PROBABILISTIC BEHAVIOR OF SHORTEST PATHS OVER UNBOUNDED REGIONS

Andrew Chi-Chih Yao

CS-TR-313-91

March 1991

Probabilistic Behavior Of Shortest Paths Over Unbounded Regions[1]

Andrew Chi-Chih Yao

Department of Computer Science
Princeton University
Princeton, New Jersey 08544

Abstract

Let $k > 1$ and P be a probability distribution over R^k with all its absolute μ-th moments being finite for some $\mu > k/(k-1)$. Let $v_1, v_2, \cdots$ be an infinite sequence of random points, each indepdently distributed according to P. It is shown that the length of the shortest traveling-salesman's tour through $v_1, v_2, \cdots, v_n$ is, for large n, almost surely around $\alpha_P n^{(k-1)/k}$ for some constant α_P. This proves a conjecture of Beardwood, Halton and Hammersley (*Proc. Camb. Phil. Soc.* **55** (1959), 299-327).

[1]This research was supported in part by the National Science Foundation under grant number CCR-8813283.

1 Introduction

Let $k > 1$ be an integer, and P be any probability distribution over a bounded region in R^k. Let f be a density function for the absolute continuous part of P. In 1959, Beardwood, Halton, and Hammersley [BHH] showed that, for an infinite sequence of random points each distributed independently according to P, the length X_n of the shortest traveling salesman's tour through the first n points has the property that $X_n/n^{(k-1)/k}$ converges to $\beta_k k^{1/2} \int_{R^k} (f(v))^{(k-1)/k} dv$ almost surely, where β_k is a positive constant dependent only on k. They conjectured [BHH] that the theorem remains true for P defined over unbounded regions in R^k, provided that all the $k/(k-1)$-th absolute moments of P are finite. The purpose of the present paper is to prove this conjecture.

Let $k > 1$, and P be a probability distribution over R^k. Let $\alpha_P = \beta_k k^{1/2} \int_{R^k} (f(v))^{(k-1)/k} dv$ (in general may be infinite), where f is a density function for the absolute continuous part of P. Let $v_1, v_2, \cdots$ be an infinite sequence of random points distributed independently according to P. Let X_n be the minimum length of any traveling salesman's tour passing through $v_1, v_2, \cdots, v_n$.

Theorem 1 If P has all its finite absolute μ-th moments for some $\mu > k/(k-1)$, then $X_n/n^{(k-1)/k}$ converges to α_P almost surely as $n \to \infty$. Furthermore, $\lim_{n\to\infty} \mathbf{E} X_n/n^{(k-1)/k} = \alpha_P$.

Note that for P satisfying the condition in the theorem, $\int_{R^k} r^\mu dP$ is finite, where r stands for $(\sum_{1 \le i \le k} x_i^2)^{1/2}$. This implies that α_P is finite as $\int (f(v))^{(k-1)/k} dv$ can be written as $\int (r^\mu f)^{(k-1)/k} (r^{-(k-1)\mu})^{1/k} dv \le (\int r^\mu f dv)^{(k-1)/k} (\int r^{-(k-1)\mu} dv)^{1/k}$ by Hölder's inequality (see e.g. [HLP, p.22]).

In Steele [S1], the theorem of Beardwood, Halton, and Hammersley was extended to a class of general functionals which includes as a special case the length of the shortest traveling salesman's tour. We will see that Theorem 1 is also true for some of these functionals.

The idea used in the proof of Theorem 1 can be used to give a fast approximate algorithm which, with high probability, produces traveling salesman's tours to within a factor of $1 + \epsilon$ of the optimal length, for probability distributions P under the same constraints. This extends a result of Karp [K] for probability distributions P over bounded regions, and will be discussed elsewhere.

We state some useful elementary facts in Section 2. Theorem 1 is proved for a special family of probability distributions in Section 3. The general proof of the theorem is given in Section 4. An extension of Theorem 1 to a class of functionals will be given in Section

5. Some concluding remarks are given in Section 6.

2 Preliminaries

We collect some well-known facts. Throughout this paper, k, n denote integers greater than 1, and ℓ denotes any positive real number. Let a_k be the constant $\pi^{k/2}(\Gamma(1+k/2))^{-1}$, where $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ for $x > 0$.

Fact 1 (see, e.g. [Si, p.26]) The volume of a ball of radius r in R^k is equal to $a_k r^k$.

Definition 1 For any finite set S of points in R^k, let $T(S)$ denote the length of a shortest traveling salesman's tour through all the points in S. Let $M(S)$ stand for the length of a shortest Steiner tree connecting all the points in S. It is agreed that $T(\emptyset) = M(\emptyset) = 0$. For any $\ell > 0$, let $C_{k,\ell}$ denote the cube $[-\ell, \ell]^k$; let e_ℓ denote the point $(\ell, 0, 0, \cdots, 0)$, which is on a face of the cube $C_{k,\ell}$. Let $T_-(S;\ell) = T(S \cap C_{k,\ell})$ and $T_+(S;\ell) = T((S \setminus C_{k,\ell}) \cup \{e_\ell\})$.

Fact 2 For any finite sets $S \subseteq S' \subseteq R^k$, $M(S) \leq M(S')$ and $T(S) \leq T(S')$.

Fact 3 For any finite set $S \subseteq R^k$, $T(S) \leq 2M(S)$.

Fact 4 There exists a constant $b_k > 0$ such that the following are true: any set S of n points satisfies $T(S) \leq b_k n^{(k-1)/k}\ell$ if the points are contained in the cube $[-\ell, \ell]^k$, and satisfies $T(S) \leq b_k n^{(k-2)/(k-1)}\ell$ if the points are on a face of the cube $[-\ell, \ell]^k$.

Definition 2 For any $v = (x_1, x_2, \cdots, x_k) \in R^k$. let $\|v\| = (\sum_{1 \leq i \leq k} x_i^2)^{1/2}$ denote the norm of v in the L_2-metric.

3 A Special Case

In this section, we prove Theorem 1 for a special family of probability distributions. This will illustrate the basic ingredients in the general proof, and in the process derive some results useful for the general proof.

We first describe the approach. Consider a large cube of side-length ℓ_n centered at the origin. Divide the n random points $v_1, v_2, \cdots, v_n$ into two sets S_- and S_+, depending on whether they lie inside or outside the cube. As the probability density tapers off away from the origin, it is plausible that the best traveling salesman's tour T for the n points is about the same length as that for S_-. Therefore, to obtain information on X_n, it suffices to study the behavior of the shortest traveling salesman's tour T_- on S_-, with ℓ_n going to ∞ in some fashion. If ℓ_n is only required to grow with n at a moderate rate, then the existing methods of analysis of the problem for bounded regions can be utilized to analyze T_-.

In the next four lemmas, we derive some quantitative results in preparation for carrying out the above outline. Lemmas 1 and 2 show how T is related to T_- and T_+ (the length of the shortest traveling salesman's tour for S_+), and gives estimate on T_+ so that we know when it can be safely ignored. Lemmas 3 and 4 extend the existing analysis of the traveling salesman's tour over bounded regions to that over regions growing with n. The information provided by these four lemmas then enables us to prove the theorem by choosing ℓ_n judiciously.

