J. Appl. Prob. **24**, 809–826 (1987)
Printed in Israel

ON THE NUMBER OF LEAVES OF A EUCLIDEAN MINIMAL SPANNING TREE

J. MICHAEL STEELE,* *Princeton University*
LAWRENCE A. SHEPP,** *AT&T Bell Laboratories*
WILLIAM F. EDDY,*** *Carnegie-Mellon University*

Abstract

Let $V_{k,n}$ be the number of vertices of degree k in the Euclidean minimal spanning tree of X_i, $1 \leqq i \leqq n$, where the X_i are independent, absolutely continuous random variables with values in R^d. It is proved that $n^{-1}V_{k,n}$ converges with probability 1 to a constant $\alpha_{k,d}$. Intermediate results provide information about how the vertex degrees of a minimal spanning tree change as points are added or deleted, about the decomposition of minimal spanning trees into probabilistically similar trees, and about the mean and variance of $V_{k,n}$.

SUBADDITIVE EUCLIDEAN FUNCTIONALS; VERTEX DEGREES; SELF-SIMILAR PROCESSES; EFRON–STEIN INEQUALITY

1. Introduction

The initial purpose of this investigation was to provide an asymptotic understanding of the number of leaves of the minimal spanning tree of n points chosen at random from the unit square. To set this problem precisely, as well as to make clear the more general results which will be established, we begin with some notation.

By $G = (V, E)$ we denote a connected graph with vertex set V and edge set E. We also assume that there is a *weight function* $w : E \to R$ which assigns a real number to each edge in E. A *minimal spanning tree* T of G is a connected graph with vertex set V and edge set $E' \subset E$ such that

$$\sum_{e \in E'} w(e) = w(T)$$

Received 9 May 1984; revision received 2 October 1986.

* Postal address: Program in Engineering Statistics, Princeton University, Princeton, NJ 08544, USA.

Research supported in part by National Science Foundation Grant DMS-8414069.

** Postal address: AT&T Bell Laboratories, Murray Hill, NJ 07974, USA.

*** Postal address: Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA 15213, USA.

is minimal, i.e. no connected subgraph of G has a smaller total edge weight than $T = (V, E')$.

One probabilistic model of interest here is specified by taking $V = \{X_1, X_2, \cdots, X_n\}$ where X_i, $1 \leqq i \leqq n$, are independent random variables with the uniform distribution $[0, 1]^2$, letting E consist of all $\binom{n}{2}$ pairs (i, j), $1 \leqq i < j \leqq n$, and setting $w(x_i, x_j) = |x_i - x_j|$, where the bars denote the usual Euclidean distance. We denote by $\mathrm{MST}(X_1, X_2, \cdots, X_n)$ a minimal spanning tree of $X_1, X_2, \cdots, X_n$ under the weight function w. As this language acknowledges, there can be more than one minimal spanning tree for a given sample $\{x_1, x_2, \cdots, x_n\}$. However, for X_i with continuous support, the minimal spanning tree is unique with probability 1, since any two sets of $n - 1$ edges will have different total lengths with probability 1.

Our main focus will be on the number of leaves (i.e., vertices of degree 1) in the minimal spanning tree determined by the uniform sample model. We also consider the more general problem of general vertex degrees of random vertices X_i, $1 \leqq i \leqq n$, which are chosen according to an arbitrary density with support in R^d, $d \geqq 2$. Still, the combinatorial and probabilistic essence of our problem is already to be found in the case of uniformly distributed random variables in the unit square.

The motivation for our investigation comes from several sources. First, there has already been interesting work done on the number of leaves of a random tree where the random tree is less intricate than the Euclidean minimal spanning tree. Second, there is a close connection with previous work on subadditive Euclidean functionals. After introducing the notation for our more general results, it will be possible to give an explicit application of our main theorem to the theory of multivariate non-parametric hypothesis tests.

The earliest work on the number of leaves of a random tree is due to Rényi (1959), where the specified model was that of choosing a tree T at random uniformly from the set of all n^{n-2} labeled trees with n vertices. If L_n denotes the number of leaves of such a tree, Rényi (1959) established that

$$EL_n \sim n/e \quad \text{as } n \to \infty \tag{1.1}$$

and also that L_n is asymptotically normal with approximate variance $n(e-2)e^{-2}$.

Najock and Heyde (1982) also established a linear growth rate for the expected number of leaves in the context of a certain class of $(n-1)!$ rooted trees which arise in a linguistic application. Under their model for a random tree, they proved that

$$\lim_{n \to \infty} n^{-1} EL_n = \tfrac{1}{2}. \tag{1.2}$$

In contrast to the tradition of Rényi (1959) and Najock and Heyde (1982), the investigation of minimal spanning trees has a strong geometric flavor. Also,

the successful analytical treatment of such Euclidean trees rests on articulating an approximate self-similarity between a subset of the tree and the whole tree. This method is closely related to subadditive techniques for Euclidean functionals, and the present application to the *number* of leaves also extends the domain of problems to which subadditive methods are effective. Earlier work with the type of subadditive method used here always seemed to deal with the *lengths* determined by optimality conditions. (See e.g. Beardwood et al. (1959), Steele (1981a,b), or Hochbaum and Steele (1982)). For some understanding of the applications of minimal spanning trees in computer science, one can consult Bentley (1978), Chin (1978), Jung (1974), Kang (1977), Katajainen (1983), and Whitney (1972). For some examples of applications of minimal spanning trees in physical sciences and biology, one can consult Mallion (1975), Penny (1980), Romane (1977), and Wu (1977).

Three main results will be proved here, but the following readily digested theorem motivates and underlies the whole development.

Theorem 1. If X_i, $1 \leqq i < \infty$, are independent and uniformly distributed on $[0, 1]^2$, then for L_n, the number of leaves of the MST$(X_1, X_2, \cdots, X_n)$, we have with probability 1,

$$\lim_{n\to\infty} n^{-1} L_n = \alpha > 0. \tag{1.3}$$

The exact value of the constant α is not known, but Monte Carlo simulation results suggest that $\alpha = 2/9$ is a reasonable approximation. We will provide a proof that $\alpha > 0$, but the only analytical bounds we can provide on α are very conservative. Further comments on the possibility that α exactly equals 2/9 are given in Section 6.

The conclusion of Theorem 1 can be extended to arbitrary dimension and arbitrary vertex degree. Naturally, the associated constants depend upon the dimension d and on the vertex degree.

