E(L_n - T^i|A_{i-1}) = E(\Delta_i|A_i). \end{aligned} \tag{3.6}$$

Now, $L_n - EL_n = \sum_{i=1}^n d_i$ and the d_i are orthogonal random variables (in fact martingale differences), so

$$\operatorname{Var} L_n = E\left(\sum_{i=1}^n d_i\right)^2 = \sum_{i=1}^n E(d_i^2) \leqslant \sum_{i=1}^n E\Delta_i^2 = nE\Delta_1^2 \tag{3.7}$$

where, in the last two steps, we first applied Jensen's inequality and then used the fact that L_n is symmetric when viewed as a function of the X_i, $1 \leqslant i \leqslant n$.

We have the simple bound $\Delta_1 \leqslant 2\min_{1 < j \leqslant n}|X_1 - X_j|$, so by elementary calculus, there is a c_0 such that $E\Delta_1^2 < c_0 n^{-1}$. Finally, going back to inequality (3.7) we obtain the rather surprising result that $\operatorname{Var} L_n$ is bounded by c_0 for all n, a fact which was first obtained in Steele (1981b) by means of the Efron-Stein inequality.

From the uniform boundedness of $\operatorname{Var} L_n$ and Chebyshev's inequality, we have

$$P(|L_n - EL_n| \geqslant \epsilon n^{1/2}) \leqslant c_0 \epsilon^{-2} n^{-1},$$

and if we let n_k be the greatest integer not exceeding $k^{5/4}$, then the Borel-Cantelli Lemma further tells us $L_{n_k} \sim EL_{n_k}$ with probability one. Also, by applying Lemma 1 repeatedly, we can check that for $n_k \leqslant n \leqslant n_{k+1}$ we have

$$L_{n_k} \leqslant L_n \leqslant L_{n_k} + c(n_{k+1} - n_k)n_k^{-1/2}, \tag{3.8}$$

and for our choice of n_k, inequality (3.8) shows $\max_{n_k \leqslant n \leqslant n_{k+1}}|L_n - L_{n_k}| \to 0$ as $k \to \infty$. Moreover, if we first take expectations in (3.8), we have $\max_{n_k \leqslant n \leqslant n_{k+1}}|EL_n - EL_{n_k}| \to 0$ as $k \to \infty$. Together with the almost sure asymptotic relation $L_{n_k} \sim EL_{n_k}$, these observations complete the proof of our second step.

This proof differs from earlier proofs of the Beardwood, Halton, Hammersley theorem in several respects, but in particular there is no introduction of a Poisson process to produce independence of the tours within the Q_i. Even the simplified proof of the BHH theorem given in Karp and Steele (1985) relied on the traditional Poissonization which has been carried along since 1959.

The Poisson embedding technique has some conceptual advantages over a bare-knuckles proof, and probably the main virtue of the present method is that it is honestly easy enough to be given in a beginning graduate course.

One particular benefit of avoiding Poisson embedding and the conditioning argument is that it circumvents the need for an explicit Tauberian theorem. In Hochbaum and Steele (1982) and Karp and Steele (1985), a Tauberian theorem for Borel averages was used to back out of the smoothing introduced by the Poissonization. Such a path provides for some economy of thought, but the covert Tauberian step used here is more elementary and direct. Nevertheless, one is forced to own up to the fact that there was no explicit Tauberian step in the original BHH proof.

This consideration of Poissonization and Tauberian arguments for Borel averages should not leave the impression that such techniques are automatically to be avoided. As Bingham (1981) shows, there is a powerful armory of results which can be used for backing out of Borel averages. Also, Hartmann (1987) provides Tauberian theorems which are designed specifically for problems like the probabilistic analysis of the TSP.

The martingale argument introduced by Rhee and Talagrand and applied here in a simplified form will probably become the standard approach to the variances and tail behaviors of random variables like L_n. The application of the inequality of Efron and Stein (1981) by Steele (1981b) provided a simplification and strengthening of earlier methods, but, with all of the machinery of martingales in play, the Efron-Stein inequality is not likely to continue to be the tool of choice. Still, one should note that the martingale method and the Efron-Stein method are closely related.

To give a brief idea of this relationship we first recall a variant of the Efron-Stein inequality from Steele (1986b):

For any function $Y = Y(X_1, X_2, \ldots, X_n)$ of n independent identically distributed random variables X_i one has

$$\operatorname{Var} Y \leqslant {}^1\!/_2 E \sum_{j=1}^{n} (Y - Y^{(j)})^2$$

where $Y^{(j)}$ is identical to Y except X_i has been redrawn, i.e. $Y^{(j)} = Y(X_1, X_2, \ldots, X_{j-1}, \tilde{X}_j, X_{j+1}, \ldots, X_n)$ where the $2n$ random variables $X_i, \tilde{X}_i$, $1 \leqslant i \leqslant n$, are i.i.d. It should be evident that for $L_n = Y$ the combinatorial facts one needs to use to provide bounds on $Y - Y^{(j)}$ are closely related to those used to bound the martingale differences d_i of (3.6). As final evidence of the closeness of the two approaches to $\operatorname{Var} L_n$, we should note that, in fact, Rhee and Talagrand (1986) have shown that the Efron-Stein inequality can be proved by martingale methods. Other proofs and generalizations of the Efron-Stein inequality are provided by Karlin and Rinott (1982), Vitale (1984), and Steele (1986b).

Even at this early point, it may be useful to record a basic open problem. Since $\operatorname{Var} L_n$ is bounded for $d = 2$, it seems inevitable that one has a genuine limit,

$$\lim_{n \to \infty} \operatorname{Var} L_n = \alpha > 0. \tag{3.9}$$

So far, the best we know in this direction is the result due to Rhee (1988) that

$$\liminf_{n \to \infty} \operatorname{Var} L_n > 0. \tag{3.10}$$

A less certain conjecture than (3.9), but one that is still very likely, is that a central limit theorem holds:

$$L_n - EL_n \sim N(0, \alpha). \tag{3.11}$$

As it happens, there is a close relationship between (3.9) and (3.11) through the central limit theorem for martingales, and if (3.9) can be established, (3.11) is liable to fall to the same insights. All one needs to add flesh to this speculation is to consider L_n as a sum of martingale differences as in (3.6) and to think through the underpinnings of the martingale central limit theorem as organized by Rootzén (1983).

4. Refinements of the BHH. Weide (1978) observed that a number of subtleties lie behind the probabilistic models which motivate methods like Karp's partitioning algorithm for the traveling salesman problem. For example, one can think of problems as being *incrementing* where $\{X_1, X_2, \dots, X_n\}$ grows to $\{X_1, X_2, \dots, X_n, X_{n+1}\}$, or, *independent* where a problem of size n given by $S = \{X_1, X_2, \dots, X_n\}$ is replaced by a problem of size $n+1$ given by $S' = \{X_1', X_2', \dots, X_{n+1}'\}$ where S' is independent of S. For models of the second type, one needs a notion of convergence which is stronger than almost sure convergence. Fortunately, the classical notion of *complete* convergence seems well suited to the task.

In Steele (1981b) it was proved that for independent random variables X_i with the uniform distribution on $[0,1]^2$ one has for any $\epsilon > 0$ that

$$\sum_{n=1}^{\infty} P\big(|L_n/n^{1/2} - \beta| > \epsilon\big) < \infty. \tag{4.1}$$

In classical terms, (4.1) says that $L_n/n^{1/2}$ *converges completely* to β.

The proof of (4.1) was achieved by combining the inequality of Efron and Stein (1981) with a recursion argument pulled along by Hölder's inequality. For $d = 2$, it was also proved that for all $k \geqslant 0$ there is a constant c_k such that

$$E(L_n - EL_n)^k \leqslant c_k \tag{4.2}$$

for all $n \geqslant 1$.