Let $\rho_{k,\lambda}(v) = c_{k,\lambda}(1 + \|v\|)^{-\lambda}$ be a probability density over R^k, where $\lambda > k^2/(k-1)$ and $c_{k,\lambda} = \left(\int_{R^k}(1+\|v\|)^{-\lambda}dv\right)^{-1}$. (Note that $c_{k,\lambda}$ is well-defined and nonzero.) We will prove Theorem 1 for P defined by the density function $\rho_{k,\lambda}$.

Let $v_1, v_2, \cdots$ be an infinite sequence of random points in R^k, where each v_i is independently distributed according to $\rho_{k,\lambda}$. Define random variables $V_n = \{v_1, v_2, \cdots, v_n\}$, and $X_n = T(V_n)$. For any $\ell > 0$, let $Y_{n,\ell} = T_-(V_n; \ell)$ and $Z_{n,\ell} = T_+(V_n; \ell)$.

Lemma 1 $Y_{n,\ell} \leq X_n \leq Y_{n,\ell} + Z_{n,\ell} + 2k^{1/2}\ell$.

Corollary If $0 < \ell \leq \ell'$, then $Y_{n,\ell} \leq Y_{n,\ell'} \leq Y_{n,\ell} + Z_{n,\ell} + 2k^{1/2}\ell$.

Proof The inequality $Y_{n,\ell} \leq X_n$ follows immediately from Fact 2. To derive the other inequality, let $S_- = V_n \cap C_{k,\ell}$ and $S_+ = (V_n \setminus C_{k,\ell}) \cup \{e_\ell\}$. We can assume that S_- is nonempty; otherwise the lemma is clearly true. Take any $v_i \in S_-$ and connect the shortest traveling salesman's tours for S_- and S_+ by two copies of the edge $\{v_i, e_\ell\}$. This creates a connected Eulerian graph on the points in $V_n \cup \{e_\ell\}$ of total length no greater than $Y_{n,\ell} + Z_{n,\ell} + 2k^{1/2}\ell$. It is well known (see e.g. [K]) that it can be transformed into a traveling salesman's tour on $V_n \cup \{e_\ell\}$ of equal or less length. Lemma 1 now follows from Fact 2. The corollary is true since $Y_{n,\ell} \leq Y_{n,\ell'} \leq X_n$ by Fact 2. □

Let $d_{k,\lambda} = 2\,k^{3/2}\,\Gamma(1/k)\,a_k\,(c_{k,\lambda})^{(k-1)/k}\,((k-1)\,\lambda - k^2)^{-1}$. Clearly $d_{k,\lambda} > 0$.

Lemma 2 $\mathbf{E}Z_{n,\ell} \leq 2\,d_{k,\lambda} n^{(k-1)/k}\,\ell^{-(\lambda-k-\lambda/k)} + 4\,b_k\,\ell\,n^{(k-2)/(k-1)}$.

Proof For each $u \in R^k \setminus C_{k,\ell}$, let J_u denote the point closest to u among all points u' such that (a) $\|u'\| < \|u\|$, and (b) u' lie in either $V_n \setminus C_{k,\ell}$ or on the faces of the cube $C_{k,\ell}$. For each $1 \leq i \leq n$, define a random variable D_i which takes on the value 0 if $v_i \in C_{k,\ell}$ and, otherwise, the value $\|v_i - J_{v_i}\|$.

Let $S_+ = (V_n \setminus C_{k,\ell}) \cup \{e_\ell\}$. Then a Steiner tree for S_+ can be obtained by adding the set of edges $\{v_i, J_{v_i}\}$, $v_i \in V_n \setminus C_{k,\ell}$, to a shortest traveling salesman's tour for the points in J, where $J = \{J_{v_i} \mid v_i \in V_n \setminus C_{k,\ell};\ J_{v_i} \in C_{k,\ell}\} \cup \{e_\ell\}$. From Facts 3 and 4, we have

$$\begin{aligned} Z_{n,\ell} &\leq 2M(S_+) \\ &= 2(\sum_{1\leq i\leq n} D_i) + 2T(J) \\ &\leq 2(\sum_{1\leq i\leq n} D_i) + 2\, b_k\, \ell\, |J|^{(k-2)/(k-1)} \\ &\leq 2(\sum_{1\leq i\leq n} D_i) + 2\, b_k\, \ell\, (n+1)^{(k-2)/(k-1)}. \end{aligned}$$

Thus,

$$\mathbf{E}Z_{n,\ell} \leq 2 \sum_{1\leq i\leq n} \mathbf{E}D_i + 4\, b_k\, \ell\, n^{(k-2)/(k-1)}. \tag{1}$$

To estimate $\mathbf{E}D_i$, let r_u denote the Euclidean distance between any point $u \in R^k$ and the set $C_{k,\ell}$. Observe that, for $u \in R^k \setminus C_{k,\ell}$ and $0 \leq s \leq r_u$, one can construct a cube K of side-length $s\, k^{-1/2}$ such that its main diagonal is the line segment connecting u and the point $(1 - s/\|u\|)\, u$. As every point u' in K is within a distance s from u and $\rho_{k,\lambda}(u') \leq \rho_{k,\lambda}(u)$, the conditional probability distribution of D_i satisfies

$$\begin{aligned} \Pr\{\, D_i > s \mid v_i = u \,\} &\leq (1 - \rho_{k,\lambda}(u)\, (sk^{-1/2})^k)^{n-1} \\ &\leq \exp(-(n-1)\, k^{-k/2}\, c_{k,\lambda}\, s^k\, (1 + \|u\|)^{-\lambda}). \end{aligned} \tag{2}$$

Clearly, inequality (2) is valid for $s > r_u$, since $\Pr\{\, D_i > s \mid v_i = u \,\} = 0$. It follows that

$$\begin{aligned} \mathbf{E}(D_i \mid v_i = u\,) &= \int_0^\infty \Pr\{\, D_i > s \mid v_i = u \,\} ds \\ &\leq \int_0^\infty \exp(-(n-1)\, k^{-k/2}\, c_{k,\lambda}\, s^k\, (1 + \|u\|)^{-\lambda}) ds \\ &= k^{-1/2}\, \Gamma(1/k)\, (c_{k,\lambda}(n-1))^{-1/k}\, (1 + \|u\|)^{\lambda/k}. \end{aligned} \tag{3}$$

Hence,

$$\begin{aligned} \mathbf{E}D_i &\leq k^{-1/2}\, \Gamma(1/k)\, (c_{k,\lambda}\, (n-1))^{-1/k} \int_{Q_\ell} (1 + \|u\|)^{\lambda/k} \rho_{k,\lambda}(u)\, du \\ &\leq k^{1/2}\, \Gamma(1/k)\, a_k\, (c_{k,\lambda})^{(k-1)/k}\, (n-1)^{-1/k} \int_\ell^\infty (1 + \|u\|)^{-\lambda+\lambda/k}\, \|u\|^{k-1} d\|u\| \\ &\leq k^{1/2}\, \Gamma(1/k)\, a_k\, (c_{k,\lambda})^{(k-1)/k}\, (n-1)^{-1/k} \int_\ell^\infty r^{-\lambda+\lambda/k+k-1} dr \\ &\leq d_{k,\lambda}\, n^{-1/k}\, \ell^{-(\lambda-k-\lambda/k)}. \end{aligned} \tag{4}$$

Lemma 2 follows from (1) and (4). □

Let $d'_{k.\lambda} = 32\, (\Gamma(1/k))^2\, k^{-1} (c_{k,\lambda})^{-2/k}$. Let $\psi_k(n) = 1 + \ln n$ if $k = 2$, and $n^{(k-2)/k}\, k/(k-2)$ if $k > 2$.