Theorem 2. Let X_i, $1 \leqq i < \infty$, be independent and uniformly distributed on $[0, 1]^d$, $d \geqq 2$. If $V_{k,n} = V_k(X_1, X_2, \cdots, X_n)$ denotes the number of vertices of degree k in the MST$(X_1, X_2, \cdots, X_n)$, then there are constants $\alpha_{k,d}$ such that with probability 1

$$\lim_{n\to\infty} n^{-1} V_{k,n} = \alpha_{k,d} \tag{1.4}$$

for all k.

In the case of $k = 1$ and $d = 2$, one has $V_{1,n} = L_n$ and the constant $\alpha_{1,2}$ is simply α, the constant of Theorem 1. One should note also that for each dimension d there is an integer D_d such that D_d is the largest possible degree of any vertex of any minimal spanning tree in R^d so that $\alpha_{k,d} = 0$ for $k > D_d$. It is easy to see that $D_d < \infty$ for any d, but only a few values of D_d are known

exactly. For proofs that $D_2 = 6$ and $D_3 = 13$, one can consult Hilbert and Cohn-Vossen (1952), p. 47, or Rogers (1964).

We are able to show $\alpha_{1,d} > 0$ for all $d \geqq 2$; but, since even the determination of D_d is open for large d, it is unreasonable to expect any determination of $\alpha_{k,d}$ for large k and d.

The third theorem shows that one can relax the assumption that the X_i, $1 \leqq i \leqq n$, are uniformly distributed. Curiously, it does not seem possible to extend our results to completely arbitrary distributions. In particular, considerable mystery remains about the number of leaves of minimal spanning tree samples from a distribution with singular support. The difficulty of dealing with the contributions due to a singular component of the distribution of the X_i is peculiar to the problem of vertex degrees and has not arisen in the earlier work on subadditive Euclidean functionals.

One interesting aspect of the limit theory of $V_{k,n}$ for absolutely continuous random variables with compact support is that the limit of $n^{-1}V_{k,n}$ does not depend on the underlying distribution. In particular, we have the following result.

Theorem 3. If X_i, $1 \leqq i < \infty$ are i.i.d. with a density f in R^d, then with probability 1

$$\lim_{n\to\infty} n^{-1}V_{k,n} = \alpha_{k,d} \qquad \text{for all } k \geqq 1, d \geqq 2. \tag{1.5}$$

Before digging into the proof of these results, it seems worth giving more details on application of Theorem 3 to the theory of non-parametric multivariate tests. In particular, we consider the work of Friedman and Rafsky (1979) which gives an elegant extension of the Wald–Wolfowitz runs test to multivariate samples by using the minimal spanning tree to suggest a proper analogue for the number of runs in a sequence. If one is given two samples of sizes n and m from two distributions F_x and F_y in R^d, one can test $H_0 : F_x = F_y$ by constructing the minimal spanning tree of the joint sample, removing all the edges of the tree which join observations from different samples, and letting R denote the resulting number of subtrees. Small values of R would lead to the rejection of H_0. The expectation of R does not depend on the degrees, and equals $1 + 2nm/(n+m)$ just as in the classical test of Wald and Wolfowitz. In the general case, Friedman and Rafsky found the variance of R *conditionally* on C, the number of edge pairs that share a common node. If we let d_i be the degree of the ith vertex of the MST of the joint sample, then since $\frac{1}{2}\sum_{i=1}^{n+m} d_i = n + m - 1$ we have under H_0 that

$$C = \frac{1}{2}\sum_{i=1}^{n+m} d_i(d_i - 1) = 1 - n - m + \frac{1}{2}\sum_k k^2 V_{n+m,k}. \tag{1.6}$$

Consequently, the results given here show that the conditional tests of

Friedman and Rafsky are asymptotically unconditional under H_0. The complete details of this fact would take us away from the main issues, still Equations (1.5) and (1.6) should make the claim credible.

The proof of preceding theorems will be given in the next four sections. In Section 2 we collect some basic combinatorial observations. Section 3 proceeds to obtain the behavior of the expectations $EV_{k,n}$ by using a Poisson embedding, subadditivity arguments, and the application of a Tauberian theorem.

The fourth section uses the Efron–Stein inequality to obtain a good bound on the variances of $V_{k,n}$. Traditional Chebyshev and interpolation techniques are then used to complete the proof of Theorem 1 and Theorem 2. Section 5 shows how a lemma due to Strassen (1965) can be used to extend Theorem 2 first to densities with compact support, then to general densities.

Section 6 collects the information which is available on the constants $\alpha_{k,d}$. The final section reviews some open problems and provides perspective on the way the asymptotic theory of vertex degrees fits into the theory of subadditive functionals and relates to the idea of fractals.

2. Combinatorial observations

The next four lemmas are free of any probability theory. They focus on the basic combinatorial geometry of minimal spanning trees.

Lemma 2.1. If B is a subset of $S = \{x_1, x_2, \cdots, x_n\}$, and if $e = \{X_i, X_j\}$ is an edge of MST(S) such that $e \subset B$, then $e \in \mathrm{MST}(B)$.

Proof. Since $e \in \mathrm{MST}(S)$, there is a partition (A, A^c) of S such that $e \cap A \neq \varnothing$, $e \cap A^c \neq \varnothing$ and

$$|e| = \min\{|e'| : e' \cap A \neq \varnothing \text{ and } e' \cap A^c \neq \varnothing\}. \tag{2.1}$$

(This is a traditional necessary and sufficient condition for an edge to be in a MST.)

Since $e \subset B$, we see $A \cap B$ and $A^c \cap B$ is a non-trival partition of B. Because of (2.1) the edge e will then satisfy

$$|e| \leqq \min\{|e'| : e' \cap A \cap B \neq \varnothing \text{ and } e' \cap A^c \cap B \neq \varnothing\}. \tag{2.2}$$

By the cited necessary and sufficient condition for an edge to be in the minimal spanning tree of B, we see that (2.2) guarantees $e \in \mathrm{MST}(B)$.

One important consequence of Lemma 2.1 is that it shows that only a few of the elements of $\{x_1, x_2, \cdots, x_n\}$ can have different degrees in the MST$(x_1, x_2, \cdots, x_n, x_{n+1})$ than they have in the smaller tree MST$(x_1, x_2, \cdots, x_n)$. This is made quantitative as follows.

Lemma 2.2. For any $x_i \in R^d$, $1 \leqq i \leqq n+1$, if w denotes the number of

elements of $\{x_1, x_2, \cdots, x_{n+1}\}$, which have different degrees in $\mathrm{MST}(x_1, x_2, \cdots, x_n, x_{n+1})$ and $\mathrm{MST}(x_1, x_2, \cdots, x_n)$, then $w \leqq 4D_d - 2$.

