THE CENTRAL LIMIT THEOREM FOR WEIGHTED MINIMAL SPANNING TREES ON RANDOM POINTS

By Harry Kesten and Sungchul Lee

Cornell University and National University of Singapore

Let $\{X_i,\ 1 \le i < \infty\}$ be i.i.d. with uniform distribution on $[0,1]^d$ and let $M(X_1,\ldots,X_n;\alpha)$ be $\min\{\sum_{e \in T'} |e|^\alpha;\ T'$ a spanning tree on $\{X_1,\ldots,X_n\}\}$. Then we show that for $\alpha > 0$,

$$\frac{M(X_1,\ldots,X_n;\alpha) - EM(X_1,\ldots,X_n;\alpha)}{n^{(d-2\alpha)/2d}} \to N(0,\sigma_{\alpha,d}^2)$$

in distribution for some $\sigma_{\alpha,d}^2 > 0$.

1. Introduction. Let $\{X_1,\ldots,X_n\}$ be a finite subset of $\mathbb{R}^d$, $d \ge 2$. For given $\alpha > 0$, a minimal spanning tree (MST) on $\{X_1,\ldots,X_n\}$ with weight function $\psi_\alpha(y) = y^\alpha$ is a connected graph T such that

$$M(X_1,\ldots,X_n) = M(X_1,\ldots,X_n;\alpha)$$
$$= \sum_{e \in T} |e|^\alpha$$
$$= \min\left\{ \sum_{e \in T'} |e|^\alpha;\ T' \text{ a spanning tree on } \{X_0,\ldots,X_n\} \right\},$$

where $|e| = |X_i - X_j|$ is the Euclidean length of the edge $e = (X_i, X_j)$. When $\{X_i,\ 1 \le i < \infty\}$ are i.i.d. with common distribution μ, which has compact support in $\mathbb{R}^d$, $d \ge 2$, Steele (1988) showed that for $0 < \alpha < d$,

$$(1.1) \qquad n^{-(d-\alpha)/d} M(X_1,\ldots,X_n;\alpha) \to c(\alpha,d) \int_{\mathbb{R}^d} f(x)^{(d-\alpha)/d}\, dx \quad \text{a.s.,}$$

where f is the density of the absolutely continuous part of μ and $c(\alpha,d)$ is a strictly positive but finite constant which depends only on the power α and the dimension d. Moreover, if $\{X_i,\ 1 \le i < \infty\}$ are i.i.d. with uniform distribution on $[0,1]^d$, then Aldous and Steele (1992) showed that

$$(1.2) \qquad M(X_i,\ldots,X_n;d) \to c(d,d) \text{ in } L^2.$$

In this paper we prove a central limit theorem for minimal spanning trees on uniformly distributed points in $[0,1]^d$. Ramey (1982) already demonstrated one approach to this problem. He argued that a certain property of continuum percolation would imply a central limit theorem for minimal spanning trees for $d = 2$, $\alpha = 1$. However, he did not prove this property of continuum percolation, and only very recently did Alexander (1996) prove a central limit theorem by

Ramey's approach for (the Poissonized version of) minimal spanning trees for $d = 2$, $\alpha = 1$.

Here we use a completely different approach to the problem. We represent $M(X_1, \ldots, X_n; \alpha) - EM(X_1, \ldots, X_n; \alpha)$ as the sum of martingale differences [see (4.2) below] and apply Lévy's martingale central limit theorem [see Lévy (1937), Theorem 67.2]. In this way the proof of the central limit theorem for minimal spanning trees is reduced to a kind of weak law of large numbers estimate for certain conditional variances. Even though a weak law of large numbers is much easier to obtain, in general, than a central limit theorem, it still requires some independence. The required independence will be obtained from a (deterministic) monotonicity property for minimal spanning trees which allows us to approximate the conditional variances by quantities which "depend only locally" on the X_i. This takes a surprising amount of work. This is somewhat annoying because the monotonicity and the approximation which we prove in Propositions 3 and 4 seems much stronger than needed for the weak law of large numbers statement (4.19). Another drawback of our approach is that it is not quantitative. Further ideas are needed to obtain an error estimate in our central limit theorem.

The next theorem is our principal result.

THEOREM 1. *Let* $\{X_i, \ 1 \le i \le n\}$ *be i.i.d. with uniform distribution on* $[0, 1]^d$, $d \ge 2$. *Then, for any given* $\alpha > 0$,

$$(1.3) \qquad \frac{M(X_1, \ldots, X_n; \alpha) - EM(X_1, \ldots, X_n; \alpha)}{n^{(d-2\alpha)/2d}} \Rightarrow N(0, \sigma_{\alpha, d}^2)$$

for some $\sigma_{\alpha, d}^2 > 0$.

It is more natural for the proof to rescale the variables, so that the density of the point set $\{X_1, \ldots, X_n\}$ remains constant. In our proof we shall therefore work with n i.i.d. points which are uniform in $[0, n^{1/d}]^d$. This also allows us to work with more general weight functions than $\psi_\alpha(y) = y^\alpha$. Theorem 1 is then a simple rescaled version of the following result.

THEOREM 2. *Let* $\psi \colon [0, \infty) \to [0, \infty)$ *be a strictly increasing function such that*

$$(1.4) \qquad \frac{1}{y} \log \psi(y) \to 0 \quad \text{as } y \to \infty.$$

Let $X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}$ *be i.i.d. and uniform on* $[0, n^{1/d}]^d$ *and define*

$$
\begin{aligned}
M(X_1^{(n)}, \ldots, X_n^{(n)}) &= M(X_1^{(n)}, \ldots, X_n^{(n)}; \psi) \\
&= \min\left\{ \sum_{e \in T'} \psi(|e|) \colon T' \text{ a spanning tree on } \{X_1^{(n)}, \ldots, X_n^{(n)}\} \right\}.
\end{aligned}
$$

(1.5)

Then

$$(1.6) \qquad \frac{M(X_1^{(n)}, \ldots, X_n^{(n)}; \psi) - EM(X_1^{(n)}, \ldots, X_n^{(n)}; \psi)}{n^{1/2}} \Rightarrow N(0, \sigma_{\psi, d}^2)$$

for some $\sigma_{\psi, d}^2 > 0$.

To avoid a few technicalities, one step in the proof (Lemma 10) will be carried out only for $\psi = \psi_\alpha$. Details for general strictly increasing ψ are available from the authors.

A byproduct of our proof concerns a linearity property of $n^{\alpha/d} EM(X_1, \ldots, X_n; \alpha)$. Alexander (1994) and Redmond and Yukich (1994) proved in the case $\alpha = 1$ with i.i.d. uniform X_i on $[0, 1]^d$ that for a suitable constant $0 < c(1, d) < \infty$,

$$(1.7) \qquad |EM(X_1, \ldots, X_n; 1) - c(1, d)n^{(d-1)/d}| = O(n^{(d-2)/d}).$$

The following is a corollary to the proof of Lemma 13.

COROLLARY 1. *Under the hypotheses of Theorem* 1, *there exists for each* $\alpha > 0$ *a constant* $\overline{\rho}(\alpha, d) > 0$ *such that*

$$(1.8) \qquad \lim_{n \to \infty} n^{\alpha/d - 1} E\{M(X_1, \ldots, X_n; \alpha)\} = \overline{\rho}(\alpha, d)$$

and

$$(1.9) \qquad \begin{aligned} &\lim_{n \to \infty} [(n+1)^{\alpha/d} E\{M(X_1, \ldots, X_{n+1}; \alpha)\} - n^{\alpha/d} E\{M(X_1, \ldots, X_n; \alpha)\}] \\ &\qquad = \overline{\rho}(\alpha, d). \end{aligned}$$

REMARK. We would like to mention here that Janson (1995) has proven a central limit theorem for a minimal spanning tree on a large number of points $X_1, \ldots, X_n$ under the assumption that all the "distances" between all the pairs of points are i.i.d. uniform on $[0, 1]$. This setup is sometimes called the *complete graph case*, whereas our situation is described as the *Euclidean case*.

In Section 2 we collect all deterministic properties of MST which we need. In particular we prove the monotonicity of $D(\mathscr{A}, \mathscr{B})$ [see (2.5) for the definition of $D(\mathscr{A}, \mathscr{B})$]. In Section 3 we estimate various moments of $D(\mathscr{A}, \mathscr{B})$. In Section 4 we prove Theorem 2 in a Poissonized version first and then the original version of Theorem 2.

In this paper there are lots of strictly positive but finite constants whose specific values are not of interest to us. We denote them by C_i, $C(q)$ or D_d. For any random variable X and probability distribution μ, we write $\mu\{X\}$ for the expectation of X with respect to μ.

2. Monotonicity of $D(\mathscr{A}, \mathscr{B})$.

In this section we study two algorithms for an MST: Kruskal's greedy algorithm and the "add and delete" algorithm. Using these algorithms we establish the monotonicity of $D(\mathscr{A}, \mathscr{B})$.

KRUSKAL'S GREEDY ALGORITHM.   Let $G = (V, E, w)$ be a connected weighted graph with vertex set V, edge set E and weight function w: $E \to [0, \infty)$. Assume that the cardinalities $|V|$, $|E|$ of V, E, respectively, are finite. A minimal spanning tree on (G, w) is a tree T with vertex set V (which makes it a spanning tree) and edge set contained in E (which we express by $T \subset E$ for brevity) and such that

$$w(T) := \sum_{e \in T} w(e) = \min \left\{ \sum_{e \in T'} w(e) \colon\ T' \subset E,\ T' \text{ a spanning tree} \right\}.$$

A minimal spanning tree T on G can be constructed by the following greedy algorithm due to Kruskal (1956).

Step 1. Let $T_0 = \phi$ and $E_0 = E \setminus \{e \in E \colon \{e\} \text{ is a circuit}\}$.

Step 2. Once T_i and E_i have been determined, choose an edge $e_{i+1} \in E_i$ such that $w(e_{i+1}) = \min_{e \in E_i} w(e)$ and construct T_{i+1} by adding the edge e_{i+1} to T_i, that is, take $T_{i+1} = T_i \cup \{e_{i+1}\}$. Then take

$$E_{i+1} = \{e \in E_i \setminus \{e_{i+1}\} \colon\ T_{i+1} \cup \{e\} \text{ does not contain a circuit}\}.$$

If there exists no such edge e_{i+1}, that is, if $E_i = \phi$, then let $T = T_i$ and stop.

Step 3. Replace i by $i + 1$. Return to Step 2.

PROPOSITION 1.   *Let* $G = (V, E, w)$ *be a connected weighted graph with* $|V| < \infty$, $|E| < \infty$ *and let* $T \subset E$ *be a tree obtained by Kruskal's greedy algorithm. Then* T *is an MST on* (G, w).

See Theorem 3A of Chartrand and Lesniak (1986) for the proof.

ADD AND DELETE ALGORITHM.   Let $G = (V, E, w)$ and $G' = (V', E', w')$ be two connected weighted graphs with $|V| < \infty$, $|E| < \infty$ and such that $V' = V$ or $V' = V \cup \{v'\}$ for a single vertex $v' \notin V$ and $E' = E \cup \{e_1', \ldots, e_n'\}$ for a finite number of edges $e_i' \notin E$. Note that the connectedness of G' requires that if v' is present, at least one edge e_i' connects v' to some $v \in V$. We assume that $w'(e) = w(e)$ for all $e \in E$ so that w' is an extension of w. We shall therefore drop the prime from the w' in the sequel and denote the weight function on G' also by w.