Lemma 3 Let $s > 0$. Then $\Pr\{ |Y_{n,\ell} - \mathbf{E}Y_{n,\ell}| > s \} \leq 2\, e^{-s^2/\nu}$, where $\nu = 128k\ell^2 + d'_{k.\lambda}\,(1 + k^{1/2}\,\ell)^{2\lambda/k}\,\psi_k(n)$.

Proof The proof employs an approach used by Rhee and Talagrand ([RT]) (also see Steele [S2]) to obtain similar bounds for the traveling salesman problem over the unit cube.

For $1 \leq i \leq n$, let σ_i denote the sigma field generated by $v_1, v_2, \cdots, v_i$. Let g_i, $1 \leq i \leq n$, be the sequence of martingale differences defined as $\mathbf{E}(Y_{n,\ell} \,|\, \sigma_i) - \mathbf{E}(Y_{n,\ell} \,|\, \sigma_{i-1})$. Let $w_i = \sup |g_i|$. Then, by Azuma's Inequality (Hoeffding [H], Azuma [A]), we have

$$\Pr\{ |\sum_{1\leq i\leq n} g_i| > s \} \leq 2\,\exp(-s^2/(2 \sum_{1\leq i\leq n} w_i^2)).$$

Since $Y_{n,\ell} - \mathbf{E}Y_{n,\ell} = \sum_{1\leq i\leq n} g_i$, we have

$$\Pr\{ |Y_{n,\ell} - \mathbf{E}Y_{n,\ell}| > s \} \leq 2\,\exp(-s^2/(2 \sum_{1\leq i\leq n} w_i^2)). \tag{5}$$

It remains to evaluate w_i.

Let $V_n = \{v_1, v_2, \cdots, v_n\}$ and $V'_n = (V_n \setminus \{v_i\}) \cup \{v'_i\}$. It is not hard to see that

$$|T_-(V_n;\ell) - T_-(V'_n;\ell)| \leq 2 \sum_{u\in\{v_i,v'_i\}\cap C_{k,\ell}} A_u,$$

where $A_u = 2\,k^{1/2}\,\ell$ if none of the v_j, $i+1 \leq j \leq n$, is in the cube $C_{k,\ell}$, and otherwise equal to $\min\{\|u - v_j\| \,|\, i+1 \leq j \leq n\}$. This implies

$$w_i \leq 4 \sup_{u\in C_{k,\ell}} \mathbf{E}A_u. \tag{6}$$

Let $1 \leq i \leq n-1$. Adopting the method used in the proof of Lemma 2, we find for any $u \in C_{k,\ell}$ and $0 \leq s \leq 2\,k^{1/2}\,\ell$

$$\begin{aligned} \Pr\{ A_u > s \} &\leq (1 - (s/k^{1/2})^k\, \rho_{k,\lambda}(\ell,\ell,\cdots,\ell)\,)^{n-i} \\ &\leq \exp(-(n-i)\,k^{-k/2}\,s^k\,c_{k,\lambda}\,(1 + k^{1/2}\,\ell)^{-\lambda}). \end{aligned}$$

Clearly, the above inequality is also valid for $s > 2\,k^{1/2}\,\ell$, since in this case $\Pr\{ A_u > s \,|\, v_i = u \} = 0$. It follows that

$$\begin{aligned} \mathbf{E}A_u &= \int_0^\infty \Pr\{ A_u > s \} ds \\ &\leq \int_0^\infty \exp(-(n-i)\,k^{-k/2}\,s^k\,c_{k,\lambda}\,(1 + k^{1/2}\,\ell)^{-\lambda}) ds \\ &= k^{-1/2}\,\Gamma(1/k)\,(c_{k.\lambda})^{-1/k}\,(1 + k^{1/2}\,\ell)^{\lambda/k}\,(n-i)^{-1/k}. \end{aligned}$$

Therefore,

$$\begin{aligned} 2\sum_{1\le i\le n-1} w_i^2 &\le d'_{k,\lambda}(1+k^{1/2}\ell)^{\lambda/k}\sum_{1\le i\le n-1}(n-i)^{-2/k} \\ &\le d'_{k,\lambda}(1+k^{1/2}\ell)^{\lambda/k}(1+\int_1^{n-1} x^{-2/k}dx) \\ &\le d'_{k,\lambda}(1+k^{1/2}\ell)^{\lambda/k}\psi_k(n)\,. \end{aligned} \tag{7}$$

Lemma 3 follows from (5)-(7) and the fact that $w_n = 8k^{1/2}\ell$. □

Lemma 4 For any fixed $t>0$, $\lim_{n\to\infty} \mathbf{E}Y_{n,t}/n^{(k-1)/k} = \beta_k\, k^{1/2} \int_{C_{k,t}} (\,\rho_{k,\lambda}(u))^{(k-1)/k}\, du\,.$

Proof Let θ be the probability that a random point v distributed according to $\rho_{k,\lambda}$ falls into the cube $C_{k,t}$. Clearly, $\theta = \int_{C_{k,t}} \rho_{k,\lambda}(u)du > 0$.

Let $\rho_{k,\lambda,t}$ denote the probability density when $\rho_{k,\lambda}$ is restricted to the cube $C_{k,t}$. Then $\rho_{k,\lambda,t}(v) = \theta^{-1}\rho_{k,\lambda}(v)$ for $v \in C_{k,t}$ and 0 otherwise. Let $u_1, u_2, \cdots$ be a sequence of random points independently distributed according to $\rho_{k,\lambda,t}$, and let $W_{n,t}$ denote the length of the shortest traveling salesman's tour through the first n points. Let $\gamma_{k,\lambda,t} = \beta_k\, k^{1/2} \int_{C_{k,t}} (\,\rho_{k,\lambda}(u))^{(k-1)/k}\, du$. The result of Beardwood, Halton and Hammersley ([BHH]) implies that

$$\lim_{n\to\infty} \mathbf{E}W_{n,t}/n^{(k-1)/k} = \theta^{-(k-1)/k}\,\gamma_{k,\lambda,t}. \tag{8}$$