Proof. We apply Lemma 2.1 where

$$B = \{x_1, x_2, \cdots, x_n\} \subset \{x_1, x_2, \cdots, x_{n+1}\} = S$$

and observe that every edge of $\mathrm{MST}(S)$ which is not an edge of $\mathrm{MST}(B)$ must be incident to x_{n+1}. There are at most D_d such edges.

Since $\mathrm{MST}(S)$ has n edges and $\mathrm{MST}(B)$ has $n-1$ edges, we see therefore that there are at most $2D_d - 1$ edges of $\mathrm{MST}(S)$ and $\mathrm{MST}(B)$ which are not common to both $\mathrm{MST}(S)$ and $\mathrm{MST}(B)$. Since the edges which fail to be in $\mathrm{MST}(S)$ or $\mathrm{MST}(B)$ are incident to at most $4D_d - 2$ vertices, there are at most $4D_d - 2$ vertices which fail to have the same degree in $\mathrm{MST}(B)$ and $\mathrm{MST}(S)$.

One way in which Lemma 2.2 will be used is expressed in the following corollary.

Corollary. If $\{x_1, x_2, \cdots, x_n\}$ and $\{x_1', x_2', \cdots, x_m'\}$ are any two finite subsets of R^d, and if h is the cardinality of the symmetric difference of these sets, then there are at most $2h(4D_d - 2)$ vertices which have different degrees in $\mathrm{MST}\{x_1, x_2, \cdots, x_n\}$ and $\mathrm{MST}\{x_1', x_2', \cdots, x_m'\}$.

As given above, the proof of the corollary is immediate. With more thought, one can provide a bound of $h(4D_d - 1)$, but the essential point is the linearity in h.

Although we are concerned with the degrees of vertices, we will find it useful to have some information about the length of minimal spanning trees.

Lemma 2.3. There is a constant c_1 such that for any $x_i \in [0, 1]^d$, $1 \leqq i \leqq n$, there exists an i and j, $1 \leqq i < j \leqq n$, such that

$$|x_i - x_j| \leqq c_1 n^{-1/d}. \tag{2.3}$$

Proof. This is a consequence of a geometric pigeonhole argument. Suppose each of the points x_i were the center of a ball of radius r. Such balls have volume $\omega_d r^d$, where ω_d is the volume of the unit d-ball. If $x_i \in [0, 1]^d$, then the ball centered at x_i must cover at least $2^{-d}\omega_d r^d$ of $[0, 1]^d$; so, if

$$n 2^{-d} \omega_d r^d > 1, \tag{2.4}$$

then two r-balls must intersect. Thus, if there are n points in $[0, 1]^d$, we see that some pair of points must be within a distance of $2r$ if inequality (2.4) holds. In other words, if there are n points in $[0, 1]^d$, we see that some pair of points must be closer than $4\omega_d^{-1/d} n^{-1/d}$, so the lemma is established with $c_1 = 4\omega_d^{-1/d}$.

One easy consequence of the preceding lemma is that the length of a

minimal spanning tree of n points in $[0, 1]^d$ cannot be too large. More precisely, one has the following result.

Lemma 2.4. For any n points $\{x_1, x_2 \cdots, x_n\} \subset [0, 1]^d$, the minimal spanning tree has total length majorized by $c_2 n^{(d-1)/d}$.

Proof. Let m_n be the maximum length of any minimal spanning tree of n points in $[0, 1]^d$, and consider a set $\{x_1, x_2, \cdots, x_{n+1}\}$ of size $n+1$ for which m_{n+1} is attained. By Lemma 2.3 there are x_i and x_j, $1 \leqq i < j \leqq n+1$, such that $|x_i - x_j| \leqq c_1 n^{-1/d}$. Since joining x_i to x_j and forming the minimal spanning tree of $\{x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{n+1}\}$ provides a spanning tree, we have

$$m_{n+1} \leqq m_n + c_1 n^{-1/d}. \tag{2.5}$$

By summing (2.5) and using the fact that $m_1 = 0$, we see

$$m_n \leqq c_1 \sum_{k=1}^{n-1} k^{-1/d} \leqq c_1 d(d-1)^{-1} n^{(d-1)/d}, \tag{2.6}$$

which establishes the lemma with $c_2 = 2c_1$

3. Asymptotics of $EV_{k,n}$

Instead of working directly with uniformly distributed random variables in $[0, 1]^d$, we will draw strength from the independence properties of the Poisson process in R^d. The first observation is that if we let $\phi(t)$ denote the expected number of vertices of degree k of the MST of the points of the Poisson process in $[0, t]^d$, then by the usual formula for expectations in terms of conditional expectations,

$$\phi(t) = \sum_{n=0}^{\infty} (EV_{k,n}) t^{dn} \exp(-t^d)/n! \tag{3.1}$$

where $V_{k,n}$ are exactly as defined before, i.e., $V_{k,n}$ is the number of vertices of degree k in the $\mathrm{MST}(X_1, X_2, \cdots, X_n)$ where X_i are independent uniform random variables on $[0, 1]^d$.

To obtain asymptotic information about $\phi(t)$, we will first get an approximate recurrence relation for the function $\phi(t)$. To aim toward that goal, we partition $[0, t]^d$ into $M = m^d$ subcubes Q_i of side t/m. We will let Q_i' denote the points of the Poisson process which are in Q_i. By Lemma 2.1 we see that if $e \in \mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ and $e \subset Q_i'$, then $e \in \mathrm{MST}(Q_i')$. We therefore have the set inclusion

$$\mathrm{MST}\left(\bigcup_{i=1}^{M} Q_i'\right) \subset \bigcup_{i=1}^{M} \mathrm{MST}(Q_i') \cup D \tag{3.2}$$

where D is the set of all edges of $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ which intersect two different subcubes Q_i and Q_j.

We now move toward showing that the number of vertices of degree k in $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ is majorized by the number of vertices of degree k in $\bigcup_{i=1}^{M} \mathrm{MST}(Q_i')$ plus a small error term. Since $B = \bigcup_{i=1}^{M} \mathrm{MST}(Q_i')$ has fewer edges than $A = \mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$, Equation (3.2) implies that the difference set of $\bigcup_{i=1}^{M} \mathrm{MST}(Q_i')$ and $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ contains at most $2|D|$ edges, where $|D|$ is the number of edges in D. This follows from the observations that $A - B = D$ and $|B - A| = |B \cap A^c| = |B| - |B \cap A| \leqq |A| - |B \cap A| = |D|$.