Now let T be an MST on G and assume that we want to construct an MST T' on G'. We may, of course, construct an MST T' by directly applying the greedy algorithm to the graph G'. However, the greedy algorithm is not "effective" in this case, because it does not use the fact that T is already an MST on G. We propose the following add and delete algorithm to construct an MST T' on G' from T.

Step 1'. Let $E' \setminus E = \{e_1', e_2', \ldots, e_n'\}$ in some order and take $T_0' = T$.

Step 2'. When T_i' is given, add e_{i+1}' to T_i', that is, form $T_i' \cup \{e_{i+1}'\}$. The first time we add an edge to T_i' which is incident to $v' \notin V$ there will be no circuit in $T_i' \cup \{e_{i+1}'\}$ and we take $T_{i+1}' = T_i' \cup \{e_{i+1}'\}$. In all other cases, $T_i' \cup \{e_{i+1}'\}$ contains a unique circuit C_{i+1}'. Choose an edge $f_{i+1}' \in C_{i+1}'$ such that $w(f_{i+1}') = \max_{e \in C_{i+1}'} w(e)$. Delete f_{i+1}' from $T_i' \cup \{e_{i+1}'\}$. That is, define

$T'_{i+1} = T'_i \cup \{e'_{i+1}\} \setminus \{f'_{i+1}\}$. If $i = n$ and there exists no more edges e'_{i+1}, then let $T' = T'_i$ and stop.

Step $3'$. Replace i by $i + 1$. Return to Step 2.

In order to prove that a tree T' on V' obtained by the add and delete algorithm is an MST on G', we need the following criterion for a spanning tree to be an MST.

LEMMA 1.   *Let $G = (V, E, w)$ be a connected weighted graph with $|V| < \infty$, $|E| < \infty$, and let $T \subset E$ be a spanning tree on V. Then T is an MST on V if and only if for each edge $e \in E \setminus T$,*

$$(2.1) \qquad\qquad w(e) \geq \max\{w(f)\colon f \in C \setminus \{e\}\},$$

where C is the unique circuit in $T \cup \{e\}$.

PROOF.   Let T be an MST on V. If there exists an edge $e \in E \setminus T$ such that (2.1) fails, then there exists an edge $f \in C \setminus \{e\}$ such that $w(e) < w(f)$. Note that $f \in T$. Let $T' = T \cup \{e\} \setminus \{f\}$. Then T' is a spanning tree on V such that

$$\sum_{g \in T'} w(g) - \sum_{g \in T} w(g) = w(e) - w(f) < 0.$$

This contradicts the fact that T is an MST on V, so for each $e \in E \setminus T$, (2.1) holds.

Conversely, let T be a spanning tree on V such that for each $e \in E \setminus T$, (2.1) holds. Assume that T is not an MST on V. For the moment assume that $w(f_1) \neq w(f_2)$ for all $f_1 \neq f_2$. Let $T_g = \{e_1, \dots, e_n\}$ be an MST on V obtained by the greedy algorithm, where e_i is the ith edge obtained by the algorithm. Let e_{i_0} be the first edge in T_g such that $e_{i_0} \notin T$, that is, $e_i \in T$ for $1 \leq i < i_0$ but $e_{i_0} \notin T$ (note that necessarily $T \neq T_g$ if T is not an MST). Add e_{i_0} to T; then $T \cup \{e_{i_0}\}$ contains a circuit C. This circuit C must contain an edge $e \notin T_g$ because T_g does not contain a circuit. By (2.1) and by the assumption that $w(f_1) \neq w(f_2)$ for $f_1 \neq f_2$, $w(e) < w(e_{i_0})$. Now $e \notin T_g$ and $e_1, \dots, e_{i_0-1}, e$ do not form a circuit, because all those edges occur in T. Moreover, $w(e) < w(e_{i_0})$. Therefore, e_{i_0} cannot be the i_0th edge obtained by the algorithm for T_g. This contradiction shows that T is an MST on V.

If $w(f_1) = w(f_2)$ for some $f_1 \neq f_2$, then we apply the preceding argument to an MST $T_g = \{e_1, \dots, e_n\}$ obtainable by the greedy algorithm and with the following additional property:

$$\tau(T_g, T) = \max\{\tau(T'_g, T)\colon T'_g \text{ an MST obtainable by the greedy algorithm}\},$$

where, for $T'_g = (e'_1, e'_2, \dots, e'_n)$,

$$\tau(T'_g, T) = \max\{j + 1\colon e'_1 \in T,\ e'_2 \in T,\ \dots, e'_j \in T\}.$$

Let e_{i_0} be the first edge in T_g such that $e_{i_0} \notin T$, that is, $\tau(T_g, T) = i_0$. The tree $T \cup \{e_{i_0}\}$ contains a circuit C, and C contains an edge $e \notin T_g$. By (2.1), $w(e) \leq w(e_{i_0})$. If we can find such an e with $w(e) < w(e_{i_0})$, then we reach a contradiction as before. Assume then that $w(e) = w(e_{i_0})$ and construct an MST

T''_g by the following greedy algorithm. For $i \leq i_0 - 1$ pick e_i as the ith edge of T''_g. At stage i_0 pick e instead of e_{i_0} as the i_0th edge of T''_g. Then complete T''_g by the greedy algorithm. This construction leads to

$$\tau(T''_g, T) \geq \tau(T_g, T) + 1,$$

which is again a contradiction, so T is an MST on V. □

PROPOSITION 2. *Let $G = (V, E, w)$ be a connected weighted graph with $|V| < \infty$, $|E| < \infty$, and let T be an MST on V. Let $G' = (V', E', w)$ be a new connected weighted graph obtained by adding at most one vertex $v' \notin V$ to V and a finite number of weighted edges $e' \notin E$ to E, and with w' an extension of w. Let T' be a spanning tree of V' obtained from T by the add and delete algorithm. Then T' is an MST on V'.*

PROOF. Let $e \notin T'$. We will show that (2.1) holds, that is, $w(e) \geq \max\{w(f): f \in C \setminus \{e\}\}$, where C is the unique circuit in $T' \cup \{e\}$. Then by Lemma 1, T' is an MST on V'. Let the edges added to E be $\{e'_1, \ldots, e'_n\}$ and let T'_i be the ith tree obtained in the add and delete algorithm as described above. Thus $T'_{i+1} = T'_i \cup \{e'_{i+1}\} \setminus \{f'_{i+1}\}$ except for the first i for which e'_{i+1} is incident to v'; in this case $T'_{i+1} = T'_i \cup \{e'_{i+1}\}$. If $e \in E \setminus T$, then, since T is an MST on V, by Lemma 1 there exists a circuit C'_0 in $T'_0 \cup \{e\} = T \cup \{e\}$ with $w(e) \geq \max\{w(f): f \in C'_0 \setminus \{e\}\}$. If $e \in T$, but $e \notin T'$, then $e \in T'_0$ but $e \notin T'_n$ so that there is some i with $e \in T'_{i-1}$ but $e \notin T'_i$, that is, $e = f'_i$ and e belongs to a circuit C'_i in $T'_i \cup \{e'_i\}$ and

(2.2) $$w(e) = \max\{w(f): f \in C'_i\}.$$

Such a C'_i also exists if e is one of the edges $e'_1, \ldots, e'_n$ which does not belong to the final T'.

Now let $e \notin T'$ and C'_i as just described. If $C'_i \setminus \{e\} \subset T'$, then, by virtue of (2.2), (2.1) holds. If not, let T'_j, $j > i$, be the first spanning tree after T'_i in our sequence such that $C'_i \setminus \{e\} \not\subset T'_j$. That is, $C'_i \setminus \{e\} \subset T'_j \cup \{f'_j\}$ but $C'_i \setminus \{e\} \not\subset T'_j$. Then $f'_j \in C'_i \setminus \{e\}$. By choice of f'_j, $T'_j \cup \{f'_j\}$ contains a circuit C'_j, $f'_j \in C'_j$, and $w(f'_j) \geq \max\{w(f): f \in C'_j \setminus \{f'_j\}\}$. Therefore (see Figure 1), $T'_j \cup \{e\}$ contains a circuit $C_j \subset ((C'_i \setminus \{f'_j\}) \cup (C'_j \setminus \{f'_j\}))$ such that

$$w(e) \geq \max\{w(f): f \in C'_i \setminus \{e\}\}$$

$$= \max\{w(f): f \in C'_i \setminus \{e, f'_j\}\} \vee w(f'_j)$$

$$\geq \max\{w(f): f \in C'_i \setminus \{e, f'_j\}\} \vee \max\{w(f): f \in C'_j \setminus \{f'_j\}\}$$

$$\geq \max\{w(f): f \in C_j \setminus \{e\}\}.$$

Therefore, if $C'_i \setminus \{e\} \not\subset T'$, then for some $j > i$ there exists a circuit C_j in $T'_j \cup \{e\}$ with $w(e) \geq \max\{w(f): f \in C_j \setminus \{e\}\}$. We can now repeat the argument with C'_i replaced by C_j. Applying this argument finitely many times, we see that there exists a circuit C in $T' \cup \{e\}$ with $w(e) \geq \max\{w(f): f \in C \setminus \{e\}\}$, that is, (2.1) holds. □
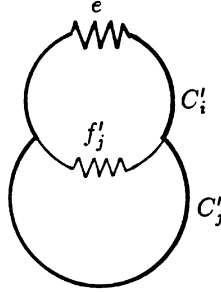
FIG. 1. *The circuit C_j in $(C'_i \setminus \{f'_j\}) \cup (C'_j \setminus \{f'_j\})$. Circuit C_j consists of the boldly drawn parts.*

LEMMA 2. *Let $G = (V, E, w)$ be a connected weighted graph with $|V| < \infty$ and $|E| < \infty$. If there exists a path $\pi = (e_1, \ldots, e_n)$ in G from $v_1 \in V$ to $v_2 \in V$ such that $w(e_i) \le \lambda$, $1 \le i \le n$, then in any MST T on G there exists a path $\pi' = (e'_1, \ldots, e'_m)$ in T from v_1 to v_2 such that $w(e'_j) \le \lambda$, $1 \le j \le m$.*

PROOF. Let T be an MST on G. Then either e_i is an edge in T or, if e_i is not an edge in T, by Lemma 1, $T \cup \{e_i\}$ contains a circuit C_i such that for all edges f in C_i we have $w(f) \le w(e_i) \le \lambda$. Thus the endpoints of e_i can be connected in T by a path with edges e with $w(e) \le \lambda$. This is true for each i, so also v_1 and v_2 are connected in T by a path with edges e with $w(e) \le \lambda$. $\square$

Let $\psi: [0, \infty) \to [0, \infty)$ be a given function. From now on we shall take our weight function to be of the following form: for the edge (x, y) between $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^d$,

$$(2.3) \qquad w(x, y) = \psi(|(x, y)|),$$

where $|(x, y)|$ is the Euclidean length of the line segment from x to y. Accordingly, we make the following definition for a finite subset $\mathscr{A}$ of $\mathbb{R}^d$:

$$(2.4) \qquad \begin{aligned} M(\mathscr{A}) &= M(\mathscr{A}; \psi) \\ &= \min\left\{ \sum_{(x, y) \in T} \psi(|(x, y)|) \colon T \text{ a spanning tree on } \mathscr{A} \right\}. \end{aligned}$$

Note that $M(\mathscr{A}; \psi)$ is just the weight of a minimal spanning tree for $G = (\mathscr{A}, E, w)$ when w is given by (2.3) and E consists of all edges between any two points of $\mathscr{A}$. Define further, for two disjoint finite subsets $\mathscr{A}$, $\mathscr{B}$ of $\mathbb{R}^d$,

$$(2.5) \qquad D(\mathscr{A}, \mathscr{B}) = D(\mathscr{A}, \mathscr{B}; \psi) = M(\mathscr{A} \cup \mathscr{B}) - M(\mathscr{B}).$$

Let Q_n be the cube

$$Q_n = [0, n^{1/d}]^d.$$

For $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$ and $L, l > 0$, we write

$$B[x, y] = \prod_{i=1}^{d} [x_i, y_i],$$

$$C(x) = \prod_{i=1}^{d} [x_i, x_i + 1],$$

$$A(B[x, y]; l, L) = \prod_{i=1}^{d} [x_i - l - L, y_i + l + L] \setminus \prod_{i=1}^{d} (x_i - l, y_i + l).$$

We shall call $B[x, y]$, $C(x)$ and $A(B[x, y], l, L)$ a *rectangle*, a *cube* and an *annulus*, respectively.