These moment inequalities imply strong bounds on the tail probabilities $P(|L_n - EL_n| \geqslant t)$, but still sharper bounds were obtained by Kern (1986), who combined moment generating functions with the Efron-Stein inequality. The most general and powerful approach to the tail probability of L_n is the martingale method introduced in Rhee and Talagrand (1987). For the TSP with $d = 2$, Kern showed there are constants $\beta > 0$ and $\gamma > 0$ such that

$$P\big(|L_n - EL_n| \geqslant t\big) \leqslant \beta \exp\big(-\gamma t/n^{1/4}\big), \tag{4.3}$$

while Rhee and Talagrand (1987) showed there is a constant $\gamma > 0$ such that for $n \geqslant 2$,

$$P\big(|L_n - EL_n| \geqslant t\big) \leqslant 2\exp\big(-\gamma t^2/\log n\big) \quad \text{and} \tag{4.4}$$

$$P\big(|L_n - EL_n| \geqslant t\big) \leqslant 2\exp\big(-\gamma t\big). \tag{4.5}$$

By combining the last two inequalities with tools from the interpolation theory of linear operators, Rhee and Talagrand (1989a) further showed there is a $\delta > 0$ such that

$$P\big(|L_n - EL_n| \geqslant t\big) \leqslant 2\exp\big(-\delta t^2/\log(1+t)\big). \tag{4.6}$$

The inequalities (4.3), (4.4), (4.5), and (4.6) are all much stronger than one needs to establish the complete convergence (4.1), and they also provide basic improvement on the information available from (4.2).

The final leg in this development was the proof of the full large deviation result of Gaussian type; there is a $K > 0$ such that for all $t \geqslant 0$ we have that

$$P(|L_n - EL_n| \geqslant t) \leqslant K \exp(-t^2/K). \tag{4.7}$$

This result was established in Rhee and Talagrand (1989b) where the martingale methods were coupled with delicate combinatorial arguments. By considering the interesting "dual" variable

$$Y_s = \max\{k\colon L(X_{i_1}, X_{i_2}, \ldots, X_{i_k}) \leqslant s, \text{ for all } k\text{-subsets of } X_1, X_2, \ldots, X_n\},$$

Shamir (1989) has provided an alternative approach to the tail behavior of the TSP, but so far the approach falls short of (4.7). The details are a bit too long to trace here, but the key insight is that Azuma's inequality applies more robustly to Y_s than to L_n because changing one of the X_i always changes Y_s by either 1 or 0.

5. Generalization of the BHH. The first proof of a generalization of the BHH theorem was given in Steele (1981a), but an important intermediate step was taken in Papadimitriou (1978) where pains were taken to articulate the properties which were central to the original proof of the BHH theorem.

How can one specify a general class of variables for which BHH type results will apply? How might one prove a limit result using only abstract properties of the TSP functional?

We will consider functions L which are defined for all the finite subsets of $\mathbb{R}^d$. The properties we require of L are reasonably few, and one can easily check that they hold for the TSP. In particular, we make the following assumptions:

A1. $L(\alpha x_1, \alpha x_2, \ldots, \alpha x_n) = \alpha L(x_1, x_2, \ldots, x_n)$ for all real $\alpha > 0$.

A2. $L(x_1 + x, x_2 + x, \ldots, x_n + x) = L(x_1, x_2, \ldots, x_n)$ for all $x \in \mathbb{R}^d$.

Since L is a function on the finite subsets of $\mathbb{R}^d$, we also note that $L(x_1, x_2, \ldots, x_n)$ is the same as $L(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$ for any permutation $\sigma\colon [1, n] \to [1, n]$. The function L is also assumed to be *monotone*, i.e.,

A3. $L(x \cup A) \geqslant L(A)$ for any $x \in \mathbb{R}^d$ and finite subset A of $\mathbb{R}^d$.

If ϕ denotes the empty set, we always suppose $L(\phi) = 0$ and note that the monotonicity of L entails positivity, i.e., $L(A) \geqslant 0$ for all finite sets $A \subset \mathbb{R}^d$.

Some boundedness of L is required, and a simple (but excessive) choice is provided by an assumption of *finite variance*,

A4. $\mathrm{Var}(L(X_1, X_2, \ldots, X_n)) < \infty$ whenever X_i, $1 \leqslant i \leqslant n$, are independent and uniformly distributed in $[0, 1]^d$.

The preceding assumptions are almost trivial to verify for a large number of problems, and one cannot prove very much using the bare assumptions A1–A4. What is needed is a more serious restriction powerful enough to get to the essence of the functionals like the shortest tour length.

We take $\{Q_i\colon 1 \leqslant i \leqslant m^d\}$ to be a partition of the d-cube $[0, 1]^d$ into m^d similar cubes with edges parallel to the axis and let tQ_i denote the dilated set defined by $\{x\colon x = ty,\ y \in Q_i\}$. The *subadditivity hypothesis*, which is the key to our main result, can be given as follows:

A5. There exists a $C > 0$, such that for all positive integers m and positive reals t, one has

$$L(\{x_1, x_2, \ldots, x_n\} \cap [0, t]^d) \leqslant \sum_{i=1}^{m^d} L(\{x_1, x_2, \ldots, x_n\} \cap tQ_i) + Ctm^{d-1}.$$

Functionals satisfying A1–A5 are fortunate indeed, since they satisfy the following generalization of the theorem of Beardwood, Halton and Hammersley.

THEOREM 2. *Suppose L is a monotone, Euclidean functional on* $\mathbb{R}^d$ *with finite variance which satisfies the subadditivity hypothesis. If* $\{X_i: 1 \leqslant i < \infty\}$ *are independent and uniformly distributed in* $[0,1]^d$, *then there is a constant* $\beta(L)$ *such that*

$$\lim_{n\to\infty} L(X_1, X_2, \dots, X_n)/n^{(d-1)/d} = \beta(L)$$

with probability one.

The proof of this result is too involved to repeat here. In spirit it rests upon a subsequence and subadditivity argument used earlier in different contexts by Kesten (1973) and Richardson (1973), but of course it also draws on many of the insights of the original proof given by Beardwood, Halton, and Hammersley (1959).

To extend Theorem 2 to random variables which are not uniformly distributed, one must make additional assumptions. We will call *L scale bounded* provided

A6. There is a constant B such that

$$L(x_1, x_2, \dots, x_n)/tn^{(d-1)/d} \leqslant B \quad \text{for all } n \geqslant t,\ t \geqslant 1, \quad \text{and}$$

$$\{x_1, x_2, \dots, x_n\} \subset [0,t]^d.$$

Also, we call *L simply subadditive* provided

A7. There is a constant B such that

$$L(A_1 \cup A_2) \leqslant L(A_1) + L(A_2) + tB$$

for any finite subsets A_1 and A_2 of $[0,t]^d$.

The last assumption we need is *upper linearity*. A Euclidean function L is called *upper-linear* provided

A8. For any finite collection of cubes Q_i, $1 \leqslant i \leqslant s$ with edges parallel to the axes and for any infinite sequence x_i, $1 \leqslant i < \infty$, in $\mathbb{R}^d$ one has

$$\sum_{i=1}^{s} L(\{x_1, x_2, \dots, x_n\} \cap Q_i) \leqslant L\left(\{x_1, x_2, \dots, x_n\} \cap \bigcup_{i=1}^{s} Q_i\right) + o(n^{(d-1)/d}).$$

The set of supplemental conditions is not as tidy as our first set, but eight is enough, and the conditions just laid out suffice to provide a full-fledged generalization of the Beardwood, Halton and Hammersley Theorem.