Let $p_{n,n'}$ be the probability that n' out of n random points v_i, each of which independently distributed according to $\rho_{k,\lambda}$, fall into $C_{k,t}$. Then

$$\mathbf{E}Y_{n,t} = \sum_{n'} p_{n,n'}\mathbf{E}W_{n',t}. \tag{9}$$

Also, by Chernoff's bound [C], we have

$$\sum_{n',|n'-\theta\, n|>\theta\, n^{2/3}} p_{n,n'} \le 2\, e^{-2\theta^2\, n^{1/3}}. \tag{10}$$

For any n' satisfying $|n'-\theta n| \le \theta n^{2/3}$, it is elementary to show that, for all sufficiently large n,

$$\left|(n')^{(k-1)/k} - (\theta\, n)^{(k-1)/k}\right| \le 4\, n^{-1/3}\, (\theta\, n)^{(k-1)/k}. \tag{11}$$

It follows from (8)-(11) that

$$\mathbf{E}Y_{n,t} \;=\; \sum_{n',|n'-\theta n|\le\theta n^{2/3}} p_{n,n'}\mathbf{E}W_{n',t} + O(n\, e^{-2\theta^2\, n^{1/3}})$$

$$
\begin{aligned}
&= \sum_{n',|n'-\theta n|\le\theta n^{2/3}} p_{n,n'}\,(\theta)^{-(k-1)/k}\,\gamma_{k,\lambda,t}\,(n')^{(k-1)/k}(1+o(1)) + O(n\,e^{-2\theta^2 n^{1/3}}) \\
&= \sum_{n',|n'-\theta n|\le\theta n^{2/3}} p_{n,n'}\,(\theta)^{-(k-1)/k}\,\gamma_{k,\lambda,t}\,(\theta n)^{(k-1)/k}(1+o(1)) + o(1) \\
&= (1+o(1))\,\gamma_{k,\lambda,t}\,n^{(k-1)/k} \sum_{n',|n'-\theta n|\le\theta n^{2/3}} p_{n,n'} + o(1) \\
&= (1+o(1))\,\gamma_{k,\lambda,t}\,n^{(k-1)/k}.
\end{aligned}
$$

This proves Lemma 4. □

To apply the above lemmas to prove the theorem, we choose an ℓ_n such that Y_{n,ℓ_n} is a close approximation to X_n, and at the same time C_{k,ℓ_n} is small enough that we can extend the known method used for bounded regions to analyze Y_{n,ℓ_n}.

Let $\epsilon = (40\,\lambda\,k^2)^{-1}$ and $\epsilon' = \epsilon\,(\lambda - k - \lambda/k)$. Then $\epsilon, \epsilon' > 0$. Let $\ell_n = n^{\epsilon}$ for all $n > 1$.

Lemma 5 $\lim_{n\to\infty} \mathbf{E}X_n/n^{(k-1)/k} = \lim_{n\to\infty} \mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k} = \alpha_P$, where P is the probability distribution with density function $\rho_{k,\lambda}$.

Proof Let K_n stand for either X_n or Y_{n,ℓ_n}. Let $t > 0$. By the corollary to Lemma 1, we have

$$\mathbf{E}Y_{n,t} \le \mathbf{E}K_n \le \mathbf{E}Y_{n,t} + \mathbf{E}Z_{n,t} + 2k^{1/2}\,t\,. \tag{12}$$

Let $\gamma_{k,\lambda,t} = \beta_k\,k^{1/2}\int_{C_{k,t}} (\,\rho_{k,\lambda}(v))^{(k-1)/k}\,dv$. By Lemmas 2, 4 and inequality (12) we have

$$\gamma_{k,\lambda,t} \le \liminf_{n\to\infty} \mathbf{E}K_n/n^{(k-1)/k} \le \limsup_{n\to\infty} \mathbf{E}K_n/n^{(k-1)/k} \le \gamma_{k,\lambda,t} + 2\,d_{k,\lambda}\,t^{-(\lambda-k-\lambda/k)}. \tag{13}$$

Letting $t \to \infty$, we obtain the lemma. □

We have thus proved one of the equations stated in Theorem 1. It remains to prove the other equation: almost surely,

$$\lim_{n\to\infty} X_n/n^{(k-1)/k} = \alpha_P, \tag{14}$$

where P is the probability distribution with density $\rho_{k,\lambda}$.

We specialize the general bounds discussed in Lemmas 2 and 3 with the choice of $\ell = \ell_n$. We will see that $T_+(V_n;\ell_n)$ is small (and hence $T(V_n)$ is essentially equal to $T_-(V_n;\ell_n)$), and that the probabilistic behavior of $T_-(V_n;\ell_n)/n^{(k-1)/k}$ is highly concentrated around its average value (and hence around α_P by Lemma 5).

Lemma 6 $\mathbf{E}Z_{n,\ell_m} \le 6\,d_{k,\lambda}\,n^{(k-1)/k-\epsilon'}$ for all sufficiently large n and $n/2 \le m \le n$.

Proof It follows from Lemma 2 and the fact that $(k-1)/k - \epsilon(\lambda - k - \lambda/k) > (k-2)/(k-1) + \epsilon$. □

Lemma 7 $\Pr\{\,|Y_{n,\ell_n} - \mathbf{E}Y_{n,\ell_n}| > n^{1/2-\epsilon}\,\} \le e^{-n^{1/k}}$ for all sufficiently large n.

Proof It follows from Lemma 3 by setting $\ell = \ell_n$, $s = n^{1/2-\epsilon}$, and observing that $s^2/\nu = \Omega(n^t/\log n)$ where $t = 2/k - 2\epsilon - 2\lambda\epsilon/k > 1/k$. □

We are now ready to prove (14). Let $\delta > 0$ be any fixed number, and B_n be the event that $|X_n/n^{(k-1)/k} - \alpha_P| > \delta$. We need to prove that

$$\lim_{m\to\infty} \Pr\{\cup_{n\ge m} B_n\} = 0. \tag{15}$$

Let F_n be the event that $|Y_{n,\ell_n} - \mathbf{E}Y_{n,\ell_n}| > n^{1/2-\epsilon}$, and G_n be the event that $Z_{n,\ell_n} > n^{(k-1)/k-\epsilon'/2}$.

From Lemma 1, we have

$$\begin{aligned} |X_n/n^{(k-1)/k} - \alpha_P| &\le |X_n/n^{(k-1)/k} - Y_{n,\ell_n}/n^{(k-1)/k}| + |Y_{n,\ell_n}/n^{(k-1)/k} - \mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k}| \\ &\quad + |\mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k} - \alpha_P| \\ &\le Z_{n,\ell_n}/n^{(k-1)/k} + 2\,k^{1/2}\,\ell_n\,/n^{(k-1)/k} + |Y_{n,\ell_n}/n^{(k-1)/k} - \mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k}| \\ &\quad + |\mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k} - \alpha_P|\,. \end{aligned} \tag{16}$$

Now, $2\,k^{1/2}\,\ell_n\,/n^{(k-1)/k} < \delta/4$, and by Lemma 5, $|\mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k} - \alpha_P| < \delta/4$ for all large n. We conclude that, for all sufficiently large n, B_n implies $F_n \cup G_n$.