DEFINITION. Let $\delta > 0$. A *separating set of width* δ for the rectangle $B[x, y]$ is a finite set $\mathscr{S}$ in the annulus $A(B[x, y]; 1, \delta)$ with the property (see Figure 2):

(2.6)   Each line segment from $\partial\Pi(x_i - 1, y_i + 1)$ to $\partial\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$ passes within distance 1/3 of some point of $\mathscr{S}$.

Most of the time we shall consider situations in which all vertices of interest lie in $Q_n = [0, n^{1/d}]^d$. For weight functions w of the form $w(e) = \psi(|e|)$ with ψ strictly increasing, it is reasonable to think of the edge (x, y) as the straight line segment between x and y. Thus the edges of interest will also lie in Q_n. In this situation it is useful to relax (2.6) to the following property:

(2.7)   Each straight line segment from a point in $\partial\Pi(x_i - 1, y_i + 1) \cap Q_n$ to another point in $\partial\Pi[x_i - 1 - \delta, y_i + 1 + \delta] \cap Q_n$ passes within distance 1/3 of some point $\mathscr{S} \subset Q_n$.
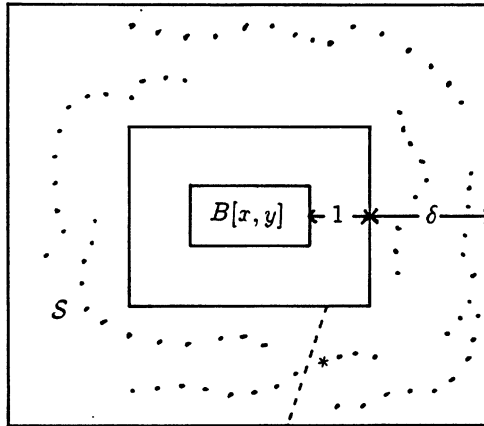

FIG. 2. *A separating set $\mathscr{S}$ for the rectangle $B[x, y]$. The asterisk* (∗) *denotes a point of $\mathscr{S}$ within distance 1/3 of the dashed line segment.*

We call a set $\mathscr{S} \subset Q_n \cap A(B[x, y]; 1, \delta)$ with this property an *n-separating set of width* δ for the rectangle $B[x, y]$. The reader should note that if $\partial\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$ lies outside Q_n, then $\mathscr{S} = \phi$ satisfies (2.7) so that ϕ is an *n*-separating set. The term "separating set" is justified by the next lemma.

LEMMA 3. *Let* ψ *be strictly increasing. If* $\mathscr{S} \subset A(B[x, y]; 1, \delta)$ *is a separating set for the rectangle* $B[x, y]$, *then in any MST* T (*with weight function* ψ) *on any finite subset* $\mathscr{A}$ *of* $\mathbb{R}^d$ *containing* $\mathscr{S}$ (*with all edges between a pair of vertices in* $\mathscr{A}$ *allowed*), *there are no edges between vertices in* $B[x, y]$ *and vertices in* $\mathbb{R}^d \setminus \Pi(x_i - 2 - \delta, y_i + 2 + \delta)$. *If* $\mathscr{A} \subset Q_n$, *then the same result holds for an n-separating set* $\mathscr{S}$.

PROOF. We only deal with the first case, when $\mathscr{S}$ is unrestricted; the case when all vertices and $\mathscr{S}$ are restricted to Q_n is essentially the same.

Assume T contains an edge (v_1, v_2) between $v_1 \in B[x, y]$ and $v_2 \in \mathbb{R}^d \setminus \Pi(x_i - 2 - \delta, y_i + 2 + \delta)$. This edge (v_1, v_2) contains a line segment from $\partial\Pi(x_i - 1, y_i + 1)$ to $\partial\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$, and hence there exists an $s \in \mathscr{S} \subset \mathscr{A}$ within distance $1/3$ of this edge. Removing the edge (v_1, v_2) breaks up T into two components. If v_1 and s lie in the same component, then connect v_2 to s. If v_2 and s lie in the same component, then connect v_1 to s. In each case we obtain a new spanning tree T'. We claim that

$$(2.8) \quad \sum_{e \in T} \psi(|e|) - \sum_{e \in T'} \psi(|e|) \geq \psi(|(v_1, v_2)|) - \max\{\psi(|(v_1, s)|), \psi(|(v_2, s)|)\} > 0.$$

This leads to a contradiction. To prove our claim, note that $|(v_1, s)| \geq 1$ and $|(v_2, s)| \geq 1$ because $B[x, y]$ has distance 1 to $\partial\Pi(x_i - 1, y_i + 1)$ and $\mathbb{R}^d \setminus \Pi(x_i - 2 - \delta, y_i + 2 + \delta)$ has distance 1 to $\partial\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$. Let t be a point on the line segment (v_1, v_2) such that $|(t, s)| \leq 1/3$. Then

$$\begin{aligned}
|(v_1, v_2)| &= |(v_1, t)| + |(v_2, t)| \\
&\geq |(v_1, s)| + |(v_2, s)| - 2|(s, t)| \\
&= \max\{|(v_1, s)|, |(v_2, s)|\} + \min\{|(v_1, s)|, |(v_2, s)|\} - 2|(s, t)| \\
&\geq \max\{|(v_1, s)|, |(v_2, s)|\} + 1 - \frac{2}{3} \\
&> \max\{|(v_1, s)|, |(v_2, s)|\}.
\end{aligned}$$

The last inequality in (2.8) now follows from the strict monotonicity of ψ. □

The next proposition is about the "monotonicity" of $D(\mathscr{A}, \mathscr{B})$.

PROPOSITION 3. *Let* ψ *be strictly increasing and let* $\mathscr{S} \subset A(B[x, y]; 1, \delta)$ *be a separating set for the rectangle* $B[x, y]$. *Let* $\mathscr{A}$ *and* $\mathscr{B}$ *be finite subsets of* $\mathbb{R}^d$ *such that* $\mathscr{A} \subset B[x, y]$, $\mathscr{S} \subset \mathscr{B}$ *and* $\mathscr{A} \cap \mathscr{B} = \phi$. *Then for any finite subset* $\mathscr{B}'$ *of* $\mathbb{R}^d$ *such that* $\mathscr{B} \cap \mathscr{B}' = \phi$ *and*

$$(2.9) \quad \mathscr{B}' \cap \Pi(x_i - 2 - \delta, y_i + 2 + \delta) = \phi,$$

we have

$$(2.10) \qquad\qquad D(\mathscr{A}, \mathscr{B}) \leq D(\mathscr{A}, \mathscr{B} \cup \mathscr{B}').$$

PROOF.   If $\mathscr{B}' = \{v_1, \ldots, v_k\}$, then it suffices to prove

$$D(\mathscr{A}, \mathscr{B} \cup \{v_1, \ldots, v_r\}) \leq D(\mathscr{A}, \mathscr{B} \cup \{v_1, \ldots, v_r, v_{r+1}\})$$

for $r = 0, 1, \ldots, k-1$. Thus we may restrict ourselves to the case when we add only one point to $\mathscr{B}$, say

$$\mathscr{B} \cup \mathscr{B}' = \mathscr{B} \cup \{v\}.$$

Also, if $\mathscr{A} = \{u_1, u_2, \ldots, u_l\}$, then

$$D(\mathscr{A}, \mathscr{B}) = M(\mathscr{A} \cup \mathscr{B}) - M(\mathscr{B})$$

$$= \sum_{r=1}^{l} [M(\mathscr{B} \cup \{u_1, \ldots, u_r\}) - M(\mathscr{B} \cup \{u_1, \ldots, u_{r-1}\})]$$

and

$$D(\mathscr{A}, \mathscr{B} \cup \mathscr{B}') = \sum_{r=1}^{l} [M(\mathscr{B} \cup \mathscr{B}' \cup \{u_1, \ldots, u_r\}) - M(\mathscr{B} \cup \mathscr{B}' \cup \{u_1, \ldots, u_{r-1}\})].$$

We therefore only have to prove

$$M(\mathscr{B} \cup \{u_1, \ldots, u_r\}) - M(\mathscr{B} \cup \{u_1, \ldots, u_{r-1}\})$$
$$\leq M(\mathscr{B} \cup \mathscr{B}' \cup \{u_1, \ldots, u_r\}) - M(\mathscr{B} \cup \mathscr{B}' \cup \{u_1, \ldots, u_{r-1}\})$$

for $r = 1, 2, \ldots, l$. Thus it suffices to prove the proposition for the case $\mathscr{A} = \{u\}$, $\mathscr{B}' = \{v\}$.

Let $\{f_1, \ldots, f_m\}$ be the edges from v to all the vertices of $\mathscr{B}$ in order of increasing length; m is the cardinality of the set $\mathscr{B}$. Let E be the set of edges between all pairs of $\mathscr{B}$. Let T be an MST for $(\mathscr{B}, E)$. Then by the add and delete algorithm construct an MST T_v on $\mathscr{B} \cup \{v\}$. Next let $\{h_1, \ldots, h_m\}$ be the edges from u to $\mathscr{B}$ again in order of increasing length. Again apply the add and delete algorithm to T to construct an MST T_u on $\mathscr{B} \cup \{u\}$. Also construct an MST T_{vu} on $\mathscr{B} \cup \{u, v\}$ by applying the add and delete algorithm to T_v. Since $\mathscr{B}$ contains a separating set $\mathscr{S}$ for $B[x, y]$, by Lemma 3, T_{vu} does not contain the edge (u, v) (recall (2.9) and $u \in \mathscr{A} \subset \mathscr{B}[x, y]$). So when we apply the add and delete algorithm to T_v in order to construct T_{vu}, we do not need to add the edge (u, v). Thus when carrying out the add and delete algorithm to form T_{vu} from T_v, we only have to add the same set of edges $h_1, \ldots, h_m$ from u to $\mathscr{B}$, as when forming T_u from T. Let T_i be the tree obtained from T after adding $f_1, \ldots, f_i$ and deleting appropriate edges $g_2, \ldots, g_i$, that is,

$$T_i = T \cup \{f_1, \ldots, f_i\} \setminus \{g_2, \ldots, g_i\}.$$

By Proposition 2, T_i is an MST on $(\mathscr{B} \cup \{v\},\ E \cup \{f_1, \ldots, f_i\})$. Also, let $T_{i,j}$ be the tree obtained from T_i after adding $h_1, \ldots, h_j$ and deleting appropriate edges $k_{i,2}, \ldots, k_{i,j}$. We shall prove

$$(2.11) \qquad \sum_{e \in T_{i,j+1}} \psi(|e|) - \sum_{e \in T_{i,j}} \psi(|e|) \leq \sum_{e \in T_{i+1,j+1}} \psi(|e|) - \sum_{e \in T_{i+1,j}} \psi(|e|)$$

for $i = 1, \ldots, m-1$ and $j = 1, \ldots, m-1$. This will give (by summing over j)

$$\sum_{e \in T_{i,m}} \psi(|e|) - \sum_{e \in T_{i,1}} \psi(|e|) \leq \sum_{e \in T_{i+1,m}} \psi(|e|) - \sum_{e \in T_{i+1,1}} \psi(|e|)$$

and, by iteration,

$$(2.12) \qquad \sum_{e \in T_{1,m}} \psi(|e|) - \sum_{e \in T_{1,1}} \psi(|e|) \leq \sum_{e \in T_{m,m}} \psi(|e|) - \sum_{e \in T_{m,1}} \psi(|e|).$$