THEOREM 3. *Suppose L is a Euclidean functional which satisfies assumptions* A1–A8. *There is a constant* $\beta(L)$ *such that*

$$\lim_{n\to\infty} L(X_1, X_2, \dots, X_n)/n^{(d-1)/d} = \beta(L)\int_{\mathbb{R}^d} f(x)^{(d-1)/d}\,dx \quad a.s.$$

for any independent identically distributed random variables $\{X_i\}$ *with bounded support in* $\mathbb{R}^d$ *and absolutely continuous part* $f(x)\,dx$.

Among the functionals to which Theorem 3 can be applied directly are the TSP, the Steiner tree problem, and their rectilinear versions. Notable functionals to which Theorems 2 and 3 fail to apply include the minimal spanning tree (MST) and the minimal matching problem. Ironically, these problems fail to be covered by Theorem 2 just because they lack the simple monotonicity property A3. For these functionals and several others, the simple failure of A3 causes more trouble than one might have easily imagined.

A generalization of Theorem 3 which relaxes assumption A3 is also given in Steele (1981a), but the minimal matching problem still falls out of its range, despite the comments in Steele (1981a, pp. 372–373). The easiest way to establish the asymptotic behavior for the minimal matching problem is by direct arguments, and, in fact, many problems which fall just outside the class covered by Theorem 3 end up having enough individual character to justify specialized analysis. A feeling for the range of the techniques which have been built up to deal with such cases can be obtained by examining the results of the next three sections on minimal spanning trees, the directed TSP, the K-median problem, and minimal triangulations.

6. Minimum spanning trees and power edge weights. If we let M_n denote the length of the MST of a random sample,

$$M_n = \min_{T} \sum_{e \in T} |e| \tag{6.1}$$

where the minimum is over all spanning trees of $\{X_1, X_2, \dots, X_n\}$, one would certainly expect a result for M_n like the limit theorem obtained for L_n. Such a result does not come easily out of the general results summarized in Theorems 2 and 3, even though the classic paper of Beardwood, Halton, Hammersley (1959) already noted that the asymptotics of M_n should follow the same probability law as L_n.

Curiously, almost everything one knows about discrete optimization suggests that M_n should actually be easier to study than L_n. After all, Papadimitriou (1977) showed that the task of computing L_n is NP-complete, while the computation of M_n is easily achieved by a variety of well-known, fast, greedy algorithms. Still, because M_n lacks the basic monotonicity on which Theorem 2 depends, the probability theory of M_n is much more troublesome than that of L_n. Nevertheless, by relying on bare-handed understanding of M_n in cooperation with subadditive techniques, one can master the asymptotics of M_n and even more complicated functionals.

Since M_n should be so simple, there is natural inclination to expect results which are more extensive than those obtained for L_n. One natural step is to consider the more general functional

$$M_n^{(\alpha)} = \min_{T} \sum_{e \in T} \psi(|e|) \tag{6.2}$$

where ψ is a positive, monotone function.

Unlike the functionals considered previously, this functional fails to be homogeneous of order one, i.e. $M_n^{(\alpha)}$ fails to satisfy A1. Despite this and other differences, the asymptotic theory of this $M_n^{(\alpha)}$ fulfills our natural expectations, and in Steele (1988) the following result is established:

THEOREM 4. *Suppose X_i, $1 \leqslant i < \infty$, are independent random variables with distribution μ with compact support in $\mathbb{R}^d$, $d \geqslant 2$. If the monotone function ψ satisfies $\psi(x) \sim x^\alpha$ as $x \to 0$, for some $0 < \alpha < d$, then with probability one*

$$\lim_{n \to \infty} n^{-(d-\alpha)/d} M_n^{(\alpha)} = c(\alpha, d) \int_{\mathbb{R}^d} f(x)^{(d-\alpha)/d}\, dx. \tag{6.3}$$

Here, f denotes the density of the absolutely continuous part of μ, and $c(\alpha, d)$ denotes a strictly positive constant which depends only on the power α and the dimension d.

Part of the motivation for Theorem 4 and the more general study of power weighted edges comes from a conjecture of R. Bland, who was led by numerical experimentation to suspect that for X_i i.i.d $U[0,1]^d$, one has with probability one that

$$\lim_{n\to\infty} \min_T \sum_{e\in T} |e|^d = c < \infty. \tag{6.4}$$

Although this probabilistic conjecture emerged only recently, some earlier work was already waiting to speak in its behalf. In particular, Gilbert and Pollak (1968) showed by a delicate geometrical argument that in $d = 2$ the sum $\Sigma|e|^2$ is uniformly bounded for any minimal spanning tree and any set $\{x_1, x_2, \ldots, x_n\} \subset [0,1]^2$.

By applying some results on spacefilling curves, it turns out to be easy to prove the result of Gilbert and Pollak and even to extend it to arbitrary dimensions. The key observation is that which says there is a function f from $[0,1]$ *onto* $[0,1]^d$ which is Lipschitzian of order $1/d$, i.e.,

$$|f(x) - f(y)| \leqslant c|x-y|^{1/d} \tag{6.5}$$

for a constant $c > 0$. This bound holds in fact for several of the classical spacefilling curves, see Milne (1980).

To apply this to the bounding of edge weights of a minimal spanning tree, we suppose $V_n = \{x_1, x_2, \ldots, x_n\} \subset [0,1]^d$ and let T be a minimal spanning tree of V_n. Since f is surjective we can choose points $\{y_1, y_2, \ldots, y_n\}$ in $[0,1]$ such that $f(y_i) = x_i$. Next, we order the y_i to give $y_{(1)} \leqslant y_{(2)} \leqslant \cdots \leqslant y_{(n)}$ and define a new suboptimal spanning tree T_o on V_n by choosing the $n-1$ edges $(f(y_{(i)}), f(y_{(i+1)}))$, $1 \leqslant i < n$. By the optimality of T and the Lipschitz property of f, we find

$$\sum_{e\in T} |e|^d \leqslant \sum_{e\in T_o} |e|^d = \sum_{i=1}^{n-1} \left|f(y_{(i)}) - f(y_{(i+1)})\right|^d \tag{6.6}$$

$$\leqslant c^d \sum_{i=1}^{n-1} |y_{(i)} - y_{(i+1)}| \leqslant c^d.$$

By comparison with the complexity of the proof of Gilbert and Pollak (1968) in $d = 2$, this proof of the uniform boundedness of $\Sigma|e|^d$ is almost effortless. All of the geometry of $[0,1]^d$ has been compressed into the existence of the Lipschitz spacefilling curve. In the case of $d = 2$, the technique used above was first discovered by S. Kakutani. The broad usefulness of the spacefilling heuristic has been emphasized in Bartholdi and Platzman (1988). For a review of the oral tradition behind the spacefilling heuristic one should read the comments of Adler (1986). Also, for a discussion of power weighted edges in problems other than the MST, but within the range of Theorem 3, one should consult the recent thesis of Troyon (1988).

Although the main hope concerning the bound (6.6) is that it provides a sensible step toward proving Bland's conjecture, inequality (6.6) has applications in its own right. For example, by applying Hölder's inequality to (6.6) one finds that the sum of any k edges of an MST of $\{x_1, x_2, \ldots, x_n\}$ is bounded by $ck^{(d-1)/d}$, where c is the Lipschitz constant of (6.5). The delicacy of this fact can be appreciated by asking if the analogous fact holds for the TSP. The issue which makes the question subtle is that an optimal tour with respect to edge weights $|e|$ can be quite different from an optimal tour with edge weights $|e|^d$. In contrast, an important fact used in (6.6) is that if a tree T minimizes $\Sigma_{e \in T}|e|$ then T minimizes $\Sigma_{e \in T}\psi(|e|)$ where ψ is any monotone increasing function, and in particular, T minimizes $\Sigma_{e \in T}|e|^d$.