Let $G^{(j)}$ be the event $\cup_{2^j \le n < 2^{j+1}} G_n$. We have, for all large m,

$$\begin{aligned} \Pr\{\cup_{n\ge m} B_n\} &\le \sum_{n\ge m} \Pr\{F_n\} + \sum_{n\ge m} \Pr\{G_n\} \\ &\le \sum_{n\ge m} \Pr\{F_n\} + \sum_{j\ge\lfloor\log_2 m\rfloor} \Pr\{G^{(j)}\}. \end{aligned} \tag{17}$$

By Lemma 7,

$$\Pr\{F_n\} \le e^{-n^{1/k}}. \tag{18}$$

Also, note that by Fact 2, $Z_{n,\ell_n} \le Z_{2^{j+1},\ell_{2^j}}$ for all $2^j \le n < 2^{j+1}$. Thus, Lemma 6 implies that, for all sufficiently large j,

$$\begin{aligned} \Pr\{G^{(j)}\} &\le \Pr\{Z_{2^{j+1},\ell_{2^j}} > (2^j)^{(k-1)/k-\epsilon'/2}\} \\ &\le (\mathbf{E}Z_{2^{j+1},\ell_{2^j}})/(2^j)^{(k-1)/k-\epsilon'/2} \\ &\le 6\,d_{k.\lambda}\,(2^{j+1})^{(k-1)/k-\epsilon'}/(2^j)^{(k-1)/k-\epsilon'/2} \\ &\le 12\,d_{k.\lambda}\,(2^{-\epsilon'/2})^j\,. \end{aligned} \tag{19}$$

It follows from (17)-(19) that, for all large m,

$$
\begin{aligned}
\Pr\{\cup_{n\geq m} B_n\} &\leq \sum_{n\geq m} e^{-n^{1/k}} + 12\, d_{k,\lambda} \sum_{j\geq \lfloor \log_2 m \rfloor} (2^{-\epsilon'/2})^j \\
&\leq \sum_{n\geq m} n^{-3} + 12\, d_{k,\lambda}\, 2^{-\epsilon' \lfloor \log_2 m \rfloor /2} / (1-2^{-\epsilon'/2}) \\
&= O(m^{-\epsilon'/2}). \qquad (20)
\end{aligned}
$$

Taking $m \to \infty$, we obtain (14). This completes the proof of Theorem 1 for the case when P has density function $\rho_{k,\lambda}$.

4 Proof of Theorem 1

The structure of the proof for the general case is similar to that for the special case $\rho_{k,\lambda}$ (as described in the beginning of Section 3). The details involve additional ideas in order to carry out estimates without relying on an explicit form for P.

By assumption P is a probability distribution over R^k whose absolute μ-th moments are all finite where $\mu > k/(k-1)$; f is the probability density for the absolute continuous part of P. Let $A_{P,\mu}$ denote $\int_{R^k} \|v\|^\mu dP$, which is clearly finite.

We will use notations similar to the ones employed in the previous section. Choose any fixed λ such that $\mu + k < \lambda < k\,\mu$. (Such λ exists as $\mu > k/(k-1)$.) Define the constants $c_{k,\lambda}$, $d_{k,\lambda}$, and ϵ by the same formulas as in Section 3; also let $C_{k,\ell} = [-\ell, \ell]^k$ and $\ell_n = n^\epsilon$ as before.

Let $v_1, v_2, \cdots$ be an infinite sequence of random points in R^k, where each v_i is independently distributed according to P. Define $V_n = \{v_1, v_2, \cdots, v_n\}$ and $X_n = T(V_n)$. For any $\ell > 0$, let $Y_{n,\ell} = T_-(V_n; \ell)$ and $Z_{n,\ell} = T_+(V_n; \ell)$.

We first prove the analogous results for Lemmas 1-4.

Lemma 8 $Y_{n,\ell} \leq X_n \leq Y_{n,\ell} + Z_{n,\ell} + 2k^{1/2}\ell$.

Corollary If $0 < \ell \leq \ell'$, then $Y_{n,\ell} \leq Y_{n,\ell'} \leq Y_{n,\ell} + Z_{n,\ell} + 2k^{1/2}\ell$.

Proof The same proof as Lemma 1. □

Let $h_{P,\mu,\lambda} = 2^{\lambda/k}\, k^{-1/2}\, \Gamma(1/k)\, (c_{k.\lambda})^{-1/k}\, A_{P,\mu}$.

Lemma 9

$$
\mathbf{E} Z_{n,\ell} \leq 2\, h_{P,\mu,\lambda}\, n^{(k-1)/k}\, \ell^{-(\mu-\lambda/k)} + 2\, d_{k,\lambda} n^{(k-1)/k}\, \ell^{-(\lambda-k-\lambda/k)} + 8\, b_k\, \ell\, n^{(k-2)/(k-1)}\,.
$$

Proof Let $v_1', v_2', \cdots$ be an infinite sequence of random points in R^k, where each v_i' is independently distributed according to $\rho_{k,\lambda}$. Let $V_n' = \{v_1', v_2', \cdots, v_n'\}$.

For each $u \in R^k \setminus C_{k,\ell}$, let J_u denote the point closest to u among all points u' such that (a) $\|u'\| < \|u\|$, and (b) u' lie in either $V_n' \setminus C_{k,\ell}$ or on the faces of the cube $C_{k,\ell}$. For each $1 \leq i \leq n$, define a random variable D_i which takes on the value 0 if $v_i \in C_{k,\ell}$ and, otherewise, the value $\|v_i - J_{v_i}\|$. Similarly, for each $1 \leq i \leq n$, define a random variable D_i' which takes on the value 0 if $v_i' \in C_{k,\ell}$ and, otherwise, the value $\|v_i' - J_{v_i'}\|$.

Let $S_+ = (V_n \setminus C_{k,\ell}) \cup \{e_\ell\}$. Observe that a Steiner tree for S_+ can be obtained by adding the sets of edges $\{v_i, J_{v_i}\}$, $v_i \in V_n \setminus C_{k,\ell}$, and $\{v_i', J_{v_i'}\}$, $v_i' \in V_n' \setminus C_{k,\ell}$ to a shortest traveling salesman's tour for the points in J, where $J = \{J_{v_i} \mid v_i \in V_n \setminus C_{k,\ell}\,;\, J_{v_i} \in C_{k,\ell}\} \cup \{J_{v_i'} \mid v_i' \in V_n' \setminus C_{k,\ell}\,;\, J_{v_i'} \in C_{k,\ell}\} \cup \{e_\ell\}$. By Facts 3 and 4, we have

$$\begin{aligned} Z_{n,\ell} &\leq 2M(S_+) \\ &\leq 2 \sum_{1\leq i\leq n} D_i + 2 \sum_{1\leq i\leq n} D_i' + 2T(J) \\ &\leq 2 \sum_{1\leq i\leq n} D_i + 2 \sum_{1\leq i\leq n} D_i' + 2\, b_k\, \ell(2n+1)^{(k-1)/k}. \end{aligned}$$