The number of vertices of degree k in $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ is bounded by the number of vertices of degree k in $\bigcup_{i=1}^{M} \mathrm{MST}(Q_i')$ plus $4|D|$, since for a vertex degree to change, it must be incident to one of the edges in $A \triangle B$, there are at most $2|D|$ such edges, and each edge is incident to two vertices. Taking expectations and summarizing in terms of ϕ, we have

$$\phi(t) \leqq m^d \phi(t/m) + 4E|D|. \tag{3.3}$$

For us this will be an important inequality, but for it to be effective we need a bound on $E|D|$. We fix a real parameter $0 < \lambda < t$ and consider the decomposition $Q_i = A_i \cup B_i$, where B_i is the set of all points within λ/m of the boundary of Q_i and $A_i = B_i^c$. We note that $\bigcup_{i=1}^{M} B_i$ has volume equal to $\rho = t^d - m^d(t/m - 2\lambda/m)^d = t^d - (t - 2\lambda)^d$, and, hence, the expected number of points of the Poisson process in $\bigcup_{i=1}^{M} B_i$ is also ρ. Further, we note that if $i \neq j$, then for any $x \in A_i$ and $y \in A_j$ we have $|x - y| \geqq 2\lambda m^{-1}$.

The relations for bounding $E|D|$ are now all in place. Any edge in D must either have at least one vertex in $\bigcup_{i=1}^{M} B_i$ or else have length at least $2\lambda m^{-1}$ in order to span the distance between two distinct A_i's.

By $N(x)$ we denote the number of edges of $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ of length at least x, and we note that $xN(x)$ is bounded by the total length of the $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$. By Lemma 2.4 we see the expected length of $\mathrm{MST}(\bigcup_{i=1}^{M} Q_i')$ is bounded by $c_2 EZ^{(d-1)/d}$, where Z is a Poisson random variable with mean t_d. By Jensen's inequality $EZ^{(d-1)/d}$ is majorized by t^{d-1}, so we have

$$EN(x) \leqq c_2 t^{d-1} x^{-1}. \tag{3.4}$$

We let T denote the set of points of the Poisson process which are in $\bigcup_{i=1}^{M} B_i$, and note that every edge of D must either meet T or have length at least $2\lambda/m$. Since no vertex of a minimal spanning tree meets more than D_d vertices, we have a basic bound on $E|D|$,

$$E|D| \leqq D_d E|T| + EN(2\lambda/m) \leqq D_d m^d \{(t/m)^d - (t/m - 2\lambda/m)^d\} + c_2 t^{d-1}(2\lambda)^{-1} m.$$

Now, by simply letting $\lambda = 1$, we see there is a constant c_3 such that

$$E|D| \leqq c_3 t^{d-1} m \qquad \text{for all } t > 1. \tag{3.5}$$

Here we should note that there are choices of λ which lead to sharper inequalities than (3.4), e.g. $\lambda = t^{d/2} m^{1/2}$, but inequality (3.5) is quite simple and sharp enough. Combining the bound (3.5) with (3.3), we see that for all integers $m \geqq 1$ and real $t > 1$, we have for $c_4 = 4c_3$ that

$$\phi(t) \leqq m^d \phi(t/m) + c_4 t^{d-1} m. \tag{3.6}$$

A key step in obtaining the asymptotics of $EV_{k,n}$ will be to show that any continuous function satisfying (3.6) must satisfy the asymptotic relationship

$$\phi(t) \sim \gamma t^d \tag{3.7}$$

for some constant γ.

We now prove the relation (3.7). If we let $\psi(t) = \phi(t) t^{-d}$, then (3.6) is equivalent to

$$\psi(mt) \leqq \psi(t) + c_4 t^{-1}. \tag{3.8}$$

Just by setting $t = 2$ and letting $m \to \infty$, we see

$$0 \leqq \gamma = \liminf_{t \to \infty} \psi(t) \leqq \psi(2) + c_4 < \infty. \tag{3.9}$$

Now, for any $\varepsilon > 0$ we can find an open interval (a, b) such that $\psi(t) \leqq \gamma + \varepsilon$ for all $t \in (a, b)$, and we can also choose the value a sufficiently large so that $c_4 a^{-1} < \varepsilon$. By (3.8) this means

$$\psi(mt) \leqq \gamma + 2\varepsilon \tag{3.10}$$

for all positive integers m and all $t \in (a, b)$. If we let $I_m = (ma, mb)$, we see that (3.10) says $\psi(t) \leqq \gamma + 2\varepsilon$ for all $t \in I_m$. Now, since I_{m+1} and I_m intersect for all m such that $(m+1)a \leqq mb$, we see $\bigcup_{m=1}^{\infty} I_m$ contains the infinite interval $[t_0, \infty)$, with $t_0 = a\lceil a(b-a)^{-1} \rceil$, so

$$\psi(t) \leqq \gamma + 2\varepsilon \qquad \text{for all } t \geqq t_0. \tag{3.11}$$

Since γ was defined as the limit infimum of ψ, the fact that $\varepsilon > 0$ was arbitrary in inequality (3.11) says that $\lim_{t \to \infty} \psi(t) = \gamma$.

The asymptotic behavior of $EV_{k,n}$ can be extracted from (3.7) by means of a Tauberian argument. An easy Tauberian theorem to apply is the following.

Lemma 3.1 (Schmidt's Borel–Tauber theorem). If f is any real function defined on the integers and satisfying the Tauberian condition

$$\lim_{\varepsilon \to 0^+} \liminf_{n \to \infty} \min_{n \leqq m \leqq n + \varepsilon n^{1/2}} \{ f(m) - f(n) \} \geqq 0 \tag{3.12}$$

then the limiting relationship

$$\sum_{n=0}^{\infty} \exp(-\lambda)\lambda^n f(n)/n! \to c \quad \text{as } \lambda \to \infty \tag{3.13}$$

implies

$$f(n) \to c \quad \text{as } n \to \infty.$$

Here, of course, the notation $\lim_{\varepsilon \to 0^+} g(\varepsilon)$ means the limit of $g(\varepsilon)$ as $\varepsilon \to 0$ for $\varepsilon > 0$. This result due to Schmidt (1925) can be found with a modern probabilistic proof in Bingham (1981), cf. Corollary to Theorem 3a, p. 224.

The asymptotic relationship $\phi(t) \sim \gamma t^d$ as $t \to \infty$ can now be pressed to yield information about $EV_{k,n}$ as $n \to \infty$. In particular, we have the following result.