The geometry of Euclidean minimal spanning trees is considerably richer than that of the TSP, and there are several quantities—like the number of leaves—which have a much different character than the length related quantities which have been discussed so far. Still, it turns out that subadditivity and Efron-Stein variance bounding are again effective, and in Steele, Shepp, Eddy (1987) it is proved that with probability one the number of leaves of an MST of an i.i.d. sample with compact support in $\mathbb{R}^2$ (but otherwise general distribution) is asymptotic to cn with probability one. While the value of c is not known analytically, simulations suggest that $c \sim 0.22$ for random variables with the uniform distribution on $[0, 1]^2$.

7. The directed TSP. Karp (1977) posed the problem of formulating a probabilistic model of the *directed* traveling salesman problem (DTSP) which supports an asymptotically optimal probabilistic polynomial time algorithm. One such model given in Steele (1986a) also serves to illustrate that the techniques of subadditive Euclidean functionals are sufficiently robust to be able to deal with information which is exogenous to the location of the points.

As usual, we suppose X_i, $1 \leqslant i < \infty$, are independent random variables with the uniform distribution in the unit square $[0, 1]^2$, and we take $V_n = \{X_1, X_2, \dots, X_n\}$ as the vertex set for our directed graph G_n. The edges of the complete graph on V_n are given directions just by flipping coins. Formally, we take independent Bernoulli random variables Y_{ij}, $1 \leqslant i < j \leqslant n$, which are also independent of V_n and for which $P(Y_{ij} = 1) = 1/2 = P(Y_{ij} = 0)$. The directed edge set E_n is defined by taking $(X_i, X_j) \in E_n$ if $Y_{ij} = 1$ and $(X_j, X_i) \in E_n$ if $Y_{ij} = 0$.

The random variable of interest here is D_n, the length in the usual Euclidean distance of the shortest directed path through all of the vertices V_n of G_n.

It may not be apparent that there is always a directed path through V_n, but its existence follows from a classic result of Rédei (1934). An algorithmic proof of Rédei's theorem is given in passing in Steele (1986a), but the main results of that paper are the following:

THEOREM 5. *There is a constant* $0 < \beta < \infty$ *such that as* $n \to \infty$

$$ED_n \sim \beta\sqrt{n}\,. \tag{7.1}$$

THEOREM 6. *There is a polynomial time algorithm which provides a directed path through* V_n *with length* D_n^* *which satisfies*

$$ED_n^* \leqslant (1 + \epsilon) ED_n, \tag{7.2}$$

for all $\epsilon > 0$ *and* $n \geqslant N(\epsilon)$.

The directed traveling salesman problem has only been studied formally in the plane, but it is not hard to show by the same methods that the natural d-dimensional analog $ED_n \sim \beta_d n^{(d-1)/d}$ holds for any $d \geqslant 2$. A more serious issue concerns the possibility of stronger types of convergence. Because of the secondary randomization of the Y_{ij}'s used to determine the direction of the edges, the almost sure convergence theory for the directed TSP seems substantially trickier than that for the undirected problem. Still, the complete convergence problem for the directed TSP has recently been solved in Talagrand (1989).

In the course of juxtaposing the limit theory of D_n and its algorithmic theory, it is worth recalling that one does not always need the exact limit theory in order to provide a useful algorithmic theory. This fact is nicely developed in Halton and Terada (1982) where the probabilistic partitioning algorithm for the TSP in $\mathbb{R}^d$ is given in detail. Finally, one cautionary point is that there is not much in common between the *directed* TSP just reviewed and the *asymmetric* TSP considered in Karp (1979) and Karp and Steele (1985).

8. Problems of location and triangulation. The TSP and MST are natural candidates for the development of a refined limit theory because they are among the most studied problems in the area of Euclidean combinatorial optimization. But besides the TSP, MST, and their variants, there are many other functionals of interest and importance. This section discusses two such functionals, the first of which is the k-center problem.

Given an integer k and any set of n points $\{x_1, x_2, \ldots, x_n\}$ of $\mathbb{R}^2$, one defines the cost of the K-center location problem by

$$(8.1) \qquad M(k; x_1, x_2, \ldots, x_n) = \min_{S: |S|=k} \sum_{i=1}^{n} \min_{j \in S} |x_i - x_j|.$$

For $0 < p < \infty$ we can also define the more general functional

$$M_p(k; x_1, x_2, \ldots, x_n) = \min_{S: |S|=k} \sum_{i=1}^{n} \min_{j \in S} |x_i - x_j|^p.$$

Upper and lower bounds determining the order of magnitude of M were obtained in Fisher and Hochbaum (1980) and Papadimitriou (1981), and the following analog of the BHH theorem for M_p was obtained in Hochbaum and Steele (1982).

THEOREM 7. *If $\{X_i\}$ are independent and uniformly distributed on $[0, 1]^2$, then for any $1 \leqslant p < 2$ and $0 < \alpha < 1$ we have with probability one that*

$$\lim_{n \to \infty} M_p([\alpha n]; X_1, X_2, \ldots, X_n)/n^{1-p/2} = c_{\alpha, p}$$

for some constant $0 < c_{\alpha, p} < \infty$.

This result was obtained by using subadditivity and Tauberian arguments to obtain the asymptotics of EM_p and by using the Efron-Stein inequality to bound $\mathrm{Var}(M_p)$.

The same approach was applied successfully in Steele (1982) to prove a conjecture of György Turán on the rate of growth of the minimal triangulation of n points independently and uniformly distributed in the unit square.

By a triangulation of a finite set $S = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^2$, we mean a decomposition of the square $[0, 1]^2$ into triangles such that each $x_i \in S$ and each of the square's four corners is a vertex of some triangle. An important technical aspect of this definition is that we do not require that each vertex of the triangulation be an element of S. As one might expect, the length of triangulation is taken to be the sum of the lengths of all the edges in the triangulation, and the quantity central of interest is $T(x_1, x_2, \ldots, x_n)$, the minimum length over all possible triangulations.

The main limit result for triangulation takes on a familiar cast, but—as will be reviewed shortly—there are new features of the triangulation problem for which Theorem 8 fails to give a good hint.

THEOREM 8. *If* $T_n = T(X_1, X_2, \ldots, X_n)$ *where* X_i, $1 \leqslant i < \infty$ *are independent and uniformly distributed in* $[0,1]^2$, *then*

$$\lim_{n\to\infty} \frac{T_n}{\sqrt{n}} = \beta$$

with probability one for some constant $\beta > 0$.

To dispel any sense that Theorem 8 only presents more business as usual, one should ponder some of the problems that wait in the wings. Can one avoid the Steinerization which was introduced as the technical part of the definition? Also, what is the proper analog of this result in $[0,1]^d$? An interesting aspect of the second question is that the analog of edge lengths of the triangulations in $\mathbb{R}^2$ is surely the area of the faces of the simplicial decomposition in $\mathbb{R}^d$. This problem leads to the consideration of a discrete probabilistic Plateau problem. Such problems were first posed in Beardwood, Halton, and Hammersley (1959); but, despite the near 30-year lapse, there has been no progress in dimension $d \geqslant 3$.

9. Two sample matchings. Growth rates of order $n^{(d-1)/d}$ occur so frequently in problems of combinatorial optimization in Euclidean space, one eventually yearns for problems which exhibit different behavior.

Ajtai, Komlós, and Tusnády (1984) considered the natural problem of minimal Euclidean matchings of a pair of random samples for $[0,1]^2$, and found a striking new $\sqrt{n \log n}$ behavior. An amusing twist of their discovery is that for $d \geqslant 3$, the logarithmic term is no longer present and the habitual $n^{(d-1)/d}$ growth rate again prevails.