Thus,

$$\mathbf{E}Z_{n,\ell} \leq 2 \sum_{1\leq i\leq n} \mathbf{E}D_i + 2 \sum_{1\leq i\leq n} \mathbf{E}D_i' + 8\, b_k\, \ell\, n^{(k-1)/k}. \tag{21}$$

The quantity $\mathbf{E}D_i'$ has been analyzed in the proof of Lemma 2, and we have from (4)

$$\mathbf{E}D_i' \leq d_{k,\lambda} n^{-1/k}\, \ell^{-(\lambda-k-\lambda/k)}. \tag{22}$$

We now estimate $\mathbf{E}D_i$. We start as in the analysis of $\mathbf{E}D_i'$. Let $Q_\ell = R^k \setminus C_{k,\ell}$. For $u \in Q_\ell$ the conditional probability distribution of D_i satisfies, for all $s \geq 0$

$$\begin{aligned} \Pr\{\, D_i > s \mid v_i = u \,\} &\leq (1 - \rho_{k,\lambda}(u)\,(sk^{-1/2})^k)^n \\ &\leq \exp(-n\, k^{-k/2}\, c_{k,\lambda}\, s^k\, (1+\|u\|)^{-\lambda}). \end{aligned} \tag{23}$$

This leads to

$$\begin{aligned} \mathbf{E}(D_i \mid v_i = u) &= \int_0^\infty \Pr\{\, D_i > s \mid v_i = u \,\} ds \\ &\leq k^{-1/2}\, \Gamma(1/k)\, (c_{k,\lambda}\, n)^{-1/k}\, (1+\|u\|)^{\lambda/k}. \end{aligned} \tag{24}$$

Using (24) we obtain

$$\mathbf{E}D_i = \int_{Q_\ell} \mathbf{E}(D_i \mid v_i = u) dP$$

$$\leq \; k^{-1/2}\,\Gamma(1/k)\,(c_{k,\lambda}\,n)^{-1/k} \int_{Q_\ell} (1+\|u\|)^{\lambda/k}\,dP$$
$$\leq \; 2^{\lambda/k}\,k^{-1/2}\,\Gamma(1/k)\,(c_{k,\lambda}\,n)^{-1/k} \int_{Q_\ell} \|u\|^{\lambda/k}\,dP$$

Now,

$$\begin{aligned}\int_{Q_\ell} \|v\|^{\lambda/k} dP &= \int_{Q_\ell} \|v\|^{\mu}\|v\|^{\lambda/k-\mu} dP \\ &\leq \ell^{-(\mu-\lambda/k)} \int_{Q_\ell} \|v\|^{\mu} dP \\ &\leq A_{P,\mu}\,\ell^{-(\mu-\lambda/k)}\,.\end{aligned}$$

Therefore,

$$\mathbf{E}D_i \leq 2^{\lambda/k}\,k^{-1/2}\,\Gamma(1/k)\,(c_{k.\lambda})^{-1/k}\,A_{P,\mu}\,n^{-1/k}\,\ell^{-(\mu-\lambda/k)}\,. \tag{25}$$

Lemma 9 follows from (21), (22) and (25). □

Lemma 10 If $1 < \ell < n^{1/12}/10$ and $s > 6\,b_k\,n^{(k-1)/k}\,(\ell^{-1/2} + n^{1/12}\,e^{-n^{1/2}}) + 6\,k^{1/2}\,\ell$, then

$$\Pr\{\,|Y_{n,\ell} - \mathbf{E}Y_{n,\ell}| > s\,\} \leq 2\,e^{-s^2/\nu} + 2e^{-2n/\ell^6},$$

where $\nu = 2^{11}k(n^{(k-2)/k}\ell^{2+24/k} + 2^{k+2}n^{(k-1)/k}\ell^{-7+12/k} + 2^k n\ell^{-7})$.

Proof Let $m = \lceil (n\ell^{-12})^{1/k} \rceil$ and $\Delta = 2\ell/m$. Divide the cube $C_{k,\ell}$ into m^k subcubes $C_1, C_2, \cdots, C_{m^k}$, each of side-length Δ. For each $1 \leq i \leq m^k$, let $\xi_i = \int_{C_i} dP$. Let I be the set of $1 \leq i \leq m^k$ satisfying $\xi_i m^k \geq \ell^{-3}$, and let $\bar{I} = \{1, 2, \cdots, m^k\} \setminus I$. Define $C' = \cup_{i \in I} C_i$ and $C'' = \cup_{i \in \bar{I}} C_i$.

Intuitively, few points falling into the cubes in C'' and thus can be ignored. For points falling into the cubes in C', each such point is likely to have neighbors within the cube containing the point, which enables us to use Azuma's inequality to obtain a strong error bound as required by the lemma.

Let $V'_n = V_n \cap C'$ and $V''_n = V_n \cap C''$. Let $Y'_{n,\ell} = T(V'_n)$ and $Y''_{n,\ell} = T(V''_n)$. The same argument as used in the proof of Lemma 1 gives

$$|Y_{n,\ell} - Y'_{n,\ell}| \leq Y''_{n,\ell} + 2k^{1/2}\ell\,. \tag{26}$$

Note also that, obviously,

$$|Y_{n,\ell} - \mathbf{E}Y_{n,\ell}| \leq |Y_{n,\ell} - Y'_{n,\ell}| + |Y'_{n,\ell} - \mathbf{E}Y'_{n,\ell}| + |\mathbf{E}Y'_{n,\ell} - \mathbf{E}Y_{n,\ell}|\,. \tag{27}$$

In view of (26) and (27), our plan is to derive bounds on the behavior of $|Y_{n,\ell} - Y'_{n,\ell}|$ by examining the behavior of $Y''_{n,\ell}$, and to derive bounds on $|Y'_{n,\ell} - \mathbf{E}Y'_{n,\ell}|$ by using Azuma's inequality.