Note that $T_{1,m}$ is the tree obtained by adding f_1 and then adding $h_1, \ldots, h_m$ and deleting $k_{1,2}, \ldots, k_{1,m}$. As long as we only add the single edge f_1 from v to $\mathscr{B}$, v is a leaf and f_1 is never deleted, so that $T_{1,m}$ is the same as T_u plus the single edge f_1. Therefore,

$$\sum_{e \in T_{1,m}} \psi(|e|) = M(\mathscr{B} \cup \{u\}) + \psi(|f_1|).$$

Similarly

$$\sum_{e \in T_{1,1}} \psi(|e|) = M(\mathscr{B}) + \psi(|f_1|) + \psi(|h_1|),$$

$$\sum_{e \in T_{m,m}} \psi(|e|) = M(\mathscr{B} \cup \{u, v\}),$$

$$\sum_{e \in T_{m,1}} \psi(|e|) = M(\mathscr{B} \cup \{v\}) + \psi(|h_1|).$$

Thus (2.12) says

$$M(\mathscr{B} \cup \{u\}) - M(\mathscr{B}) \leq M(\mathscr{B} \cup \{u, v\}) - M(\mathscr{B} \cup \{v\}).$$

This is the proposition for the case $\mathscr{A} = \{u\}$, $\mathscr{B}' = \{v\}$. To prove the proposition, it therefore suffices to prove (2.11).

To prove (2.11) we construct $T_{i+1,j}$ and $T_{i,j+1}$. By Proposition 2, $T_{i,j}$, which is obtained by first adding $f_1, \ldots, f_i$ to T and then adding $h_1, \ldots, h_j$ to T, is an MST on $(\mathscr{B} \cup \{u, v\},\ E \cup \{f_1, \ldots, f_i, h_1, \ldots, h_j\})$. To construct $T_{i+1,j}$ we should first add $f_1, \ldots, f_{i+1}$ and then $h_1, \ldots, h_j$. By Proposition 2, $T_{i+1,j}$ is then an MST on $(\mathscr{B} \cup \{u, v\}, E \cup \{f_1, \ldots, f_{i+1}, h_1, \ldots, h_j\})$. Consider instead of this the following construction. First construct $T_{i,j}$, then add f_{i+1} and delete an edge of maximal weight in the circuit formed in $T_{i,j} \cup \{f_{i+1}\}$. Let this circuit be $\Gamma_{i,j}$ and the deleted edge $\gamma_{i,j}$ and let the resulting tree be $\overline{T}_{i+1,j}$. By Proposition 2, $\overline{T}_{i+1,j}$ is also an MST on $(\mathscr{B} \cup \{u, v\},\ E \cup \{f_1, \ldots, f_{i+1}, h_1, \ldots, h_j\})$. We may therefore replace $\sum_{e \in T_{i+1,j}} \psi(|e|)$ in (2.11) by $\sum_{e \in \overline{T}_{i+1,j}} \psi(|e|)$. Similarly, we

may form a tree $\overline{T}_{i,\,j+1}$ by adding h_{j+1} to $T_{i,\,j}$ and deleting an edge $\delta_{i,\,j}$ of maximal weight from the circuit $\Delta_{i,\,j}$ in $T_{i,\,j} \cup \{h_{j+1}\}$. We may then replace $\sum_{e \in T_{i,\,j+1}} \psi(|e|)$ in (2.11) by $\sum_{e \in \overline{T}_{i,\,j+1}} \psi(|e|)$. Finally we can form a tree $\overline{T}_{i+1,\,j+1}$ by adding first f_{i+1} to $T_{i,\,j}$ and deleting $\gamma_{i,\,j}$; this gives $\overline{T}_{i+1,\,j}$. Then add h_{j+1} and delete the edge $\overline{\delta}_{i,\,j}$ of maximal weight in the circuit $\overline{\Delta}_{i,\,j}$ of $\overline{T}_{i+1,\,j} \cup \{h_{j+1}\}$. Again $\sum_{e \in \overline{T}_{i+1,\,j+1}} \psi(|e|) = \sum_{e \in T_{i+1,\,j+1}} \psi(|e|)$. This gives

$$\sum_{e \in \overline{T}_{i,\,j+1}} \psi(|e|) - \sum_{e \in T_{i,\,j}} \psi(|e|) = \psi(|h_{j+1}|) - \psi(|\delta_{i,\,j}|),$$

$$\sum_{e \in \overline{T}_{i+1,\,j+1}} \psi(|e|) - \sum_{e \in \overline{T}_{i+1,\,j}} \psi(|e|) = \psi(|h_{j+1}|) - \psi(|\overline{\delta}_{i,\,j}|).$$

Thus (2.11) is equivalent to

$$(2.13) \qquad\qquad \psi(|\delta_{i,\,j}|) \geq \psi(|\overline{\delta}_{i,\,j}|).$$

This, however, is easy. If $\gamma_{i,\,j} \notin \Delta_{i,\,j}$, then $\Delta_{i,\,j}$ is present in $\overline{T}_{i+1,\,j} \cup \{h_{j+1}\} = T_{i,\,j} \cup \{f_{i+1}, h_{j+1}\} \setminus \{\gamma_{i,\,j}\}$. Consequently, $\overline{\Delta}_{i,\,j} = \Delta_{i,\,j}$, $\overline{\delta}_{i,\,j} = \delta_{i,\,j}$ and (2.13) is trivial. If $\gamma_{i,\,j} \in \Delta_{i,\,j}$, the situation is as in Figure 3. Circuit $\overline{\Delta}_{i,\,j}$ is formed from a piece of $\Delta_{i,\,j}$ and a piece of $\Gamma_{i,\,j}$, but then by the choice of $\delta_{i,\,j}$,

$$\psi(|\delta_{i,\,j}|) \geq \max_{e \in \Delta_{i,\,j}} \psi(|e|);$$

$\gamma_{i,\,j} \in \Delta_{i,\,j}$ and the choice of $\gamma_{i,\,j}$ in $\Gamma_{i,\,j}$ further imply that

$$\psi(|\delta_{i,\,j}|) \geq \psi(|\gamma_{i,\,j}|) \geq \max_{e \in \Gamma_{i,\,j}} \psi(|e|).$$
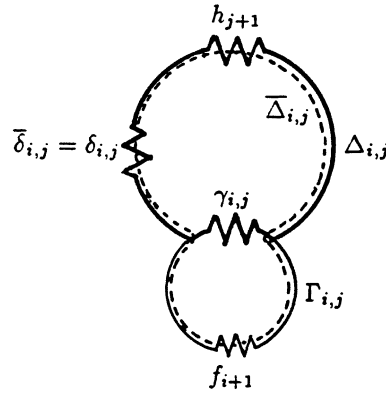

FIG. 3.   *The trees* $\overline{T}_{i+1,\,j} = T_{i,\,j} \cup \{f_{i+1}\} \setminus \{\gamma_{i,\,j}\}$, $\overline{T}_{i,\,j+1} = T_{i,\,j} \cup \{h_{j+1}\} \setminus \{\delta_{i,\,j}\}$ *and* $\overline{T}_{i+1,\,j+1} = \overline{T}_{i+1,\,j} \cup \{h_{j+1}\} \setminus \{\overline{\delta}_{i,\,j}\}$. *Edges* $\gamma_{i,\,j}$, $\delta_{i,\,j}$ *and* $\overline{\delta}_{i,\,j}$ *are the longest edges in the circuits* $\Gamma_{i,\,j}$, $\Delta_{i,\,j}$, *and* $\overline{\Delta}_{i,\,j}$, *respectively;* $\Delta_{i,\,j}$ *is the boldly drawn circuit and* $\overline{\Delta}_{i,\,j}$ *is the dashed circuit.*

Therefore

$$\psi(|\delta_{i,j}|) \geq \max_{e \in \Delta_{i,j} \cup \Gamma_{i,j}} \psi(|e|) \geq \max_{e \in \overline{\Delta}_{i,j}} \psi(|e|) = \psi(|\overline{\delta}_{i,j}|).$$

Thus (2.13) holds. [In fact this argument shows that if $\delta_{i,j} \in \Delta_{i,j} \setminus \Gamma_{i,j}$, then $\psi(|\overline{\delta}_{i,j}|) = \psi(\delta_{i,j})$, but we do not need this fact.] $\square$

We close this section by stating that the degree of an MST is "uniformly bounded."

LEMMA 4. *Let ψ be strictly increasing. Then there is a finite constant D_d, which depends only on d, such that, for any MST T on a finite subset of $\mathbb{R}^d$, T has maximum vertex degree bounded by D_d.*

The same proof works for all strictly increasing weight functions. See Lemma 4 of Aldous and Steele (1992) or Talagrand [(1995), proof of Lemma 11.3.1].

3. Moment estimates. In Section 4 we represent $M(X_1^{(n)}, \ldots, X_n^{(n)}) - EM(X_1^{(n)}, \ldots, X_n^{(n)})$ as the sum of martingale differences and then we approximate (in the sense of Proposition 4) the martingale differences by some random variables which depend only on the "local configuration of $\{X_1^{(1)}, \ldots, X_n^{(n)}\}$." This can be justified by the estimates of moments in this section.

We denote the (random) set of a Poisson point process of density λ in $\mathbb{R}^d$ by $\mathscr{P}(\lambda)$ and the corresponding probability measure by $\mu^{P(\lambda)}$. If we are interested in the process only on Q_n, we denote the Poisson points and the corresponding probability measure by $\mathscr{P}_n(\lambda)$ and $\mu^{P_n(\lambda)}$, respectively. Thus $\mathscr{P}_n(\lambda) = \mathscr{P}(\lambda) \cap Q_n$.

In the case $\lambda = 1$ we simplify the notations $\mathscr{P}(1)$, $\mu^{P(1)}$, $\mathscr{P}_n(1)$ and $\mu^{P_n(1)}$ to $\mathscr{P}$, μ^P, $\mathscr{P}_n$ and μ^{P_n}, respectively.