In the course of their work, Ajtai, Komlós, and Tusnády also developed two powerful qualitative features of random samples in $[0,1]^2$ which are intimately related to many optimization problems. But, before describing these features, we should lay out the results for which they were developed. Given X_i, $1 \leqslant i < \infty$, and Y_i, $1 \leqslant i < \infty$, two independent sequences of random variables with the uniform distribution on $[0,1]^2$, the random variable of central interest in Ajtai, Komlós, and Tusnády (1984) is defined by

$$T_n = \min_{\pi} \sum_{i=1}^{n} |X_{\pi(i)} - Y_i|. \tag{9.1}$$

In longhand, T_n is the cost of the minimal bipartite matching between the X's and the Y's. The main theorem of Ajtai, Komlós, and Tusnády (1984) sharpened an earlier upper bound of $0_p(n^{1/2} \log n)$ due to R. Karp (unpublished) and determined the exact order of T_n:

THEOREM 9. *There are constants* $0 < c_1 < c_2 < \infty$ *such that*

$$P\left(c_1\sqrt{n \log n} < T_n < c_2\sqrt{n \log n}\right) \to 1 \tag{9.2}$$

as $n \to \infty$.

There are two ingenious ideas which underlie the proof of (9.2). First, the proof of the lower bound depends upon the construction of a weight function f (which depends upon the X's and Y's) such that f overweights the X's, in the sense that

$$(9.3)\qquad P\left(\sum_{i=1}^{n} f(X_i) - \sum_{i=1}^{n} f(Y_i) > Lc_1\sqrt{n\log n}\right) \to 1.$$

The function f is also proved to be Lipschitzian,

$$(9.4)\qquad |f(x) - f(y)| \leqslant L|x - y|.$$

The facts (9.3) and (9.4) instantly justify the lower bound in (9.2), and there are doubtless other two-sample problems where the existence of such an f is a powerful tool. It also is interesting to see a Lipschitz function crop up again after the appearance of the deterministic Lip_α function which proved so useful in studying power weighted minimal spanning trees.

The second device developed to prove (9.2) concerns an elegant continuous analogue to the transportation problem. Risking a little colorful language, suppose you have a pile of dirt of weight $1/n$ located at each n random points $X_1, X_2, \dots, X_n$ in $[0,1]^2$. Further, suppose you want to spread that dirt out over a rectangular field $R_i \subset [0,1]^2$ of area exactly $1/n$, such that all of X_i's dirt is spread evenly over X_i's rectangle R_i. The total transportation cost of this redistribution can be measured by

$$(9.5)\qquad S_n^X = \sum_{i=1}^{n} \int_{R_i} |X_i - u|\,du.$$

Here, of course, the rectangles R_i are random pairwise disjoint sets that form a partition of $[0,1]^2$. We should note each R_i depends on all of the variables X_i, $1 \leqslant i \leqslant n$, and, also, there is no requirement that $X_i \in R_i$.

The main fact about the variables S_n^X is that the R_i can be constructed so that as $n \to \infty$ we have

$$(9.6)\qquad P\left(S_n^X < c\sqrt{n\log n}\right) \to 1.$$

Inequality (9.6) speaks directly to the bound sought in (9.2) since

$$(9.7)\qquad T_n \leqslant S_n^X + S_n^Y,$$

as one can see by confirming that shipping the X's dirt to $[0,1]^2$ and reversing the shipment of the Y's dirt to $[0,1]^2$ gives a real-valued solution to the transportation problem between X and Y. By a classical result of linear optimization, the minimal transportation solution must be a matching, and, therefore, the cost T_n of the minimal matching has to be less than $S_n^X + S_n^Y$.

A final aspect of the two-sample matching problem is that it has rich connections to several other problems of interest: bin packing, grid matching, and empirical discrepancies of lower layers. For these connections and many other interesting geometric insights one can rely on Shor (1985), Leighton and Shor (1986), and Coffman and Shor (1989).

10. Greedy matchings. One reason subadditive methods are effective in the probabilistic study of optimization problems is that optimality often provides an almost automatic path to the required subadditivity properties. Such circumstances no longer prevail when one studies asymptotically suboptimal heuristics, and in such cases considerable restructuring of the basic subadditive machinery may be required before limit results can be obtained. For example, the analysis of the greedy algorithm for minimal matching given in Avis, Davis, and Steele (1988) depended critically upon extracting subadditive inequalities directly from the combinatorial properties of the greedy matching process, even though the eventual result is just as one would expect:

THEOREM 10. *For each integer $d \geqslant 2$, there is a positive constant c_d such that if $X_1, X_2, \ldots,$ are i.i.d. random variables with values in R^d and bounded support, and if G_n denotes the Euclidean edge weight of the matching attained by the greedy algorithm applied to $\{X_1, X_2, \ldots, X_n\}$, then with probability one as $n \to \infty$,*

$$G_n \sim c_d n^{(d-1)/d} \int_{R^d} f(x)^{(d-1)/d}\, dx.$$

Here, f is the density with respect to d-dimensional Lebesgue measure of the absolutely continuous part of the distribution of the X_i.

One way to appreciate the analysis of the greedy matching problem is to consider heuristics for some closely related problems which still resist analysis. For example, for TSP heuristics based on nearest neighbors, minimal insertions, and many other processes (Karp et al. 1984, Lawler et al. 1985, and Rosenkrantz et al. 1977), one expects a result like Theorem 10. Such results are not yet available, and they are not likely to come easily. In the meanwhile their resistance helps illuminate the distinctive nature of those problems, like the greedy matching problem, where progress has been made.

11. Worst case growth rate results. The *worst case* behavior of a Euclidean functional is often remarkably parallel to its probabilistic behavior. This fact may come as a surprise, but the TSP and MST provide forceful illustrations.

To spell out the details, we again represent a tree or a tour by a graph $G = (V_n, E)$, where V_n denotes a set of n points in $[0, 1]^d$, E denotes a subset of the edges of the complete graph on the points of V_n, and $L(E)$ denote the sum of the lengths of the edges in E.

Conceptually, the worst case analyses of the TSP and MST are simpler than the probabilistic analysis, since the whole problem boils down to understanding the asymptotic behavior of the numerical sequences $\rho_{\mathrm{MST}}(n)$ and $\rho_{\mathrm{TSP}}(n)$ defined by

$$\rho_{\mathrm{MST}}(n) = \max_{V_n}\left\{\min_{T} \sum_{e \in T} |e|: T \text{ is a spanning tree of } V_n\right\} \quad \text{and} \tag{11.1}$$

$$\rho_{\mathrm{TSP}}(n) = \max_{V_n}\left\{\min_{T} \sum_{e \in T} |e|: T \text{ is a tour of } V_n\right\}. \tag{11.2}$$

The following result from Steele and Snyder (1989) recalls the form of the probabilistic results for the TSP and MST, but repetition of form should not obscure the substantial shift in perspective and technique due to the absence of probability theory.

THEOREM 11. *For any dimension $d \geqslant 2$, there are constants β_{MST} and β_{TSP} satisfying $1 \leqslant \beta_{\mathrm{MST}} \leqslant \beta_{\mathrm{TSP}}$ such that as $n \to \infty$,*

$$\rho_{\mathrm{MST}}(n) \sim \beta_{\mathrm{MST}} n^{(d-1)/d} \quad \textit{and} \tag{11.3}$$

$$\rho_{\mathrm{TSP}}(n) \sim \beta_{\mathrm{TSP}} n^{(d-1)/d}. \tag{11.4}$$

Exact asymptotic results for ρ_{MST} and ρ_{TSP} are recent, but considerable earlier work focused on bounds for ρ_{TSP} and ρ_{MST} which translate into bounds on β_{MST} and β_{TSP}. For example, Verblunsky (1951) proved that in $d = 2$ one has $\rho_{TSP}(n) \leqslant (2.8n)^{1/2} + 3.15$, and Fejes-Tóth (1940) established that $\rho_{TSP}(n)$ and $\rho_{MST}(n)$ are both at least as large as $(1 - \epsilon)(4/3)^{1/4}n^{1/2}$ for all $n \geqslant N(\epsilon)$. Next, Few (1955) improved the upper bound of Verblunsky (1951) to $\rho_{TSP}(n) \leqslant (2n)^{1/2} + 1.75$ in $d = 2$ and obtained $\rho_{TSP}(n) \leqslant d\{2(d-1)\}^{(1-d)/2d} n^{(d-1)/d} + 0(n^{1-2/d})$ for general $d \geqslant 2$. More recently, Moran (1984) obtained essential improvements on the upper bounds of Few (1955) for large values of d.