Let $\xi = \sum_{i \in \bar{I}} \xi_i$. Then

$$\begin{aligned} \xi &\le m^k(\ell^{-3}m^{-k}) \\ &= \ell^{-3}. \end{aligned}$$

Let N denote the cardinality of V_n''. As each v_i has a probability ξ to fall into C'', it follows that

$$\begin{aligned} \mathbf{E}N &= \xi n \\ &\le n/\ell^3. \end{aligned}$$

Also, by Chernoff's bound

$$\begin{aligned} \Pr\{N > 2n/\ell^3\} &\le \Pr\{N - \mathbf{E}N > n/\ell^3\} \\ &\le 2\exp(-2n/\ell^6). \end{aligned}$$

Now, $Y''_{n,\ell} \le b_k \, \ell \, N^{(k-1)/k}$ by Fact 4. It follows that

$$\begin{aligned} \Pr\{Y''_{n,\ell} > 2b_k n^{(k-1)/k}\ell^{-2+3/k}\} &\le \Pr\{N > 2n/\ell^3\} \\ &\le 2\exp(-2n/\ell^6). \end{aligned} \tag{28}$$

Using Fact 4 and the above inequality, we have

$$\begin{aligned} \mathbf{E}Y''_{n,\ell} &\le 2b_k n^{(k-1)/k}\ell^{-2+3/k} + 2\exp(-2n/\ell^6) b_k \ell n^{(k-1)/k} \\ &\le 2b_k n^{(k-1)/k}(\ell^{-1/2} + n^{1/12}e^{-n^{1/2}}). \end{aligned} \tag{29}$$

It follows from (26), (28), and (29) that

$$\begin{aligned} |\mathbf{E}Y'_{n,\ell} - \mathbf{E}Y_{n,\ell}| &\le 2b_k n^{(k-1)/k}(\ell^{-1/2} + n^{1/12}e^{-n^{1/2}}) + 2k^{1/2}\ell \\ &< s/3, \end{aligned} \tag{30}$$

and

$$\begin{aligned} \Pr\{|Y_{n,\ell} - Y'_{n,\ell}| > s/3\} &\le \Pr\{Y''_{n,\ell} > s/3 - 2k^{1/2}\ell\} \\ &\le \Pr\{Y''_{n,\ell} > 2b_k n^{(k-1)/k}\ell^{-2+3/k}\} \\ &\le 2\exp(-2n/\ell^6). \end{aligned} \tag{31}$$

From (27), (30), and (31), we obtain

$$\Pr\{|Y_{n,\ell} - \mathbf{E}Y_{n,\ell}| > s\} \le 2\exp(-2n/\ell^6) + \Pr\{|Y'_{n,\ell} - \mathbf{E}Y'_{n,\ell}| > s/3\}. \tag{32}$$

It remains to estimate $\Pr\{|Y'_{n,\ell} - \mathbf{E}Y'_{n,\ell}| > s/3\}$.

As in the proof of Lemma 3, consider the martingale differences $g_i = \mathbf{E}(Y''_{n,\ell}|\sigma_i) - \mathbf{E}(Y''_{n,\ell}|\sigma_{i-1})$ for $1 \le i \le n$, where σ_i is the σ-algebra generated by $v_1, v_2, \cdots, v_i$. Let $w_i = \sup |g_i|$. Using Azuma's inequality, we obtain

$$\Pr\{|Y'_{n,\ell} - \mathbf{E}Y'_{n,\ell}| > s/3\} \le 2\exp(-s^2/(18 \sum_{1\le i\le n} w_i^2)). \tag{33}$$

To evaluate w_i, define for each $u \in C_{k,\ell}$ the random variable A_u whose value is $2k^{1/2}\ell$ if none of the v_j, $i+1 \le j \le n$, is in C', and otherwise is equal to $\min\{\|u - v_j\| \mid i+1 \le j \le n, v_j \in C'\}$. Then

$$w_i \le 4 \sup_{u\in C'} \mathbf{E}A_u. \tag{34}$$

Suppose $u \in C_t$ where $t \in I$. By definition of I, $\xi_t \ge \ell^{-3}m^{-k} \ge \ell^{-3}2^{-k}n^{-1}\ell^{12} = z/n$, where $z = 2^{-k}\ell^9$. The probability that none of v_j, $i+1 \le j \le n$, falls into C_t is at most

$$\begin{aligned}(1-\xi_t)^{n-i} &\le e^{-\xi_t(n-i)}\\ &\le e^{-z(n-i)/n}.\end{aligned}$$

This implies

$$\begin{aligned}\mathbf{E}A_u &\le k^{1/2}\Delta + 2k^{1/2}\ell e^{-z(n-i)/n}\\ &\le 2k^{1/2}n^{-1/k}\ell^{1+12/k} + 2k^{1/2}\ell e^{-z(n-i)/n}.\end{aligned}$$

Hence, by (34), we obtain

$$w_i \le 8k^{1/2}(n^{-1/k}\ell^{1+12/k} + \ell e^{-z}e^{zi/n}).$$

Standard manipulation leads to

$$\sum_{1\le i\le n} w_i^2 \le 64k(n^{(k-2)/k}\ell^{2+24/k} + 2^{k+2}n^{(k-1)/k}\ell^{-7+12/k} + 2^k n\ell^{-7}). \tag{35}$$

Lemma 10 follows from (32), (33), and (35). □

Lemma 11 For any fixed $t > 0$, $\lim_{n\to\infty} \mathbf{E}Y_{n,t}/n^{(k-1)/k} = \beta_k\, k^{1/2} \int_{C_{k,t}} (f(v))^{(k-1)/k}\, dv$.

Proof Let $\theta = \int_{C_{k,t}} dP$. We can assume that $\theta > 0$; otherwise the lemma is trivially true. Let P_t denote the probability distribution when P is restricted to the cube $C_{k,t}$. Clearly, P_t has an absolute continuous part with density $\theta^{-1} f$.

Let $u_1, u_2, \cdots$ be an infinite sequence of independent random points each distributed according to P_t, and let $W_{n,t}$ denote the length of the shortest traveling salesman's tour through the first n points. Let $\gamma_{P,t} = \beta_k k^{1/2} \int_{C_{k,t}} (f(v))^{(k-1)/k} dv$. The result of [BHH] implies that

$$\lim_{n\to\infty} \mathbf{E}W_{n,t}/n^{(k-1)/k} = \theta^{-(k-1)/k} \gamma_{f,t}. \tag{36}$$

Consider n independent random points in R^k, each distributed according to P, and let $p_{n,n'}$ be the probability that n' of them fall into $C_{k,t}$. Then

$$\mathbf{E}Y_{n,t} = \sum_{n'} p_{n,n'} \mathbf{E}W_{n',t}. \tag{37}$$

Lemma 11 can be obtained from (36), (37) and Chernoff's bound on $p_{n,n'}$ in the same way as in the proof of Lemma 4. □

We now complete the proof of Theorem 1 using the preceding lemmas in essentially the same way as in Section 3. We will only sketch the proof.

Let $\ell_n = n^\epsilon$ and $s_n = n^{(k-1)/k-\epsilon/4}$. The following lemmas (analogous to Lemmas 5-7) can be proved.

Lemma 12 $\lim_{n\to\infty} \mathbf{E}X_n/n^{(k-1)/k} = \lim_{n\to\infty} \mathbf{E}Y_{n,\ell_n}/n^{(k-1)/k} = \alpha_P$.

Lemma 13 $\mathbf{E}Z_{n,\ell_m} \leq (4\, h_{P,\mu,\lambda} + 1)\, n^{(k-1)/k-\epsilon(\mu-\lambda/k)}$ for all sufficiently large n and $n/2 \leq m < n$.

Lemma 14 $\Pr\{\, |Y_{n,\ell_n} - \mathbf{E}Y_{n,\ell_n}| > n^{(k-1)/k-\epsilon/4}\} \leq e^{-n^\epsilon}$ for all sufficiently large n.