LEMMA 5. *There exist constants $0 < C_1, C_2 < \infty$ such that for all rectangles $B[x, y]$ and for $1/2 \leq \lambda \leq 2$,*

(3.1) $\quad \mu^{P(\lambda)}\{\text{there is no separating set of width } \delta \text{ of Poisson points} \\ \text{for the rectangle } B[x, y] \text{ in the annulus } A(B[x, y]; 1, \delta)\} \\ \leq C_1 \gamma^{2(d-1)} \exp(-C_2 \delta),$

where $\gamma = \left[\sum (y_i - x_i + 2 + 2\delta)^2 \right]^{1/2}$ is the diameter of the rectangle $\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$. Moreover, for $B[x, y] \cap Q_n \neq \phi$,

(3.2) $\quad \mu^{P_n(\lambda)} \{\text{there is no } n\text{-separating set of width } \delta \text{ of } \mathscr{P}_n(\lambda)\text{-points} \\ \text{for the rectangle } B[x, y] \text{ in the annulus } A(B[x, y]; 1, \delta)\} \\ \leq C_1 \gamma^{2(d-1)} \exp(-C_2 \delta).$

PROOF.   Partition $\partial \Pi(x_i - 1, y_i + 1)$ into at most $C_3[\gamma - 2\delta]^{d-1}$ $(d-1)$-dimensional cubes of edge length at most $1/[3(d-1)^{1/2}]$. Similarly partition $\partial \Pi[x_i - 1 - \delta, y_i + 1 + \delta]$ into $C_3 \gamma^{d-1}$ such cubes. Assume there exists a line segment from $\partial \Pi(x_i - 1, y_i + 1)$ to $\partial \Pi[x_i - 1 - \delta, y_i + 1 + \delta]$ such that no point from the Poisson process in the annulus $\Pi[x_i - 1 - \delta, x_i + 1 + \delta] \setminus \Pi(x_i - 1, y_i + 1)$ lies within distance $1/3$ of the segment. Let the segment run from a point p_1 in a $(d-1)$-dimensional cube F_1 on $\partial \Pi(x_i - 1, y_i + 1)$ to a point p_2 in a $(d-1)$-dimensional cube F_2 on $\partial \Pi[x_i - 1 - \delta, y_i + 1 + \delta]$. Let f_i be the center of F_i and note that $|f_i - p_i| \le 1/6$. Then no Poisson point π in the annulus lies within distance $1/6$ from the line segment connecting $f_1 \in F_1$ to $f_2 \in F_2$, because $|\pi - \alpha p_1 - (1 - \alpha) p_2| \le |\pi - \alpha f_1 - (1 - \alpha) f_2| + 1/6$, $0 \le \alpha \le 1$. For each f_1 and f_2,

$$\mu^{P(\lambda)}\{\text{there is no Poisson point in the annulus within}$$
$$\text{distance } 1/6 \text{ of the segment } (f_1, f_2)\}$$

$$\le C_4 \exp(-C_2 \delta).$$

The number of choices for each of F_1 and F_2 is at most $C_5 \gamma^{d-1}$ so that the probability in (3.1) is at most

$$C_1 \gamma^{2(d-1)} \exp(-C_2 \delta).$$

This proves (3.1) and the proof of (3.2) is essentially the same.   □

We shall also choose m vertices uniformly in the cube Q_n and independently of each other. We denote these random points by $X_1^{(n)}, \ldots, X_m^{(n)}$ and denote the set of these random points by $\mathscr{U}_n(m)$. The corresponding probability measure governing $\mathscr{U}_n(m)$ is denoted by $\mu^{U_n(m)}$. In the case $m = n$ we simplify the notations $\mathscr{U}_n(m)$ and $\mu^{U_n(m)}$ to $\mathscr{U}_n$ and μ^{U_n}, respectively.

LEMMA 6.   *There exist constants $0 < C_1, C_2, C_6, C_7 < \infty$ such that for all rectangles $B[x, y]$ with $B[x, y] \cap Q_n \ne \phi$ and for all $3n/4 \le m \le 5n/4$*

(3.3)
$$\mu^{U_n(m)}\{\text{there is no } n\text{-separating set } \mathscr{S} \subset \mathscr{U}_n(m) \text{ for the}$$
$$\text{rectangle } B[x, y] \text{ in the annulus } A(B[x, y]; 1, \delta)\}$$

$$\le C_1 \gamma^{2(d-1)} \exp(-C_2 \delta) + C_6 \exp(-C_7 n),$$

where, as in Lemma 5, $\gamma = \left[\sum (y_i - x_i + 2 + 2\delta)^2\right]^{1/2}$ is the diameter of the rectangle $\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$.

PROOF.   Choose points in Q_n according to a Poisson process with density $1/2$. If the Poisson process has l points in Q_n with $l \le m$, then add $m - l$ independent points, chosen uniformly in Q_n. If $l > m$, then choose m new independent points uniformly in Q_n. It is easy to see that the resulting set of m points in Q_n has the distribution $\mu^{U_n(m)}$. Therefore, since increasing the

number of points can only create more separating sets, the left-hand side of (3.3) is at most

$$
(3.4) \quad \mu^{P(1/2)}\{l > 3n/4\} + \mu^{P(1/2)}\{\text{there is no } n\text{-separating set of}
$$
$$
\mathscr{P}(1/2)\text{-points for the rectangle } B[x, y]
$$
$$
\text{in the annulus } A(B[x, y]; 1, \delta)\}.
$$

Now

$$
(3.5) \qquad\qquad \mu^{P(1/2)}\{l > 3n/4\} \le C_6 \exp(-C_7 n).
$$

So (3.3) follows from (3.4), (3.5) and (3.2). □

We shall apply Proposition 3 when $\mathscr{A}$ is the set of Poisson points in $C(x) \cap Q_n$ and $\mathscr{B}$ is the set of Poisson points in $Q_n \setminus C(x)$. In this case we shall want to approximate $D(C(x) \cap Q_n \cap \mathscr{P}(\lambda), (Q_n \setminus C(x)) \cap \mathscr{P}(\lambda))$ by $D(C(x) \cap Q_n \cap \mathscr{P}(\lambda), B' \cap \mathscr{P}(\lambda))$, where (for a suitable L') $B' = (\Pi[x_i - L', x_i + 1 + L'] \cap Q_n) \setminus C(x)$. We shall also want to approximate $D(\{x\}, (Q_n \setminus \{x\}) \cap \mathscr{P}(\lambda))$ by $D(\{x\}, B'' \cap \mathscr{P}(\lambda))$, where $B'' = (\Pi[x_i - L'', x_i + L''] \cap Q_n) \setminus \{x\}$. For these situations it is convenient to have some abbreviated notation. For a locally finite set $\mathscr{W}$ of points in $\mathbb{R}^d$ we define

$$
\widetilde{D}_L(x; \mathscr{W}) = D(C(x) \cap \mathscr{W}, (\Pi[x_i - L, x_i + 1 + L] \setminus C(x)) \cap \mathscr{W}),
$$
$$
D_L(x; \mathscr{W}) = D(\{x\}, (\Pi[x_i - L, x_i + L] \setminus \{x\}) \cap \mathscr{W}).
$$

The following lemma provides a simple estimate for the various functions D.

LEMMA 7.   *Let ψ be strictly increasing. Let $\mathscr{A}$ be a finite subset of the rectangle $B[x, y]$ with cardinality $|\mathscr{A}|$ and let $\mathscr{B}$ be a finite subset of $\mathbb{R}^d$ which contains a separating set $\mathscr{S}$ for $B[x, y]$ in the annulus $A(B[x, y]; 1, \delta)$. Then*

$$
(3.6) \qquad\qquad |D(\mathscr{A}, (\mathscr{B} \setminus \mathscr{A}))| \le C_8 |\mathscr{A}| \psi(\gamma + 2d^{1/2}),
$$

where γ is the diameter of $\Pi[x_i - 1 - \delta, y_i + 1 + \delta]$. Similarly, if $\mathscr{W}$ is a locally finite set of points in $\mathbb{R}^d$ which contains a separating set for $C(x)$ in the annulus $A(C(x); 1, \delta)$, then

$$
(3.7) \qquad |\widetilde{D}_L(x; \mathscr{W})| \le C_8 |C(x) \cap \mathscr{W}| \psi(((2\delta + 5) \wedge (2L + 1)) d^{1/2}),
$$

and if $\mathscr{W}$ contains a separating set for $\{x\}$ in the annulus $A(\{x\}; 1, \delta)$, then

$$
(3.8) \qquad\qquad |D_L(x; \mathscr{W})| \le C_8 \psi(((2\delta + 4) \wedge (2L)) d^{1/2}).
$$

Finally, if $\mathscr{W} \cap Q_n$ contains an n-separating set for $C(x)$ in the annulus $A(C(x); 1, \delta)$, then

$$
(3.9) \qquad
\begin{aligned}
&|D(C(x) \cap \mathscr{W} \cap Q_n, (\mathscr{W} \cap Q_n) \setminus C(x))| \\
&\le C_8 |C(x) \cap \mathscr{W} \cap Q_n| \psi(((2\delta + 5) \wedge n^{1/d}) d^{1/2}).
\end{aligned}
$$

PROOF.   We begin with (3.6). It is trivial that

$$M(\mathscr{A} \cup \mathscr{B}) \leq M(\mathscr{B}) + |\mathscr{A}|\psi(\gamma)$$

because we can connect each vertex of $\mathscr{A}$ to an MST on $\mathscr{B}$ by connecting it to some vertex of $\mathscr{S} \subset \mathscr{B}$. Conversely, if we take an MST on $\mathscr{A} \cup \mathscr{B}$, then we can obtain a spanning tree on $\mathscr{B}$ by deleting all the edges incident to any vertex in $\mathscr{A} \setminus \mathscr{B}$ (by Lemma 4 we delete at most $D_d|\mathscr{A}|$ edges) and by reconnecting the at most $(D_d|\mathscr{A}|+1)$ components, which by Lemma 3 all have a vertex in $\Pi[x_i - 2 - \delta, y_i + 2 + \delta]$. Note that the diameter of the rectangle $\Pi[x_i - 2 - \delta, y_i + 2 + \delta]$ is at most $\gamma + 2d^{1/2}$, so

$$M(\mathscr{B}) \leq M(\mathscr{A} \cup \mathscr{B}) + D_d|\mathscr{A}|\psi(\gamma + 2d^{1/2}).$$

This proves (3.6). Equations (3.7)–(3.9) can be proven in the same way if one takes into account that any pair of points in $\Pi[x_i-L, x_i+1+L]$, $\Pi[x_i-L, x_i+L]$ or Q_n are at most a distance $(2L + 1)d^{1/2}$, $(2L)d^{1/2}$ or $n^{1/d}d^{1/2}$, respectively, apart. $\square$

LEMMA 8.   *Assume that ψ is strictly increasing and that*

$$(3.10) \qquad\qquad \lim_{x \to \infty} \frac{1}{x} \log \psi(x) = 0.$$

Then, for each $q > 0$, there exists a constant $C(q)$ such that

$$(3.11) \qquad \begin{aligned} &\mu^{P(\lambda)}\{|\widetilde{D}_L(x; \mathscr{P}(\lambda))|^q\} \leq C(q), \\ &\mu^{P(\lambda)}\{|D_L(x; \mathscr{P}(\lambda))|^q\} \leq C(q) \quad \textit{uniformly in } L, x, \ 1/2 \leq \lambda \leq 2, \end{aligned}$$

$$(3.12) \qquad \begin{aligned} &\mu^{P_n(\lambda)}\{|D(C(x) \cap Q_n \cap \mathscr{P}_n(\lambda), (Q_n \setminus C(x)) \cap \mathscr{P}_n(\lambda))|^q\} \\ &\qquad \leq C(q) \quad \textit{uniformly in } n, \ x \in Q_n, \ 1/2 \leq \lambda \leq 2, \end{aligned}$$

$$(3.13) \qquad \begin{aligned} &\mu^{P_n(\lambda)}\{|\widetilde{D}_L(x; \mathscr{P}_n(\lambda))|^q\} \leq C(q) \quad \textit{uniformly in } n, \\ &\qquad\qquad\qquad\qquad\qquad\qquad x \in Q_n, \ L, \ 1/2 \leq \lambda \leq 2, \end{aligned}$$

$$(3.14) \qquad \begin{aligned} &\mu^{U_n(m)}\{|D(\{X_m^{(n)}\}, \mathscr{U}_n(m-1))|^q\} \leq C(q) \quad \textit{uniformly in } n, \\ &\qquad\qquad\qquad\qquad\qquad\qquad 3n/4 \leq m \leq 5n/4, \end{aligned}$$

$$(3.15) \qquad \begin{aligned} &\mu^{U_n(m)}\{|D_L(X_m^{(n)}; \mathscr{U}_n(m-1))|^q\} \leq C(q) \quad \textit{uniformly in } n, \\ &\qquad\qquad\qquad\qquad\qquad\qquad 3n/4 \leq m \leq 5n/4 \textit{ and } L. \end{aligned}$$