The simplest and sharpest bound on $\rho_{TSP}(n)$ in dimension two is given in Supowit, Reingold, and Plaisted (1983), where it is proved that the additive term 1.75 can be dropped from Few's upper bound to give $\rho_{TSP}(n) \leqslant (2n)^{1/2}$. That paper also gives another new proof of the appealing lower bound due to Fejes-Tóth (1940) that $\rho_{TSP}(n) \geqslant 2(12)^{-1/4}n^{1/2}$, and while the abstract of Supowit, Reingold, and Plaisted (1983) asserts $\rho_{TSP}(n) = \alpha\sqrt{n} + o(\sqrt{n})$, that statement accidentally sacrifices precision for brevity. The text of Supowit, Reingold, and Plaisted (1983) is quite clear and no claim is made concerning results like Theorem 11.

The idea which underlies the exact asymptotic analysis of $\rho_{MST}(n)$ and $\rho_{TSP}(n)$ is that both sequences satisfy inequalities which bound their rates of growth and which express an approximate recursiveness. The following lemma looks technical, but it gets at the essence of the asymptotics $\rho_{TSP}(n)$ and $\rho_{MST}(n)$.

LEMMA 2. *If* $\rho(1) = 0$ *and there is a constant* $c \geqslant 0$ *such that*

$$\text{(11.5a)} \qquad \text{(i)} \quad \rho(n+1) \leqslant \rho(n) + cn^{-1/d} \text{ and}$$

$$\text{(11.5b)} \qquad \text{(ii)} \quad m^{d-1}\rho(k) - m^{d-1}k^{(d-1)/d}r(k) \leqslant \rho(m^d k),$$

where $r(k) \to 0$ *as* $k \to \infty$, *then as* $n \to \infty$

$$\text{(11.6)} \qquad \rho(n) \sim \beta n^{(d-1)/d}$$

for a constant β.

To justify the hypotheses of Lemma 2 for ρ_{MST} and ρ_{TSP}, one has to develop some basic properties of the MST and TSP. The following lemma points out one of particular interest because it provides a bound which does not depend upon n.

LEMMA 3. *There is a constant c such that for all* $x > 0$, *one has* $\nu_{MST}(x) \leqslant cx^{-d}$, *where* $\nu_{MST}(x)$ *denotes the maximum number of edges larger than x any minimal spanning tree of* $\{x_1, x_2, \ldots, x_n\} \subset [0, 1]^d$.

The worst case behavior of the MST and TSP has a parallel for minimal and greedy matchings. Steele and Snyder (1987) studied these problems in the context of general power weighted edges, and provided exact asymptotic results to complement the bounds obtained in Avis (1981, 1983).

The sequence associated with minimal matching is naturally

$$\text{(11.7)} \qquad \rho_M(n) = \max_{V_n}\left\{\min_{M} \sum_{e \in M} |e|^{\alpha} : M \text{ is a matching of } V_n\right\},$$

but for the honest analysis of greedy matchings it turns out to be necessary to consider the two sequences:

$$(11.8)\qquad \rho_G(n) = \max_{V_n}\left\{\max_{M} \sum_{e\in M} |e|^\alpha : M \text{ is a greedy matching of } V_n\right\} \quad\text{and}$$

$$(11.9)\qquad \hat{\rho}_G(n) = \max_{V_n}\left\{\min_{M} \sum_{e\in M} |e|^\alpha : M \text{ is a greedy matching of } V_n\right\}.$$

The reason one is forced to consider both ρ_G and $\hat{\rho}_G$ is that one can have matchings M_1 and M_2 which are both *bona fide* greedy matchings of a set $V = \{x_1, x_2, \dots, x_n\}$, but for which

$$L_\alpha(M_1) = \sum_{e\in M_1} |e|^\alpha \neq L_\alpha(M_2) = \sum_{e\in M_2} |e|^\alpha.$$

This phenomenon is due to the possibility of ties, but when investigating worst case configurations, ties cannot be shrugged off. In fact, they are surely to be expected in most optimal configurations. As it happens, one can establish a minimax result which says that $\hat{\rho}_G(n) = \rho_G(n)$ for all $n \geqslant 1$, so after some worry, one comes back to the consideration of just two basic sequences. Moreover, the results for ρ_M and ρ_G are just as expected:

$$(11.10a)\qquad \rho_M(n) \sim \beta_M n^{(d-\alpha)/d} \quad\text{and}$$

$$(11.10b)\qquad \rho_G(n) = \hat{\rho}_G(n) \sim \beta_G n^{(d-\alpha)/d},$$

where β_M and β_G are nonzero constants which depend on $d \geqslant 2$.

The techniques behind (11.10a) and (11.10b) largely parallel the worst case analyses of the TSP and MST, but new turns are required, particularly for the greedy matchings. As in the stochastic analysis of the greedy matching, subadditive inequalities do not come easily from suboptimality, and one has to look hard at the underlying algorithmic process.

12. Non-Euclidean cousins. True to title and design, this review has focused on the probabilistic and worst case analyses of classical problems of combinatorial optimization in Euclidean space, but some recent results on the probabilistic analysis of closely related non-Euclidean problems are so interesting they require at least a brief look.

First, consider the weight M_n of the minimal spanning tree for the complete graph on n vertices where the weight of any edge $e = (i, j)$ is given by X_{ij} and the $\{X_{ij}: 1 \leqslant i < j < n\}$ are independent random variables. Timofeev (1984) established a number of interesting results concerning the behavior of M_n and in particular proved

$$(12.1)\qquad EM_n \leqslant 3.29$$

provided the $\{X_{ij}\}$ are uniformly distributed in the unit interval $[0, 1]$. Still, the crowning result in this direction is due to Frieze (1985):

$$(12.2)\qquad M_n \to \zeta(3)/F'(0)$$

where the convergence is in probability and in L^1, $\zeta(3) = \sum_{n=1}^{\infty} n^{-3} = 1.202\cdots$, and F is the common distribution of the nonnegative independent random variables $\{X_{ij}\}$.

The assignment problem with random costs has also had an interesting set of developments. If $K_{n,n}$ denotes the complete bipartite graph and independent uniform costs X_{ij} are associated with each edge of $K_{n,n}$, then the assignment problem is to determine the minimum value A_n of a matching of $K_{n,n}$.

The stochastic analysis of A_n was considered in the Bachelor's thesis of Lazarus (1979) who showed $EA_n \leqslant c \log n$. Independently, Walkup (1979) established the remarkable fact—in keeping with Timofeev's bound (12.1)—that for all $n \geqslant 1$ one has $EA_n \leqslant 3$. Most recently, Karp (1987) introduced a new conditioning method which was simpler than Walkup's method and which provided the sharper bound $EA_n \leqslant 2$. The development we wish to review in more detail on is the generalization and systematization of Karp's conditioning method provided by Dyer, Frieze, and McDiarmid (1986), and McDiarmid (1986).

It is well known that the value A_n—as well as all of the other functionals we have reviewed—can be expressed in terms of a linear programming problem. What is truly remarkable is that one can obtain good bounds on EA_n and similar quantities within that general framework.