Lemma 3.2. There are constants $\alpha_{k,d}$ depending only on k and d such that

$$EV_{k,n} \sim \alpha_{k,d} n \quad \text{as } n \to \infty. \tag{3.14}$$

Proof. If we let $x = t^d$ in Equation (3.7) and multiply by e^x, we have

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} EV_{k,n} \sim \gamma x e^x. \tag{3.15}$$

Integration of an asymptotic series preserves that relation, so integrating (3.15) and dividing by xe^x gives

$$\sum_{n=0}^{\infty} e^{-x} \frac{x^n}{(n+1)!} EV_{k,n} \sim \gamma. \tag{3.16}$$

By taking $f(n) = (n+1)^{-1} EV_{k,n}$, we will be able to complete the proof of (3.14) by verifying the hypothesis of Lemma 3.1 given by (3.12). If $n \leqq m \leqq n + \varepsilon n^{1/2}$, we have by Lemma 2.2 that

$$|V_{k,m} - V_{k,n}| \leqq 4D_d \varepsilon n^{1/2},$$

so, we also have

$$\begin{aligned} f(m) - f(n) &= (m+1)^{-1} EV_{k,m} - (n+1)^{-1} EV_{k,n} \\ &\geqq (n + \varepsilon n^{1/2} + 1)^{-1} \{EV_{k,n} - 4D_d \varepsilon n^{1/2}\} - (n+1)^{-1} EV_{k,n} \\ &= -E(V_{k,n})(n + \varepsilon n^{1/2} + 1)^{-1}(n+1)^{-1} \varepsilon n^{1/2} \\ &\quad - 4D_d \varepsilon n^{1/2}(n + \varepsilon n^{1/2} + 1)^{-1}. \end{aligned} \tag{3.17}$$

Since $EV_{k,n} \leqq n$, the last expression in (3.11) shows without effort that

$$\liminf_{n \to \infty} \min_{n \leqq m \leqq n + \varepsilon n^{1/2}} \{f(m) - f(n)\} \geqq 0,$$

establishing (3.12). Finally, we can invoke Lemma 3.1, and thus complete the proof of the relation (3.14) with $\alpha_{k,d} = \gamma$.

4. Almost sure behavior

Now that we know $EV_{k,n} \sim \alpha_{k,d} n$, we can obtain the almost sure relation $V_{k,n} \sim \alpha_{k,d} n$ by obtaining information about the variance of $V_{k,n}$ and then using Borel–Cantelli and interpolation arguments. First we recall as a lemma the inequality of Efron and Stein (1981), which essentially says that Tukey's jackknife estimate of variance is conservative.

Lemma 4.1. If h is any symmetric funcion of $n-1$ variables and we set $h_i = h(X_1, X_2, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n)$ and $\bar{h} = n^{-1} \Sigma_{i=1}^n h_i$, then

$$\text{Var}\{h(X_1, X_2, \cdots, X_{n-1})\} \leqq E \sum_{i=1}^{n} (h_i - \bar{h})^2. \tag{4.1}$$

The Efron–Stein inequality makes light work of bounding Var $V_{k,n}$.

Lemma 4.2. There is a constant b_d such that for all $n \geqq 1$,

$$\text{Var } V_{k,n} \leqq b_d n. \tag{4.2}$$

Proof. If we let h denote the number of vertices of degree k in the minimal spanning tree of $\{X_1, X_2, \cdots, X_{n-1}\}$, then h is certainly symmetric. We also note that to replace $\bar{h}$ in $\Sigma_{i=1}^n (h_i - \bar{h})^2$ by $V_{k,n}$ will only make the sum larger, so

$$\text{Var } V_{k,n-1} \leqq E \sum_{i=1}^{n} (h_i - V_{k,n})^2. \tag{4.3}$$

Now by Lemma 2.2, $|h_i - V_{k,n}| \leqq 4D_d$, so (4.3) implies

$$\text{Var } V_{k,n-1} \leqq n\, 16 D_d^2,$$

and this inequality implies (4.2) with $b_d = 32 D_d^2$.

We now have all the tools needed to complete the proof of Theorem 2 (and to obtain Theorem 1 as a corollary). By Lemma 4.2 we see that for $m_n = n^2$, we have for any $\varepsilon > 0$

$$\sum_{n=1}^{\infty} P(|m_n^{-1} V_{k,m_n} - \alpha_{k,d}| > \varepsilon) \leqq \sum_{n=1}^{\infty} (\varepsilon m_n)^{-2} (b_d m_n) < \infty;$$

so, by the Borel–Cantelli lemma, we have proved that $V_{k,m_n} \sim m_n \alpha_{k,d}$ with probability 1. Now for any $m_n \leqq m < m_{n+1}$, we have $m - m_n \leqq 3n \leqq 3m^{1/2}$ whence by Lemma 2.2 we have

$$\begin{aligned} |m^{-1} V_{k,m} - m_n^{-1} V_{k,m_n}| &= |(m^{-1} - m_n^{-1}) V_{k,m} + m_n^{-1} (V_{k,m} - V_{k,m_n})| \\ &\leqq m(m - 3m^{1/2})^{-1} - 1 + (m - 3m^{1/2})^{-1} 4 D_d (3m^{1/2}). \end{aligned}$$

Letting $m \to \infty$ and using the fact that $V_{k,m_n} \sim m_n \alpha_{k,d}$ therefore implies that $V_{k,m} \sim m \alpha_{k,d}$ holds with probability 1. This completes the proof of Theorem 2.

5. Extension to general densities

The proof of Theorem 2 depended heavily on the assumption that the points were uniformly distributed. In this section we will extend the result to arbitrary densities of R^d.

We first treat the case of densities with bounded support, and we begin by constructing a family of densities $\{f\}$, which we will call the *family of blocked densities*. From the construction of the family of blocked densities, we will see that, under the L^1-norm, it will be dense in the set of all densities on the unit cube $[0, 1]^d$. We first partition the cube into m^d cubical cells $\{Q_i\}$, each cell of edge length m^{-1}. Now, within each cell Q_i, we define a centered subcell $\tilde{Q}_i$ consisting of all points of Q_i which are further than δ from every face of Q_i where δ is a constant satisfying $m^{-1} > 2\delta > 0$. If for some m, δ, and β, we have (1) f is constant and bounded below by $\beta > 0$ on each $\tilde{Q}_i$, $1 \leqq i \leqq m^d$, and (2) f is 0 on $\bigcup_{1 \leqq i \leqq m^d} \{Q_i \cap \tilde{Q}_i^c\}$, we say f is a *blocked density* with parameters m, δ, and β.