PROOF.   Consider a Poisson field $\mathscr{P}(\lambda)$, $1/2 \leq \lambda \leq 2$, in $\mathbb{R}^d$. Let

$$\begin{aligned} \widetilde{\delta} = \inf\{\delta: \ &\exists \text{ separating set } \mathscr{S} \subset \mathscr{P}(\lambda) \text{ for } C(x) \\ &\text{in the annulus } A(C(x); 1, \delta)\}. \end{aligned}$$

Then, by (3.7),

$$(3.16) \quad |\widetilde{D}_L(x;\mathscr{P}(\lambda))| \leq C_8|C(x) \cap \mathscr{P}(\lambda)|\psi(((2\widetilde{\delta}+5) \wedge (2L+1))d^{1/2}).$$

Therefore

$$\mu^{P(\lambda)}\{|\widetilde{D}_L(x;\mathscr{P}(\lambda))|^q\}$$

$$= \mu^{P(\lambda)}\{|\widetilde{D}_L(0;\mathscr{P}(\lambda))|^q\} \quad \text{(by translation invariance)}$$

$$\leq C_9[\mu^{P(\lambda)}\{|C(0) \cap \mathscr{P}(\lambda)|^{2q}\}]^{1/2}[\mu^{P(\lambda)}\{\psi((2\widetilde{\delta}+5)d^{1/2})^{2q}\}]^{1/2}$$

$$\leq C_{10}[\mu^{P(\lambda)}\{\psi((2\widetilde{\delta}+5)d^{1/2})^{2q}\}]^{1/2}.$$

Now by Lemma 5 and (3.10), $\mu^{P(\lambda)}\{\psi((2\widetilde{\delta}+5)d^{1/2})^{2q}\}$ is bounded for $1/2 \leq \lambda \leq 2$. Thus we get the first part of (3.11). Almost the same argument works for the second part of (3.11). Similarly, for (3.12) and (3.13) we merely have to replace $\widetilde{\delta}$ by

$$(3.17) \quad \overline{\delta} = \inf\{\delta \colon \exists \; n\text{-separating set } \mathscr{S} \subset \mathscr{P}_n(\lambda) \text{ for } C(x)$$
$$\text{in the annulus } A(C(x);1,\delta)\}$$

and use (3.6) and (3.7). For (3.14) and (3.15) we use Lemma 6 instead of Lemma 5. □

We remind the reader that μ^{P_n} is the distribution of $\mathscr{P}_n = \mathscr{P}_n(1) = \mathscr{P}(1) \cap Q_n = \mathscr{P}(1) \cap [0, n^{1/d}]^d$.

LEMMA 9. *Let ψ be strictly increasing. For all $\varepsilon > 0$, there exists an $L = L(\varepsilon)$ such that for $x \in [L, n^{1/d} - L - 1]^d$ and $L' \geq L$,*

$$(3.18) \quad \mu^{P_n}\{|D(C(x) \cap \mathscr{P}_n(1), (Q_n \setminus C(x)) \cap \mathscr{P}_n(1)) - \widetilde{D}_{L'}(x;\mathscr{P}_n(1))| \geq \varepsilon\} \leq \varepsilon.$$

PROOF. Choose δ so large that

$$\mu^P\{\text{there is no separating set for } C(x) \text{ in } A(C(x);1,\delta)\} \leq \varepsilon/3.$$

Such a δ exists by Lemma 5. Also choose N so that

$$\mu^P\{C(x) \text{ contains more than } N \text{ Poisson points}\} \leq \frac{\varepsilon}{3}.$$

On the event

$$(3.19) \quad \{\exists \text{ separating set for } C(x) \text{ in } A(C(x);1,\delta), \; C(x)$$
$$\text{contains at most } N \text{ Poisson points}\}$$

we have, by (3.16),

$$|\widetilde{D}_L(x;\mathscr{P}(1))| \leq C_8 N\psi((2\delta+5)d^{1/2})$$

and, by Proposition 3,

$$\widetilde{D}_L(x;\mathscr{P}(1)) \text{ is increasing in } L \text{ for } L \geq \delta + 2.$$

Therefore, $\widetilde{D}_L(x; \mathscr{P}(1))$ converges to a finite limit a.e. μ^P and one can choose $L = L(\varepsilon) \geq \delta + 2$ so large that

$$(3.20) \quad \mu^P\left\{(3.19) \text{ fails or } \sup_{L' \geq L(\varepsilon)} |\widetilde{D}_{L'}(x; \mathscr{P}(1)) - \widetilde{D}_{L(\varepsilon)}(x; \mathscr{P}(1))| \geq \varepsilon/2\right\} \leq \varepsilon.$$

Note that this estimate is uniform in x, by the translation invariance of μ^P.

If $x \in [L, n^{1/d} - L - 1]^d$, $L', L'' \geq L$ and $Q_n \subset \Pi[x_i - L'', x_i + 1 + L'']$, then by Proposition 3 one has on the event

$$(3.21) \quad \left\{(3.19) \text{ holds and } \sup_{L' \geq L} |\widetilde{D}_{L'}(x; \mathscr{P}(1)) - \widetilde{D}_L(x; \mathscr{P}(1))| < \varepsilon/2\right\}$$

that

$$0 \leq D(C(x) \cap \mathscr{P}_n(1), (Q_n \setminus C(x)) \cap \mathscr{P}_n(1)) - \widetilde{D}_{L'}(x; \mathscr{P}_n(1))$$

$$= D(C(x) \cap \mathscr{P}(1), (Q_n \setminus C(x)) \cap \mathscr{P}(1)) - \widetilde{D}_{L'}(x; \mathscr{P}_n(1))$$

$$(3.22) \quad \leq D(C(x) \cap \mathscr{P}(1), (Q_n \setminus C(x)) \cap \mathscr{P}(1)) - \widetilde{D}_L(x; \mathscr{P}_n(1))$$

$$= D(C(x) \cap \mathscr{P}(1), (Q_n \setminus C(x)) \cap \mathscr{P}(1)) - \widetilde{D}_L(x; \mathscr{P}(1))$$

$$\leq \widetilde{D}_{L''}(x; \mathscr{P}(1)) - \widetilde{D}_L(x; \mathscr{P}(1)) < \varepsilon.$$

So (3.18) follows from (3.20), (3.21) and (3.22). □

We next want to use Proposition 3 when $\mathscr{A}$ is a single (random) point and $\mathscr{B}$ is the collection of Poisson points in Q_n. That is, we want to see what the influence of adding a single point is on the weight of an MST. Choose $\mathscr{P}(\lambda)$, the set of Poisson points in $\mathbb{R}^d$ of density λ, and then choose a point Y_n, independent of $\mathscr{P}(\lambda)$, uniformly in Q_n. We denote the corresponding probability by $\mu^{P(\lambda), Y_n}$. Similarly, we denote by $\mu^{P_n(\lambda), Y_n}$ the joint distribution of $\mathscr{P}_n(\lambda) = \mathscr{P}(\lambda) \cap Q_n$ and Y_n.

LEMMA 10. *Let ψ be strictly increasing. For all $\varepsilon > 0$, there exists a $K = K(\varepsilon)$ and an $n_1 = n_1(\varepsilon)$ such that uniformly for $1/2 \leq \lambda \leq 2$, $n \geq n_1$,*

$$\mu^{P_n(\lambda), Y_n}\{|D(\{Y_n\}, (Q_n \setminus \{Y_n\}) \cap \mathscr{P}_n(\lambda)) - D_K(Y_n; \mathscr{P}_n(\lambda))| \geq \varepsilon\}$$

$$(3.23) \quad = \mu^{P(\lambda), Y_n}\{|D(\{Y_n\}, (Q_n \setminus \{Y_n\}) \cap \mathscr{P}(\lambda)) - D_K(Y_n; \mathscr{P}_n(\lambda))| \geq \varepsilon\}$$

$$\leq \varepsilon$$

and

$$(3.24) \quad \mu^{P(\lambda), Y_n}\{\text{there is no separating set for } \{Y_n\}$$
$$\text{in } A(\{Y_n\}; 1, K) \cap P(\lambda)\} \leq \varepsilon/4.$$

PROOF. We give the proof only for $\psi(y) = \psi_\alpha(y) = y^\alpha$. When $\lambda = 1$, then (3.23) is very similar to (3.18) except that in (3.18) the $\{Y_n\}$ is replaced by $C(x)$, $x \in [L, n^{1/d} - 1 - L]^d$. However, the basic estimate is the same, except

that an extra term comes in to take care of the case when Y_n is too close to ∂Q_n. Specifically, choose δ, independent of $1/2 \leq \lambda \leq 2$, so large that for all $n \geq 1$,

$$\mu^{P(\lambda), Y_n}\{\text{there is no separating set for } \{Y_n\} \text{ in } A(\{Y_n\}; 1, \delta)\} \leq \varepsilon/4.$$

This can be done by virtue of Lemma 5 and the translation invariance of $\mu^{P(\lambda)}$. The inequality (3.24) will hold if we take $K \geq \delta$. On the event

$$(3.25) \qquad \{\exists \text{ separating set for } \{Y_n\} \text{ in } A(\{Y_n\}; 1, \delta)\}$$

we have, by Lemma 7, $|D_K(Y_n; \mathscr{P}(\lambda))| \leq C_8 \psi((2\delta + 4)d^{1/2})$ for $K \geq \delta + 1$ and, by Proposition 3, $D_K(Y_n; P(\lambda))$ is increasing in K for $K \geq \delta + 2$. We claim that one can choose $K(\varepsilon) \geq \delta + 2$ and $n_1(\varepsilon)$ so that uniformly for $1/2 \leq \lambda \leq 2$, $n \geq n_1$,

$$\mu^{P(\lambda), Y_n}\Big\{Y_n \notin [K, n^{1/d} - K]^d \text{ or } (3.25) \text{ fails or}$$
$$\sup_{K' \geq K(\varepsilon)} |D_{K'}(Y_n; \mathscr{P}(\lambda)) - D_{K(\varepsilon)}(Y_n; \mathscr{P}(\lambda))| \geq \varepsilon/2\Big\} \leq \varepsilon.$$

Indeed, since, for any fixed x, $\lambda^{-1/d}\mathscr{P}(1) + x$ has the same distribution as $\mathscr{P}(\lambda)$, and $\psi(\lambda^{-1/d}y) = \lambda^{-\alpha/d}\psi(y)$, we have

$$\mu^{P(\lambda), Y_n}\Bigg\{\sup_{K' \geq K} |D_{K'}(Y_n; \mathscr{P}(\lambda)) - D_K(Y_n; \mathscr{P}(\lambda))| \geq \varepsilon/2\Bigg\}$$
$$= \mu^{P, Y_n}\Bigg\{\sup_{K' \geq K\lambda^{1/d}} |D_{K'}(\mathbf{0}; \mathscr{P}) - D_{K\lambda^{1/d}}(\mathbf{0}; \mathscr{P})| \geq (\varepsilon/2)\lambda^{\alpha/d}\Bigg\}.$$

Choose $K(\varepsilon)2^{-1/d} \geq \delta + 2$ so that

$$\mu^{P, Y_n}\Bigg\{\sup_{K' \geq K(\varepsilon)2^{-\alpha/d}} |D_{K'}(\mathbf{0}; \mathscr{P}) - D_{K(\varepsilon)2^{-\alpha/d}}(\mathbf{0}; \mathscr{P})| \geq (\varepsilon/4)2^{-\alpha/d}\Bigg\} \leq \varepsilon/4.$$

Finally, we choose $n_1(\varepsilon)$ so that for $n \geq n_1(\varepsilon)$,

$$P\{Y_n \notin [K, n^{1/d} - K]^d\} \leq \varepsilon/4.$$

The rest of the proof is the same as that of Lemma 9. $\quad\square$

Next we prove an analogue of Lemma 10 when $\mathscr{P}_n(\lambda)$ is replaced by $\mathscr{U}_n(m)$, a configuration of m i.i.d. points, each one uniformly distributed in Q_n. Let us write $X_1^{(1)}, \ldots, X_m^{(n)}$ for m such points in Q_n.