Dyer, Frieze, and McDiarmid (1986) consider the general problem:

$$\text{(12.3)} \qquad \begin{aligned} &\text{minimize } \sum_{j=1}^{n} c_j x_j \\ &\text{subject to } \sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i = 1, 2, \ldots, m, \\ &\qquad\qquad x_j \geqslant 0, \quad j = 1, 2, \ldots, n. \end{aligned}$$

The coefficients c_j of the objective function are assumed to be independent, nonnegative random variables (with possibly different distributions), but all of the remaining parameters are assumed to be known constants. The Dyer-Frieze-McDiarmid inequalities relate the pointwise (random) optimal solution z^* of (12.3) to the fixed feasible solutions of (12.3). For example, if the c_i are independent and uniform on $[0, 1]$ and $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$ is a fixed feasible solution of (12.3), then the simplest of the Dyer-Frieze-McDiarmid inequalities is

$$\text{(12.4)} \qquad Ez^* \leqslant m \max\{\hat{x}_j : j = 1, 2, \ldots, n\}.$$

To see how (12.4) relates to Karp's bound on $E(A_n)$ we note the assignment problem can be written as

$$\text{(12.5)} \qquad \begin{aligned} &\text{minimize } z = \sum_{1 \leqslant i, j \leqslant n} c_{ij} x_{ij} \\ &\text{subject to} \quad \sum_{1 \leqslant j \leqslant n} x_{ij} = 1, \qquad 1 \leqslant i \leqslant n, \\ &\qquad\qquad \sum_{1 \leqslant i \leqslant n} x_{ij} = 1, \qquad 1 \leqslant j \leqslant n, \quad \text{and} \\ &\qquad\qquad x_{ij} \geqslant 0, \qquad 1 \leqslant i, j \leqslant n. \end{aligned}$$

Looking at the assignment problem in terms of the formulation (12.3), we see that

$m = 2n$ and also that $\hat{x}_{ij} = 1/n$ is a feasible solution. Consequently, inequality (12.4) implies Karp's inequality $E(A_n) \leqslant 2$.

The main result of Dyer, Frieze, McDiarmid (1986) is an apt generalization of inequality (12.4) to a class of random cost variables which retain a conditioning property of the uniform. Specifically, the generalization concerns independent random variables c_j for which there is a β, $0 < \beta \leqslant 1$ such that

$$E\big(c_j | c_j \geqslant h\big) \geqslant E(c_j) + \beta h \tag{12.6}$$

for all $h \geqslant 0$ with $P(c_j \geqslant h) > 0$.

The main inequality of Dyer, Frieze, and McDiarmid (1986) can be written as

$$E(z^*) \leqslant \beta^{-1} \max_{S:|S|=m} \left\{ \sum_{i \in S} \hat{x}_i E c_i \right\} \tag{12.7}$$

To see how (12.7) contains the handy inequality (12.4), we first note that for uniformly distributed c_j we can take $\beta = 1/2$ in (12.6) and $Ec_j = 1/2$, so the right side of (12.4) certainly majorizes the right side of (12.4).

The great distinction between the non-Euclidean cousins considered in this section and the problems considered previously rests entirely in the assumption of independence in the costs $\{c_j\}$. It is still hard to see how the methods of Dyer, Frieze, and McDiarmid (1986) can be used where the dependence of the costs carries as much complexity as it must in Euclidean problems, but nevertheless it seems likely that the Dyer-Frieze-McDiarmid inequalities will become part of the standard machinery of probability as it is applied in combinatorial optimization.

13. Conclusion. One intention of this review has been to show that the work initiated by Beardwood, Halton, and Hammersley (1959) and fueled by Karp (1976) has good prospects of growing into a rich and useful theory. The review should have also conveyed the spirit of a young theory with open problems, rough edges, and many opportunities for invention.

Added in Proof. The conjecture of R. Bland discussed in §6 has been proved recently by D. Aldous and the author by techniques unrelated to those of this survey.

Acknowledgement. A preliminary version of this review supplemented three lectures given at Cornell University as part of the *Year in Discrete Optimization* organized by the Center of Applied Mathematics and the Department of Operations Research and Industrial Engineering. The exposition has profited from the comments of Robert Bland and Timothy Snyder.

References

Adler, R. (1986). Comments in *Collected Works of S. Kakutani*, Vol. II, R. R. Kallman (Ed.), Birkhauser, Boston, 444.

Ajtai, M., Komlós, J. and Tusnády, G. (1984). On Optimal Matchings. *Combinatorica* **4** 259–264.

Avis, D. (1981). Worst Case Bounds for the Euclidean Matching Problem. *Comput. Math. Appl.* **7** 251–257.

______ (1983). A Survey of Heuristics for the Weighted Matching Problem. *Networks* **13** 475–493.

______ Davis, B. and Steele, J. M. (1988). Probabilistic Analysis of a Greedy Heuristic for Euclidean Matching. To appear in *Probability in the Engineering and Information Sciences.*

Bartholdi, J. J. and Platzman, L. K. (1988). Heuristics Based on Spacefilling Curves for Combinatorial Problems in Euclidean Space. *Management Sci.* **34** 291–305.

Beardwood, J., Halton, J.H. and Hammersley, J. M. (1959). The Shortest Path Through Many Points. *Proc. Cambridge Philos. Soc.* **55** 299–327.

Bingham, N. H. (1981). Tauberian Theorems and the Central Limit Theorem. *Ann. Probab.* **9** 221–231.

Coffman, E. G. and Shor, P. W. (1989). A Simple Proof of the $0(\sqrt{n}\ \log^{3/4} n)$ Up-Right Matching Bound. Technical Report, AT & T Bell Laboratories.

Dyer, M. E., Frieze, A. M. and McDiarmid, C. J. H. (1986). On Linear Programs with Random Costs. *Math. Programming* **35** 3–16.

Efron, B. and Stein, C. (1981). The Jackknife Estimate of Variance. *Ann. Statist.* **9** 586–596.

Fejes-Tóth, L. T. (1940). Uber einen geometrische Satz. *Math. Z.* **46** 83–85.

Few, L. (1955). The Shortest Path and the Shortest Road through n Points in a Region. *Mathematika* **2** 141–144.

Fisher, M. L. and Hochbaum, D. S. (1980). Probabilistic analysis of the Planar K-Median Problem. *Math. Oper. Res.* **5** 27–34.

Frieze, A. M. (1985). On the Value of a Random Spanning Tree Problem. *Discrete Appl. Math.* **10** 47–56.

Ghosh, H. W. (1949). Expected Travel among Random Points. *Bull. Calcutta Statist. Assoc.* **2** 83–87.

Gilbert, E. N. and Pollak, H. O. (1968). Steiner random minimal trees. *SIAM J. Appl. Math.* **16** 1–29.

Halton, J. H. and Terada, R. (1982). A Fast Algorithm for the Euclidean Traveling Salesman Problem, Optimal with Probability One. *SIAM J. Comput.* **11** 28–46.

Hartmann, M. (1987). Tauberian-like Theorems for Upper and Lower Bounds. Report No. 87471-OR, Institut für Okonometrie und Operationes Research, Universität Bonn.

Hochbaum, D. and Steele, J. M. (1982). Steinhaus' Geometric Location Problem for Random Samples in the Plane. *Adv. in Appl. Probab.* **14** 55–67.

Jessen, R. J. (1942). Statistical Investigation of a Sample Survey for Obtaining Farm Facts. *Bull. Iowa State Coll. Agricultural Res.* **309**.

Karlin, S. and Rinott, Y. (1982). Application of ANOVA Type Decompositions for Comparison of Variance Statistics Including Jackknife Estimates. *Ann. Statist.* **10** 485–501.

Karloff, H. (1989). "How Long Can a Euclidean Traveling Salesman Tour Be? *SIAM J. Discrete Math.* **2** 91–99.