Lemma 5.1. If $X_1, \cdots, X_n$ are i.i.d. with a blocked density, then with probability 1 we have

$$\lim_{n \to \infty} n^{-1} V_{k,n} = \alpha_{k,d}. \tag{5.1}$$

Before beginning the proof of Lemma 5.1, it will be useful to collect a few geometric facts. First, if Q and Q' are subcubes of $[0, 1]^d$ which have edge length r and which share a common face, then the Pythagorean theorem shows that any two points of Q and Q' are at most a distance apart of $r(d+3)^{1/2}$. Second, if a cube like $[0, 1]^d$ is divided into s^d labeled subcells C_i of edge s^{-1}, then there is a permutation $\sigma: \{1, 2, \cdots, s^d\} \to \{1, 2, \cdots, s^d\}$ such that for all $1 \leqq i < s^d$, the cells $C_{\sigma(i)}$ and $C_{\sigma(i+1)}$ share a common face. Third, if we define a graph $G = (V, E)$ by taking a vertex set $V = \{y_1, y_2, \cdots, y_n\}$ and edge set $E = \{e = \{y_i, y_j\} : |y_i - y_j| < \delta\}$, and if G is connected, then the MST$(y_1, y_2, \cdots, y_n)$ is a subgraph of G.

We can now begin the proof of Lemma 5.1. The first observation is a structural one which can be given as follows: there is a random integer τ such that

(1) $\tau < \infty$ with probability 1, and

(2) for all $n \geqq \tau$, the MST$(X_1, X_2, \cdots, X_n)$ is composed of the union of MST$(\{X_1, X_2, \cdots, X_n\} \cap \tilde{Q}_i)$ over $1 \leqq i \leqq m^d$, together with exactly $m^d - 1$ edges which join elements of $\{X_1, X_2, \cdots, X_n\}$ that are in different cells $\tilde{Q}_i$.

To prove the existence of τ we decompose each $\tilde{Q}_i$ into k^d subcells where k is an integer satisfying $k^{-1}m^{-1}(d+3)^{1/2} < \delta$. We then define τ to be the least value of n such that for all $1 \leqq i \leqq m^d$, each of the k^d subcells of $\tilde{Q}_i$ is occupied by at least one element of $\{X_1, X_2, \cdots, X_n\}$. Standard multinomial bounds

(and the fact that $\beta > 0$) are enough to show $E\tau < \infty$, so τ is certainly finite with probability 1.

To verify the second property required of τ, we consider the graph $G = (V, E)$ defined by $V = \{X_1, X_2, \cdots, X_n\}$ and $E = \{e = \{x_i, x_j\} : |x_i - x_j| < \delta\}$. By the relation between k and δ, there will be an edge of E between any two X_i's in the same subcells or in two subcells which share a face. By the definition of τ and applying the second and third geometrical facts, we see that the union of $\text{MST}(\{X_1, X_2, \cdots, X_n\} \cap \tilde{Q}_i)$, $1 \leqq i \leqq m^d$, is a subgraph of G. Since G has no edges which meet two different $\tilde{Q}_i$, we see that an algorithmic construction of $\text{MST}(X_1, X_2, \cdots, X_n)$ for $n \geqq \tau$ will first construct complete minimal spanning trees on each of the point sets $\{X_1, X_2, \cdots, X_n\} \cap \tilde{Q}_i$. Since at the time of completion of these m^d trees there are m^d connected components in the forest which is created by the edges-ordered algorithm, we see there are $m^d - 1$ edges greater than δ in $\text{MST}(X_1, X_2, \cdots, X_n)$. Since for the m^d sets, $\tilde{Q}_i \cap \{X_1, X_2, \cdots, X_n\}$ to be joined in a tree requires at least $m^d - 1$ edges which hit two different Q_i's, we see that for $n \geqq \tau$ there are exctly $m^d - 1$ edges of $\text{MST}(X_1, X_2, \cdots, X_n)$ which join two distinct $Q_i \cap \{X_1, X_2, \cdots, X_n\}$. We have thus verified the second required property of τ.

We now have all of the tools needed to complete the proof of Lemma 5.1. If we let A denote the set of edges of $\text{MST}(X_1, X_2, \cdots, X_n)$ which intersect two Q_i's, then for $n \geqq \tau$ we have $|A| = m^d - 1$. The number of vertices which have different degrees in the tree $\text{MST}(X_1, X_2, \cdots, X_n)$ than in the forest F defined by

$$F = \bigcup_{i=1}^{m^d} \text{MST}(Q_i \cap \{X_1, X_2, \cdots, X_n\}) \tag{5.2}$$

is therefore at most $2(m^d - 1)$. Now, if h_i denotes the value of f on $\tilde{Q}_i$, then

$$N_i \equiv |Q_i \cap \{X_1, X_2, \cdots, X_n\}| \sim nh_i(1 - 2\lambda)^d m^{-d},$$

and we have by Theorem 2 and the law of large numbers that with probability 1, the numbers of vertices of degree k in F is asymptotic to

$$\sum_{i=1}^{m^d} \alpha_{k,d} N_i = \alpha_{k,d} \sum_{i=1}^{m^d} N_i \sim \alpha_{k,d} n \int f(x)dx.$$

This completes the proof of Lemma 5.1.

To move from the class of blocked densities which are treated by Lemma 5.1 to the general densities of Lemma 5.2, we use a lemma from Strassen (1965).

Lemma 5.2 (Strassen). Suppose μ_1 and μ_2 are probability measures on a bounded subset of R^d satisfying $\mu_1(A) \leqq \mu_2(A) + \varepsilon$ for some $\varepsilon > 0$ and all closed

A. Then there is a probability measure ν on $R^d \times R^d$ satisfying $\nu(A, R^d) = \mu_1(A)$, $\nu(R^d, A) = \mu_2(A)$, and $\nu(B) \leqq \varepsilon$ where $B = \{(x, x') : x \neq x'\}$.