LEMMA 11. *Let ψ be strictly increasing. For all $\varepsilon > 0$, there exists a $K = K(\varepsilon)$ and an $n_2 = n_2(\varepsilon)$ such that uniformly for $n \geq n_2$, $n - n^{3/4} \leq m \leq n + n^{3/4}$,*

$$(3.26) \quad \mu^{U_n(m)}\{|D(\{X_m^{(n)}\}, (Q_n \setminus \{X_m^{(n)}\}) \cap \mathscr{U}_n(m)) - D_K(X_m^{(n)}; \mathscr{U}_n(m))| \geq \varepsilon\} \leq \varepsilon.$$

PROOF. Consider two independent Poisson variables N_1 and W with means $n - 2n^{3/4}$ and $4n^{3/4}$, respectively. Then $N_2 := N_1 + W$ is a Poisson variable with mean $n + 2n^{3/4}$. We now choose in succession N_2 i.i.d. uniform points on Q_n. The first N_1 of these points forms a realization of $\mathscr{P}(1 - 2n^{-1/4})$ on Q_n and the total forms a realization of $\mathscr{P}(1 + 2n^{-1/4})$ on Q_n. On the event

$$(3.27) \qquad \{N_1 \le n - n^{3/4} < n + n^{3/4} \le N_2\}$$

we further take the first m points as our realization of $\mu^{U_n(m)}$, say, $X_1^{(n)}, \ldots, X_m^{(n)}$ (recall that $n - n^{3/4} \le m \le n + n^{3/4}$). If (3.27) is not the case, we pick m new i.i.d. uniform points on Q_n to get a realization of $\mu^{U_n(m)}$. Finally let Y_n be another uniform point on Q_n, independent of N_1, W and the $X_i^{(n)}$.

Let $K = K(\varepsilon)$ and $n_1 = n_1(\varepsilon)$ be such that (3.23) holds and such that (3.24) holds even with K replaced by $K - 2$. Assume that for some given $n \ge n_1$, (3.27) occurs and

$$(3.28) \qquad |D(\{Y_n\}, \mathscr{P}_n(1 + 2n^{-1/4})) - D_K(\{Y_n\}; \mathscr{P}_n(1 + 2n^{-1/4}))| < \varepsilon.$$

Assume further that

$$(3.29) \qquad \begin{array}{l} \text{none of the } W \text{ points numbered } N_1, \ldots, N_2 \\ \text{fall in } \Pi[(Y_n)_i - K - 2, (Y_n)_i + K + 2] \end{array}$$

and that

$$(3.30) \qquad \exists \text{ separating set in } \mathscr{P}_n(1 - 2n^{-1/4}) \text{ for } \{Y_n\} \text{ in } A(\{Y_n\}; 1, K - 2).$$

We claim that, if (3.27)–(3.30) hold, then

$$(3.31) \qquad |D(\{Y_n\}, \mathscr{U}_n(m - 1)) - D_K(\{Y_n\}; \mathscr{U}_n(m - 1))| \le \varepsilon$$

for $n - n^{3/4} \le m \le n + n^{3/4}$. Indeed, by (3.27) and (3.29), for each $N_1 \le m \le N_2$,

$$(3.32) \qquad \begin{aligned} & D(\{Y_n\}, \mathscr{U}_n(m - 1) \cap \Pi[(Y_n)_i - K, (Y_n)_i + K]) \\ & = D(\{Y_n\}, \mathscr{P}_n(1 + 2n^{-1/4}) \cap \Pi[(Y_n)_i - K, (Y_n)_i + K]). \end{aligned}$$

Moreover, by (3.30) and Proposition 3, as we keep adding points,

$$(3.33) \qquad D(\{Y_n\}, \mathscr{U}_n(r)) \text{ increases in } r, \quad N_1 \le r \le N_2.$$

Therefore, if (3.27)–(3.30) hold, then for $n \ge n_1$, $n - n^{3/4} \le m \le n + n^{3/4}$,

$$0 \le D(\{Y_n\}, \mathscr{U}_n(m - 1)) - D_K(\{Y_n\}; \mathscr{U}_n(m - 1))$$

$$\text{[by (3.30) and Proposition 3]}$$

$$\le D(\{Y_n\}, \mathscr{P}_n(1 + 2n^{-1/4})) - D_K(\{Y_n\}; \mathscr{U}_n(m - 1))$$

$$\text{[by (3.27) and (3.33)]}$$

$$= D(\{Y_n\}, \mathscr{P}_n(1 + 2n^{-1/4})) - D_K(\{Y_n\}; \mathscr{P}_n(1 + 2n^{-1/4})) \qquad \text{[by (3.32)]}$$

$$< \varepsilon.$$

This proves our claim (3.31). Next we show that the probability that (3.27)–(3.30) hold is close to 1. It is trivial that $P\{(3.27) \text{ fails}\} \leq \varepsilon/4$ for large n, say $n \geq n_2$. By Lemma 10, for $n \geq n_1(\varepsilon/4)$, $P\{(3.28) \text{ fails}\} \leq \varepsilon/4$. Also

$$P\{(3.29) \text{ fails}\} \leq (EW)\frac{(2K+4)^d}{n} = (2K+4)^d 4n^{-1/4}.$$

We assumed that (3.24) holds with K replaced by $K-2$ and we may further assume that $n_2 \geq n_1$ is so large that, for $n \geq n_2$, $P\{(3.29) \text{ or } (3.30) \text{ fails}\} \leq \varepsilon/2$. Then for $n \geq n_2$,

$$(3.34) \qquad\qquad P\{(3.27)\text{--}(3.30) \text{ hold}\} \geq 1 - \varepsilon.$$

Since the joint distribution of Y_n and $\mathscr{U}_n(m-1)$ is the same as that of $X_m^{(n)}$ and $\mathscr{U}_n(m-1)$, (3.31) and (3.34) together imply

$$\mu^{U_n(m)}\{|D(\{X_m^{(n)}\}, \mathscr{U}_n(m-1)) - D_K(X_m^{(n)}; \mathscr{U}_n(m-1))| \geq \varepsilon\} \leq \varepsilon.$$

The inequality (3.26) follows because $D(\{X_m^{(n)}\}, (Q_n \setminus \{X_m^{(n)}\}) \cap \mathscr{U}_n(m)) = D(\{X_m^{(n)}\}, \mathscr{U}_n(m-1))$ and $D_K(X_m^{(n)}; \mathscr{U}_n(m-1)) = D_K(X_m^{(n)}; \mathscr{U}_n(m))$. □

4. Central limit theorem for an MST. We order the vertices v in $\mathbb{Z}^d \cap [0, n^{1/d}]^d$ in some way, say lexicographically, as $v(1), \ldots, v((\lfloor n^{1/d} \rfloor + 1)^d)$ and define $\mathscr{F}_k$ by

$$\mathscr{F}_k = \sigma\left(\mathscr{P}_n(1) \cap \left[\bigcup_{i \leq k} C(v(i))\right]\right)$$

($\mathscr{F}_0$ is the trivial σ-field). Also define $\widetilde{M}_n$ and Δ_k by

$$(4.1) \qquad \widetilde{M}_n = \min\left\{\sum_{e \in T'} \psi(|e|) \colon T' \text{ a spanning tree on } \mathscr{P}_n(1)\right\},$$

$$\Delta_k = \mu^{P_n}\{\widetilde{M}_n \mid \mathscr{F}_k\} - \mu^{P_n}\{\widetilde{M}_n \mid \mathscr{F}_{k-1}\}.$$

From now on we assume that ψ is strictly increasing, continuous, and satisfies the growth condition (3.10).

LEMMA 12. *The quantity $\widetilde{M}_n$ is a function of the independent random variables $\mathscr{A}_1 = \mathscr{P}_n(1) \cap Cv(1), \ldots, \mathscr{A}_l = \mathscr{P}_n(1) \cap Cv(l)$, $l = (\lfloor n^{1/d} \rfloor + 1)^d$. Moreover,*

$$(4.2) \qquad\qquad \widetilde{M}_n - E\widetilde{M}_n = \sum_{k=1}^{l} \Delta_k$$

and

$$(4.3) \quad \begin{aligned} \Delta_k = \int \mu^{P_n}(da_k, da_{k+1}, \dots, da_l) \\ \times \left[D\left(\mathscr{A}_k, \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - D\left(a_k, \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) \right], \end{aligned}$$

where $\mu^{P_n}(da_k, \dots, da_l)$ is short for $\mu^{P_n}(\mathscr{A}_k \in da_k, \dots, \mathscr{A}_l \in da_l)$.

PROOF. Clearly $\widetilde{M}_n$ is a function of $\mathscr{A}_1, \dots, \mathscr{A}_l$. In fact, with the appropriate topology on the range of $(\mathscr{A}_1, \dots, \mathscr{A}_l)$, $\widetilde{M}_n$ is a Borel function of $\mathscr{A}_1, \dots, \mathscr{A}_l$. Equation (4.2) is then immediate from the fact that $\widetilde{M}_n$ is $\mathscr{F}_l$-measurable. Moreover, by definition,

$$\begin{aligned} \Delta_k = \int \mu^{P_n}(da_k, da_{k+1}, \dots, da_l) \\ \times \left[\widetilde{M}_n\left(\left[\bigcup_{i \le k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - \widetilde{M}_n\left(\left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i \ge k} a_i \right] \right) \right]. \end{aligned}$$

Now subtract and add $\widetilde{M}_n(\left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right])$ to the expression in square brackets and note that

$$\begin{aligned} \widetilde{M}_n\left(\left[\bigcup_{i \le k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - \widetilde{M}_n\left(\left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) \\ = D\left(\mathscr{A}_k, \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right), \end{aligned}$$

and similarly if $\mathscr{A}_k$ is replaced by a_k. □

Now we write $\Delta_{k,L}$ for the expression in (4.3) when $D(\mathscr{A}_k, \mathscr{B})$ and $D(a_k, \mathscr{B})$ are replaced by $\widetilde{D}_L(v(k); \mathscr{A}_k \cup \mathscr{B})$ and $\widetilde{D}_L(v(k); a_k \cup \mathscr{B})$, respectively. That is,

$$(4.4) \quad \begin{aligned} \Delta_{k,L} = \int \mu^{P_n}(da_k, da_{k+1}, \dots, da_l) \\ \times \left[\widetilde{D}_L\left(v(k); \left[\bigcup_{i \le k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - \widetilde{D}_L\left(v(k); \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i \ge k} a_i \right] \right) \right]. \end{aligned}$$