Karp, R. M. (1976). The Probabilistic Analysis of Some Combinatorial Search Algorithms. *Algorithms and Complexity: New Directions and Recent Results* J. F. Traub (Ed.), Academic Press, New York, 1–19.

______ (1977). Probabilistic Analysis of Partitioning Algorithms for the Traveling-Salesman Problem in the Plane. *Math. Oper. Res.* **2** 209–224.

______ (1979). A Patching Algorithm for the Non-symmetric Traveling Salesman Problem. *SIAM J. Comput.* **8** 561–573.

______ (1987). An Upper Bound on the Expected Cost of an Optimal Assignment. 1–4, in *Discrete Algorithms and Complexity: Proc. Japan-US Joint Seminar* D. Johnson, et al. (Eds.), Academic Press, New York.

______ Lenstra, J. K., McDiarmid, C. J. H., and Rinnooy Kan, A. H. G. (1984). Probabilistic Analysis of Combinatorial Algorithms: An Annotated Bibliography in *Combinatorial Optimization: Annotated Bibliographies* M. O'hEigeartaigh, J. K. Lenstra and A. H. G. Rinnooy Kan (Eds.), John Wiley and Sons, New York.

______ and Steele, J. M. (1985). Probabilistic Analysis of Heuristics. *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization* Lawler, E. L., et al. (Eds.), John Wiley and Sons, New York, 181–206.

Kern, W. (1986). On the Rate of Convergence of Some Stochastic Processes. Technical Report No. 85.26, Mathematisches Institut, Universität zu Köln, West Germany.

Kesten, H. (1973). Contribution to J. F. C. Kingman. Subadditive ergodic theory. *Ann. Probab.* **1** 883–909.

Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G. and Shmoys, D. B. (1985). *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. John Wiley and Sons, New York.

Lazarus, A. (1979). The Assignment Problem with Uniform (0, 1) Cost Matrix. Unpublished B.A. Thesis, Department of Mathematics, Princeton University, Princeton, NJ.

Leighton, T. and Shor, P. (1986). Tight Bounds for Minimax Grid Matching with Applications to the Average Case Analysis of Algorithms. *Proc. 18th ACM Symp. Theory of Computing*, 91–103.

Mahalanobis, P. C. (1940). A Sample Survey of the Acreage under Jute in Bengal. *Sankyā* **4** 511–531.

Marks, E. S. (1948). A Lower Bound for the Expected Travel among m Random Points. *Ann. Math. Statist.* **19** 419–422.

McDiarmid, C. (1986). On the Greedy Algorithm with Random Costs. *Math. Programming* **36** 245–255.

Milne, S. C. (1980). Peano Curves and Smoothness of Functions. *Adv. in Math.* **35** 129–157.

Moran, S. (1984). On the Length of Optimal TSP Circuits in Sets of Bounded diameter. *J. Combinatorial Theory Ser. B* **37** 113–141.

Papadimitriou, C. H. (1977). The Euclidean Travelling Salesman Problem is NP-Complete. *Theoretical Computer Sci.* **4** 237–244.

Papadimitriou, C. H. (1978). The Probabilistic Analysis of Matching Heuristics. *Proc. 15th Annual Conf. Comm. Contr. Computing*, Univ. Illinois, Champaign-Urbana.

______ (1981). Worst-Case and Probabilistic Analysis of a Geometric Location Problem. *SIAM J. Comput.* 10, **3** 542–557.

Rédei, L. (1934). Ein kombinatorischer Satz. *Acta Szeged* **7** 39–43.

Rhee, W. T. (1988). On the Variance of the Stochastic Traveling Salesperson Problem. Technical Report, Ohio State University.

______ and Talagrand, M. (1986). Martingale Inequalities and the Jackknife Estimate of Variance. *Statist. Probab. Lett.* **4** 5–6.

______ and ______ (1987). Martingale Inequalities and NP-Complete Problems. *Math. Oper. Res.* **12** 177–181.

______ and ______ (1989a). Martingale Inequalities, Interpolation, and NP-Complete Problems. To appear in *Math. of Oper. Res.*

______ and ______ (1989b). A Sharp Deviation for the Stochastic Traveling Salesman Problem. *Ann. Probab.* **17** 1–8.

Richardon, D. (1973). Random Growth in a Tessellation. *Proc. Cambridge Philos. Soc.* **74** 515–528.

Rootzén, H. (1983). Central Limit Theory for Martingales via Random Change of Time. *Probab. Math. Statist. Essays in Honor of Carl-Gustav Esseen.* A. Gut and L. Holst (Eds.), Department of Mathematics, Uppsala University, Uppsala, Sweden.

Rosenkrantz, D. J., Stearns, E. and Lewis, P. M. (1977). An Analysis of Several Heuristics for the Traveling Salesman Problem. *SIAM J. Comput.* **6**, **3** 563–581.

Shamir, E. (1989). Presentation at Bell Communications Research, Morristown, NJ, February.

Shor, P. W. (1985). Random Planar Matching and Bin Packing. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Steele, J. M. (1981a). Subadditive Euclidean Functionals and Non-linear Growth in Geometric Probability. *Ann. Probab.* **9** 365–376.

______ (1981b). Complete Convergence of Short Paths and Karp's Algorithm for the TSP. *Math. Oper. Res.* **6** 374–378.

______ (1982). Optimal Triangulation of Random Samples in the Plane. *Ann. Probab.* **10** 548–553.

______ (1986a). Probabilistic Algorithm for the Directed Traveling Salesman Problem. *Math. Oper. Res.* **11** 343–350.

______ (1986b). An Efron Stein Inequality for Non-symmetric Statistics. *Ann. Statist.* **14** 753–758.

______ (1988). Growth Rates of Euclidean Minimal Spanning Trees with Power Weighted Edges. *Ann. Probab.* **16** 1767–1787.

______ Shepp, L. A. and Eddy, W. F. (1987). On the Number of Leaves of a Euclidean Minimal Spanning Tree. *J. Appl. Probab.* **24** 809–826.

______ and Snyder, T. L. (1987). Worst Case Growth Rates for Minimal and Greedy Matchings with Power Weighted Edges. Technical Report, Princeton University, Program in Operations Research and Statistics.

______ and ______ (1989). Worst Case Growth Rates of Some Classical Problems of Computational Geometry, *SIAM J. Computation* **18** 278–287.

Supowit, K. J., Reingold, E. M. and Plaisted, D. A. (1983). The Traveling Salesman Problem and Minimum Matching in the Unit Square. *SIAM J. Comput.* **12** 144–156.

Talagrand, M. (1989). Complete Convergence of the directed TSP preliminary manuscript.

Timofeev, E. A. (1984). Random Minimal Trees. *Theory Probab. Appl.* **29** 134–141.

Troyon, M. (1988). Quelques Heuristiques et Resultats Asymptotiques pour Trois Problems d'Optimisation Combinatoire. Ph.D. Thesis, Ecole Polytechnique Federale de Lausanne, Laussane, Switzerland.

Verblunsky, S. (1951). On the Shortest Path through a Number of Points. *Proc. Amer. Math. Soc.* **2** 904–913.

Vitale, R. A. (1984). An Expansion for Symmetric Statistics and the Efron-Stein Inequality. *Inequalities in Statistics and Probability IMS Lecture Notes—Monograph Series* Vol. 5, Y. L. Tong (Ed.), 112–114.

Walkup, D. W. (1979). On the Expected Value of a Random Assignment Problem, *SIAM J. Comput.* **8**, **3** 440–442.

Weide, B. W. (1978). Statistical Methods in Algorithm Design and Analysis. Unpublished Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, PA.

DEPARTMENT OF STATISTICS, WHARTON SCHOOL, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PENNSYLVANIA 19104