To complete the proof of Theorem 3, we choose an $\varepsilon > 0$ and let f' denote a blocked density such that $\| f' - f \|_1 \leqq \varepsilon$. Let μ_1 be the measure with density f and μ_2 be the measure with density f', and define an independent sequence of random vectors $\{(X_i, X_i')\}$ with distribution given by ν from Strassen's lemma. By the law of large numbers, we then have with probability 1 that

$$\limsup_{n \to \infty} n^{-1} \sum_{i=1}^{n} 1(X_i \neq X_i') \leqq \varepsilon.$$

If we let $V_{k,n}'$ denote the number of vertices of degree k in the MST$\{X_1', X_2', \cdots, X_n'\}$, we then have by the corollary to Lemma 2.2 that

$$\limsup_{n \to \infty} n^{-1} | V_{k,n} - V_{k,n}' | \leqq 8 D_d \varepsilon.$$

By Lemma 4.1, we have already seen that with probability 1 $V_{k,n}' \sim n\alpha_{k,d}$; so, by the arbitrariness of $\varepsilon > 0$, we also have with probability 1 that $V_{k,n} \sim n\alpha_{k,d}$ since there was nothing special about $[0, 1]^d$ except its boundedness. The proof of Theorem 3 is thus complete in the case of random variables with compact support.

In fact, there is almost nothing left to show that Theorem 3 holds for arbitrary densities on R^d. If f is any given density, then there is a density f' which has compact support and such that $\| f' - f \|_1 < \varepsilon$. We can then argue exactly as before in the passage from blocked densities on $[0, 1]^d$ to general densities on $[0, 1]^d$; one just uses Strassen's Lemma and Lemma 2.2.

6. Information on the constants $\alpha_{k,d}$

The first task is to show that the constant α of Theorem 1 is strictly positive. It will be clear from the proof that for all $d \geqq 2$, we also have $\alpha_{1,d} > 0$.

Because of the relation $\phi(t) \sim \gamma t^d$ given by (3.7) and the identification $\gamma = \alpha_{k,d}$ made in Lemma 3.2, it will suffice to prove (in $d = 2$) that if L_t denotes the number of leaves of the minimal spanning tree M_t of the points of the uniform Poisson process on $[0, t]$, then there is a constant α_0 such that $\alpha_0 > 0$ and

$$EL_t \geqq \alpha_0 t^2 \tag{6.1}$$

for all sufficiently large t.

To prove (6.1) we will show that for any square S of area 1, there is a positive probability p that S contains a leaf of M_t. Since $[0, t]^2$ contains at least $\lfloor t \rfloor^2$ such disjoint squares, we would then have

$$EL_t > p\lfloor t \rfloor^2, \tag{6.2}$$

and (6.2) is sufficient to show (6.1).

We may as well consider $S = [0, 1]^2$ and partition S into m^2 subcells, which we label as Q_{ij}, $1 \leqq i \leqq m$, $1 \leqq j \leqq m$, where we use the natural lattice point labels. When m is odd, S has a unique *center cell* $Q_{k+1,k+1}$, if $m = 2k + 1$, and there are $4m - 4$ *boundary cells* Q_{ij} which share a face with S. We now let q be the probability that the center cell and all boundary cells have exactly one point of the Poisson process, but all of the other subcells of S are empty. Geometrical arguments which are now quite familiar give us that the point in the center cell must have degree 1 if $m \geqq 7$. This gives us that $p \geqq q > 0$, so (6.1) is established.

Since the bounds on α which are provided by constructions like the one just given are so intrinsically crude, we do not pursue the numerical evaluation of q.

It seems hopeless to find an analytic approach that will determine the values of $\alpha_{k,d}$. Consequently, we have performed some limited computer simulations for $d = 2$ in an effort to begin to understand the behavior of the $\alpha_{k,2}$. In Table 1 we report the results. The size of the simulated tree is denoted by n, and for each $n = 2^k$, $4 \leqq k \leqq 16$, 20 MST were simulated. In our estimates for $\alpha_{k,2}$, the $\hat{\alpha}_j$ were taken to be the average fraction of observed vertices of degree j, $1 \leqq j \leqq 5$.

TABLE 1
Monte Carlo estimate of frequencies

Sample size	Vertex degrees				
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$
16	0.303	0.528	0.159	0.009	0.000
32	0.266	0.542	0.181	0.011	0.000
64	0.243	0.555	0.191	0.010	0.000
128	0.223	0.575	0.197	0.005	0.000
256	0.226	0.561	0.206	0.006	0.000
512	0.223	0.564	0.208	0.006	0.000
1024	0.224	0.562	0.205	0.009	0.000
2048	0.223	0.563	0.207	0.007	0.000
4096	0.221	0.566	0.206	0.007	0.000
8192	0.221	0.565	0.206	0.008	0.000
16384	0.221	0.566	0.206	0.007	0.000
32768	0.221	0.566	0.206	0.007	0.000
65536	0.221	0.566	0.206	0.007	0.000

A vertex of degree 5 is obviously very rare. Still, they have been observed in

trees of $8K$ and $16K$ vertices. Another point to be made by the simulation is that vertex frequencies seem to change very little once n is as big as 100.

Probably the most intriguing speculation to emerge out of the Monte Carlo estimates is that $\alpha = \alpha_{1,2}$ (the proportion of leaves) is decently approximated by 2/9. The simulations do not strongly support the possibility that α exactly equals 2/9; but, in this connection, it is still interesting to note that Prodinger (1986) established that, if $h_n^{(d)}$ denotes the expected height of the dth highest leaf of a random planted plane tree, then for all $\varepsilon > 0$,

$$h_n^{(d)} = \sqrt{\pi n} - \frac{1}{2} - \frac{2}{3}\sum_{s=1}^{d-1}\left(\frac{2}{9}\right)^s \binom{2s+1}{s} + O(n^{-\frac{1}{2}+\varepsilon}). \tag{6.3}$$

We see no way to connect the two models, but the coincidental appearance of the unusual fraction 2/9 seems remarkable enough to mention.

Some caution should be given concerning any conjecture that $\alpha = 2/9$. In fact, before careful simulations were done, we felt the hypothesis $\alpha = 0$ was reasonable. This speculation arose because we erroneously thought that leaves had to be near the edge of the point cloud, not in the middle. This false start made clear for us that simulation and analogy are not always trustworthy guides in this area.

Roberts (1968), (1969) give some simulation results which are related to those we have done. In particular, Roberts (1968) gives results concerning the average length of a branch of the MST when the data are sampled uniformly from the unit sphere in two and three dimensions. The results of Roberts (1969) give interesting information concerning the probability that a point of a Poisson process is the nearest neighbor of n other points.

7. Concluding remarks

As the last section reveals, there are many open issues concerning the constants $\alpha_{k,d}$. Progress on these problems would be very interesting, but such progress is not likely to be forthcoming. For that reason, we focus here on the interesting analytical problems which remain concerning $V_{k,n}$ in view of the relative completeness of our analysis assuming the X_i are absolutely continuous random variables.