Lemma 12 proves $\lim_{n\to\infty} \mathbf{E}X_n/n^{(k-1)/k} = \alpha_P$. To prove the other equation in the theorem, let $\delta > 0$ be any fixed number. Let B_n be the event that $|X_n/n^{(k-1)/k} - \alpha_P| > \delta$. We need to prove that

$$\lim_{m\to\infty} \Pr\{\cup_{n\geq m} B_n\} = 0. \tag{38}$$

Let F_n be the event that $|Y_{n,\ell_n} - \mathbf{E}Y_{n,\ell_n}| > n^{(k-1)/k-\epsilon/4}$, and G_n be the event that $Z_{n,\ell_n} > n^{(k-1)/k-\epsilon(\mu-\lambda/k)/2}$. With the help of Lemmas 12-14, inequalities (16)-(20) can be derived in the same way as in Section 3, leading to the proof of (38). This completes the proof of Theorem 1.

5 Extensions

In this section we extend Theorem 1 to include a general class of functionals. Following Steele [S1], a *Euclidean functional* L in R^k is a real-valued function of the finite subsets of R^k, such that $L(\emptyset) = 0$, $L(cv_1, cv_2, \cdots, cv_m) = cL(v_1, v_2, \cdots, v_m)$, $L(v_1+u, v_2+u, \cdots, v_m+$

$u) = L(v_1, v_2, \cdots, v_m)$ for all $m \geq 1$, $c > 0$ and $v_i, y \in R^k$, and $L(S) \leq L(S')$ for all finite subsets $S \subseteq S' \subseteq R^k$.

The length of the shortest traveling salesman's tours, minimum Steiner trees, and minimum spanning trees are familiar examples of Euclidean functionals. Steele [S1] showed that the Beardwood-Halton-Hammersley theorem [BHH] holds for a broad class of Euclidean functionals satisfying certain conditions. We show that Theorem 1 is true when one more condition ((T2) below) is imposed.

Let $\tau(S, S')$ denote the minimum Euclidean distance between any point in S and any point in S', when S and S' are nonempty sets in R^k. Consider the following properties.

(T1) For any finite collection of disjoint bounded cubes $Q_1, Q_2, \cdots, Q_s$ with edges parallel to the axes, and for any infinite sequence of points $v_1, v_2, \cdots$ in R^k, one has for large n,

$$\sum_{1 \leq i \leq s} L(V_n \cap Q_i) \leq L\left(V_n \cap (\cup_{1 \leq i \leq s} Q_i)\right) + o(n^{(k-1)/k}),$$

where $V_n = \{v_1, v_2, \cdots, v_n\}$.

(T2) For some positive constant η, $L(S \cup S') \leq L(S) + L(S') + \eta \cdot \tau(S, S')$ for all nonempty finite sets S and S'.

Theorem 2 Let $k > 1$ and L be any Euclidean functional in R^k that satisfies properties (T1) and (T2). Then there exists a positive constant β_L such that, for any probability distribution over R^k with all its absolute μ-th moments being finite for some $\mu > k/(k-1)$, the following is true: if $v_1, v_2, \cdots$ is an infinite sequence of independent random points each distributed according to P, then $L_n/n^{(k-1)/k}$ converges to $\alpha_{L,P}$ almost surely and $\lim_{n \to \infty} \mathbf{E} L_n/n^{(k-1)/k} = \alpha_{L,P}$, where L_n is the random variable $L(v_1, v_2, \cdots, v_n)$ and $\alpha_{L,P} = \beta_L \int_{R^k} (f(v))^{(k-1)/k} dv$ with f being any density function for the absolute continous part of P.

Note that (T1) was called the *upper-linearity* property in [S1]. The Euclidean functionals corresponding the traveling salesman's tour, minimum Steiner tree, minimum spanning tree satisfy (T1) as was demonstrated in [S1]. It is easy to see that (T2) is also true for these three examples.

We make two further observations. First, it is not hard to show that any Euclidean functional obeying (T1) and (T2) automatically satisfies all the properties (A1)-(A8) listed in Steele [S1], which implies by the results in [S1] that Theorem 2 is valid for any probability distribution P with a *bounded* support.

Second, it is easy to see that, if L satisfies (T2), then $L(S)$ is no greater than $\max\{1, \eta\}$ times the length of the minimum spanning tree on S. It follows that $L(S) \leq \eta' M(S)$

for some positive constant η', where $M(S)$ is the length of the minimum Steiner tree connecting all the points in S. (Fact 3 is a special case when L is T, the functional corresponding to the shortest traveling salesman's tour.)

With the help of the above two observations, the proof of Theorem 2 is basically the same as Theorem 1, and will not be repeated here.

6 Concluding Remarks

It would be of interest to investigate the asymptotic behavior of functionals not covered by the results in this paper. For example, what can be said of the optimum matching of points with distributions over unbounded regions? (See Papadimitriou [P] for discussions of the optimum matching question over bounded regions.) Are there natural functionals whose behavior over bounded regions, say the unit cube, is different from $n^{(k-1)/k}$? How do they behave?

We have proved that a subclass of the Euclidean functionals discussed in Steele [S1] have the Beardwood-Halton-Hammersley limiting behavior even when the underlying distribution P has unbounded support. Can this result be extended to cover the entire class discussed in [S1] (i.e. replace (T1) and (T2) by properties (A1)-(A8) listed in [S1])?

References

[A] K. Azuma, "Weighted sums of certain dependent random variables," *Tôhoku Mathematical Journal* **19** (1967), 359-367.

[BHH] J. Beardwood, J.H. Halton, and J.M. Hammersley, "The shortest path through many points," *Proceedings of the Cambridge Philosophical Society* **55** (1959), 299-327.

[C] H. Chernoff, "A measure of asymptotic efficiencey for tests of a hypothesis based on the sum of observables," *Annals of Mathematical Statistics* **23** (1952), 493-507.

[HLP] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, second edition, Cambridge University Press, Cambridge, UK, 1952.

[H] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of American Statistics Association* **58** (1963), 13-30.

[K] R.M. Karp, "Probabilistic analysis of partitioning algorithms for the traveling salesman problem in the plane," *Mathematics of Operations Research* **2** (1977), 209-224.

[P] C.H. Papadimitriou, "The probabilistic analysis of matching heuristics," *Proceedings of the 15th Annual Conference on Communication, Control, and Computing*, Urbana, Illinois, 1978.

[RT] W.T. Rhee and M. Talagrand, "Martingale inequalities and NP-complete problems," *Mathematics of Operations Research* **12** (1987), 177-181.

[Si] C.L. Siegel, *Lectures on the Geometry of Numbers*, Springer-Verlag, Berlin, 1989.

[S1] J.M. Steele, "Subadditive Euclidean functionals and nonlinear growth in geometric probability," *The Annals of Probability* **9** (1981), 365-376.

[S2] J.M. Steele, "Seedlings in the theory of shortest paths," in *Disorder in Physical Systems: A Volume in Honour of John M. Hammersley*, edited by G.R. Grimmett and D.J.A. Welsh, Oxford University Press, Oxford, 1990, 277-306.