In the first place, we should acknowledge that we have treated a result like the strong law of large numbers, yet we have no knowledge concerning a possible central limit theorem. This is an inversion of traditional developments, but there seems to be serious difficulty in developing a central limit theory for non-linear functionals like our $V_{k,n}$.

A second mystery concerns the behavior of $V_{k,n}$ for singular measures. To be specific, suppose $X_i^{(1)} = \Sigma_{j=1}^{\infty} \xi_{i,j} 3^{-j}$ and $X_i^{(2)} = \Sigma_{j=1}^{\infty} \xi'_{i,j} 3^{-j}$, where $\xi_{i,j}$ and $\xi'_{i,j}$ are identically distributed, jointly independent sequences such that

$P(\xi_{i,j} = 0) = P(\xi_{i,j} = 2) = \frac{1}{2}$. Now, taking $X_i = (X_i^{(1)}, X_i^{(2)})$, we see the X_i are independent, identically distributed random variables in the singular support in $[0, 1]^2$. For of the subadditive Euclidean functions as defined in Steele (1981a), such as the length of the shortest tour of the points $\{X_1, X_2, \cdots, X_n\}$ or the length of the minimal matching of $\{X_1, X_2, \cdots, X_n\}$, we have no trouble dealing with singular random variables or random variables with a singular part. For those functionals, the singular part makes no contribution.

The issue with the $V_{k,n}$ functional is quite distinct. We suspect that for the X_i, $1 \leqq i < \infty$, defined above, one has $V_{k,n} \sim c_k n$ for some $c_k > 0$. This is in rather surprising contrast to the other parts of the theory of subadditive functionals. There is also an interesting second layer of subtlety which rests on the fact there is no reason to suppose that c_k equals the $\alpha_{k,2}$ of the absolutely continuous case.

The question of the number of leaves of a minimal spanning tree of a random sample from a singular distribution might well have a close relation to the fractal nature of the support. If that connection is born out, it will provide yet another instance where self-similarity (or approximate self-similarity) ties into fractal geometry. For several examples of more traditional probability theory which mixes the elements of self-similarity, fractals, and singular measures, one can consult the interesting paper of Hughes et al. (1982). For much interesting speculation, the classic source is, of course, Mandelbrot (1977).

Acknowledgement

Thanks are due to Timothy Snyder for numerous valuable comments on this revised manuscript.

References

BEARDWOOD, J., HALTON, J. H., AND HAMMERSLEY, J. M. (1959) The shortest path through many points. *Proc. Camb. Phil. Soc.* **55**, 299–327.

BENTLEY, J. L. (1978) Fast algorithms for constructing minimal spanning trees in coordinate spaces. *IEEE Comput.* **27**, 97–105.

BINGHAM, N. H. (1981) Tauberian theorems and the central limit theorem. *Ann. Prob.* **9**, 221–231.

DE BRUIJN, N. D. (1961) *Asymptotic Methods in Analysis*, 2nd edn. Wiley Interscience, New York.

CHIN, F. (1978) Algorithms for updating minimal spanning trees. *J. Comput. Syst.* **16**, 333–344.

EFRON, B. AND STEIN, C. (1981) The jackknife estimate of variance. *Ann. Statist.* **9**, 586–596.

FRIEDMAN, J. H. AND RAFSKY, L. C. (1979) Multivariate generalizations of the Wolfowitz and Smirnov two-sample tests. *Ann. Statist.* **7**, 697–717.

HILBERT, D. AND COHN-VOSSEN, S. (1952) *Geometry and the Imagination*. Chelsea, New York.

HOCHBAUM, D. AND STEELE, J. M. (1982) Steinhaus' geometric location problem for random samples in the plane. *Adv. Appl. Prob.* **14**, 55–67.

HUGHES, B. D., MONTROLL, B. W., AND SHLESINGER, M. F. (1982) Fractal random walks. *J. Statist. Phys.* **28**, 111–128.

JUNG, H. A. (1974) Determination of minimal paths and spanning trees in graphs. *Computing* **13**, 249.

KANG, A. N. C. (1977) Storage reduction through minimal spanning trees and spanning forests. *IEEE Comput.* **26**, 425–434.

KATAJAINEN, J. (1983) On the worst case of a minimal spanning tree algorithm for Euclidean space. *BIT* **23**, 2–8.

MALLION, R. B. (1975) Number of spanning trees in a molecular graph. *Chem. Phys. Lett.* **36**, 170–174.

MANDELBROT, B. B. (1977) *Fractals, Form, Chance, and Dimension*. W. H. Freeman, San Francisco.

NAJOCK, D. AND HEYDE, C. C. (1982) On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *J. Appl. Prob.* **19**, 675–680.

PENNY, D. (1980) Techniques for the verification of minimal phylogenetic trees illustrated with 10 mammalian hemoglobin sequences. *Biochem. J.* **187**, 65–74.

PRODINGER, H. (1986) The average height of the d-th highest leaf of a planted tree. *Networks* **16**, 67–75.

PYNN, C. (1972) Improved algorithm for construction of minimal spanning trees. *Elect. Lett.* **8**, 143.

RÉNYI, A. (1959) Some remarks on the theory of random trees. *MTA Mat. Ket. Imt. Kozl.* **4**, 73–85 (also *Selected Papers of Alfréd Rényi*, Akadémiai Kaidó, Budapest).

ROBERTS, F. D. K. (1968) Random minimal trees. *Biometrika* **55**, 255–258.

ROBERTS, F. D. K. (1969) Nearest neighbors in a Poisson ensemble. *Biometrika* **54**, 401–406.

ROGERS, C. A. (1964) *Packing and Covering*. Cambridge University Press, Cambridge.

ROMANE, F. (1977) Possible use of minimum spanning tree in phytoecology. *Vegetation* **33**, 99–106.

SCHMIDT, R. (1925) Über das Borelsche Summierungsverfahren. *Schriften d. Königsberger gel. Gesellschaft* **1**, 205–256.

STEELE, J. M. (1981a) Subadditive Euclidean functionals and non-linear growth in geometric probability. *Ann. Prob.* **9**, 365–376.

STEELE, J. M. (1981b) Complete convergence of short paths and Karp's algorithm for the TSP. *Math. Operat. Res.* **6**, 374–378.

STRASSEN, V. (1965) The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439.

WALD, A. AND WOLFOWITZ, J. (1940) On a test whether two-samples are from the same population. *Ann. Math. Statist.* **11**, 147–162.

WHITNEY, V. K. M. (1972) Minimal spanning tree. *Comm. ACM* **15**, 273.

WU, F. Y. (1977) Number of spanning trees on a lattice. *J. Phys. A* **10**, L113–L115.