PROPOSITION 4. *For all $\varepsilon > 0$ there exists an $L = L(\varepsilon)$ and, for each $L' \ge L$, an $n_3 = n_3(L')$ such that for $L' \ge L$, $n \ge n_3$,*

$$(4.5) \quad \frac{1}{n} \sum_k \mu^{P_n}\{|\Delta_k^2 - \Delta_{k,L'}^2|\} \le \varepsilon.$$

PROOF.

$$\Delta_k - \Delta_{k,L'}$$

$$= \int \mu^{P_n}(da_k, \ldots, da_l)$$

(4.6)
$$\times \left[\left\{ D\left(\mathscr{A}_k, \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - \tilde{D}_{L'}\left(v_k; \left[\bigcup_{i\leq k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) \right\} \right.$$

$$\left. - \left\{ D\left(a_k, \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i>k} a_i \right] \right) - \tilde{D}_{L'}\left(v_k; \left[\bigcup_{i<k} \mathscr{A}_i \right] \cup \left[\bigcup_{i\geq k} a_i \right] \right) \right\} \right].$$

By (4.6), (3.12) and (3.13), $|\Delta_k - \Delta_{k,L'}|$ is uniformly [in (n, k, L')] integrable. By (3.18), $|\Delta_k - \Delta_{k,L'}| \to 0$ as $L' \to \infty$ (uniformly in n and k with $v(k) \in [L', n^{1/d} - L' - 1]^d$) in μ^{P_n}-measure, that is,

(4.7)
$$\mu^{P_n}\{|\Delta_k - \Delta_{k,L'}|\} \to 0$$

as $L' \to \infty$ uniformly in n and k with $v(k) \in [L', n^{1/d} - L' - 1]^d$. This is close to what we want. First we note that the number of k for which $v(k) \notin [L', n^{1/d} - L' - 1]^d$ is at most $C_{11}L'n^{(d-1)/d}$ and the sum of $\mu^{P_n}\{|\Delta_k^2 - \Delta_{k,L'}^2|\}$ over those k is therefore at most $C_{12}L'n^{(d-1)/d}$, by (4.3), (4.4), (3.12) and (3.13). Thus, this part of the sum in (4.5) contributes at most $C_{12}L'n^{-1/d}$. For the remaining summands we note that

(4.8)
$$|\Delta_k^2 - \Delta_{k,L'}^2| = |\Delta_k - \Delta_{k,L'}| \cdot |\Delta_k + \Delta_{k,L'}|$$

$$\leq 2K|\Delta_k - \Delta_{k,L'}| \cdot I[|\Delta_k| \leq K, |\Delta_{k,L'}| \leq K]$$

$$+ \Delta_k^2 I[|\Delta_k| > K] + \Delta_{k,L'}^2 I[|\Delta_{k,L'}| > K].$$

By (4.7),

(4.9)
$$2K\mu^{P_n}\{|\Delta_k - \Delta_{k,L'}|\} \to 0$$

as $L' \to \infty$ uniformly in n and k with $v(k) \in [L', n^{1/d} - L' - 1]^d$. Moreover, by (4.3), (4.4), (3.12) and (3.13),

(4.10)
$$\mu^{P_n}\{\Delta_k^2 I[|\Delta_k| > K]\} \leq \frac{1}{K}\mu^{P_n}\{|\Delta_k|^3\} \leq \frac{C_{13}}{K},$$

(4.11)
$$\mu^{P_n}\{\Delta_{k,L'}^2 I[|\Delta_{k,L'}| > K]\} \leq \frac{1}{K}\mu^{P_n}\{|\Delta_{k,L'}|^3\} \leq \frac{C_{13}}{K}.$$

Now, for a given $\varepsilon > 0$, choose K large so that $C_{13}/K \leq \varepsilon/4$. Then choose L large so that $2K\mu^{P_n}\{|\Delta_k - \Delta_{k,L'}|\} \leq \varepsilon/4$ for $L' \geq L$, $v(k) \in [L', n^{1/d} - L' - 1]^d$ and finally for given $L' \geq L$ choose n_3 large so that $C_{12}L'n^{-1/d} \leq \varepsilon/4$ for $n \geq n_3$. The proposition follows from (4.8)–(4.11), together with the fact that the boundary part of the sum in (4.5) contributes at most $C_{12}L'n^{-1/d}$. $\square$

LEMMA 13. *Define ρ_K by*

(4.12)
$$\rho_K = \mu^P\{D_K(0; \mathscr{P}(1))\}.$$

Then

$$(4.13) \qquad\qquad\qquad \rho = \lim_{K \to \infty} \rho_K$$

exists and ρ is finite. Moreover, for $\varepsilon > 0$ and $A > 0$, there exists an $n_4 = n_4(\varepsilon, A)$ such that, for $n \geq n_4$,

$$
\begin{aligned}
(4.14) \qquad \mu^{U_n(n+An^{1/2})}\big\{ &|M(X_1^{(n)}, \ldots, X_s^{(n)}) \\
&- M(X_1^{(n)}, \ldots, X_n^{(n)}) - (s-n)\rho| \geq \varepsilon n^{1/2} \\
&\qquad \text{for some } n - An^{1/2} \leq s \leq n + An^{1/2} \big\} \\
&\leq (1 + 12\varepsilon^4 + 48C(2)\varepsilon + 24A\varepsilon)\varepsilon.
\end{aligned}
$$

PROOF. By the same argument as for (3.18), as $K \to \infty$, $D_K(0; \mathscr{P}(1))$ is eventually bounded and increasing, so $D_K(0; \mathscr{P}(1))$ converges a.e. as $K \to \infty$. Moreover, by (3.11), $D_K(0; \mathscr{P}(1))$ is uniformly integrable, so (4.13) holds and ρ is finite.

Fix K such that

$$|\rho_K - \rho| \leq \varepsilon / [4(2A+1)]$$

and such that (3.26), with ε replaced by ε^6/A, holds for $n \geq n_2(\varepsilon^6/A)$. Now take $t = \lfloor n - An^{1/2} \rfloor$ and note that

$$(4.15) \quad M(X_1^{(n)}, \ldots, X_s^{(n)}) = M(X_1^{(n)}, \ldots, X_t^{(n)}) + \sum_{p=t+1}^{s} D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1)),$$

and

$$
\begin{aligned}
(4.16) \qquad \sup_{n-An^{1/2} \leq s_1 \leq s_2 \leq n+An^{1/2}} &\left| \sum_{p=s_1+1}^{s_2} \big[D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1)) \right. \\
&\qquad\qquad \left. - D_K(X_p^{(n)}; \mathscr{U}_n(p-1)) \big] \right| \\
&\leq \sum_{n-An^{1/2}}^{n+An^{1/2}} \big| D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1)) - D_K(X_p^{(n)}; \mathscr{U}_n(p-1)) \big|.
\end{aligned}
$$

Introduce

$$\square_K(X_p^{(n)}) = D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1)) - D_K(X_p^{(n)}; \mathscr{U}_n(p-1)).$$

Then, by virtue of (3.14), (3.15) and the fact that

$$\mu^{U_n(n+An^{1/2})}\{f\} = \mu^{U_n(p)}\{f\}$$

if f depends only on $X_1^{(n)}, \ldots, X_p^{(n)}$, $p \le n + An^{1/2}$, one has

$$\mu^{U_n(p)}\big\{\big|\square_K(X_p^{(n)})\big|\big\}$$

$$\le \mu^{U_n(p)}\bigg\{\big|\square_K(X_p^{(n)})\big|; \big|\square_K(X_p^{(n)})\big| \le \frac{\varepsilon^6}{A}\bigg\}$$

$$+ 2\mu^{U_n(p)}\bigg\{\big|D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\big|; \big|D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\big| \ge \frac{A}{\varepsilon^3}\bigg\}$$

$$+ 2\mu^{U_n(p)}\bigg\{\big|D_K(X_p^{(n)}; \mathscr{U}_n(p-1))\big|; \big|D_K(X_p^{(n)}; \mathscr{U}_n(p-1))\big| \ge \frac{A}{\varepsilon^3}\bigg\}$$

$$+ \mu^{U_n(p)}\bigg\{\big|\square_K(X_p^{(n)})\big|; \big|\square_K(X_p^{(n)})\big| > \frac{\varepsilon^6}{A}, \big|D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\big| \le \frac{A}{\varepsilon^3},$$

(4.17)
$$\big|D_K(X_p^{(n)}; \mathscr{U}_n(p-1))\big| \le \frac{A}{\varepsilon^3}\bigg\}$$

$$\le \frac{\varepsilon^6}{A} + 2\frac{\varepsilon^3}{A}\mu^{U_n(p)}\big\{D^2(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\big\}$$

$$+ 2\frac{\varepsilon^3}{A}\mu^{U_n(p)}\big\{D_K^2(X_p^{(n)}; \mathscr{U}_n(p-1))\big\} + \frac{\varepsilon^6}{A}\frac{2A}{\varepsilon^3}$$

$$\bigg[\text{by (3.26) with } \varepsilon \text{ replaced by } \frac{\varepsilon^6}{A}\bigg]$$

$$\le \bigg(\frac{\varepsilon^6}{A} + 4\frac{\varepsilon^3}{A}C(2) + 2\varepsilon^3\bigg).$$

So, if $3An^{1/2} \ge 2An^{1/2} + 1$, say for $n \ge n_4(\varepsilon, A)$, then

$$\mu^{U_n(n+An^{1/2})}\bigg\{\sup_{n-An^{1/2}\le s_1\le s_2\le n+An^{1/2}}\bigg|\sum_{p=s_1+1}^{s_2}\square_K(X_p^{(n)})\bigg| \ge \tfrac{1}{4}\varepsilon n^{1/2}\bigg\}$$

$$\le \sum_{n-An^{1/2}}^{n+An^{1/2}}\mu^{U_n(p)}\{|\square_K(X_p^{(n)})|\}/\big(\tfrac{1}{4}\varepsilon n^{1/2}\big)$$

$$\le (12\varepsilon^4 + 48C(2)\varepsilon + 24A\varepsilon)\varepsilon.$$

Thus, it suffices to show

(4.18)
$$\mu^{U_n(n+An^{1/2})}\bigg\{\sup_{|s-n|\le An^{1/2}}\bigg|\sum_{t+1}^{s}[D_K(X_p^{(n)}; \mathscr{U}_n(p-1)) - \rho_K]\bigg| \ge \frac{1}{4}\varepsilon n^{1/2}\bigg\}$$

$$\le \varepsilon.$$

To do this we observe that the distribution of $(\mathscr{U}_n(p-1) - X_p^{(n)}) \cap [-K, K]^d$ converges to that of a Poisson field with mean 1 in $[-K, K]^d$ as $n \to \infty$ and $p/n \to 1$. In particular we have

(4.19)
$$\mu^{U_n(p)}\{D_K(X_p^{(n)}; \mathscr{U}_n(p-1))\} \to \rho_K = \mu^P\{D_K(0; \mathscr{P}(1))\}$$

uniformly in $n - An^{1/2} \le p \le n + An^{1/2}$. We can therefore further replace $[D_K(X_p^{(n)}; \mathcal{U}_n(p-1)) - \rho_K]$ by

$$Z_p := [D_K(X_p^{(n)}; \mathcal{U}_n(p-1)) - \rho_K] - \mu^{U_n(p)}[D_K(X_p^{(n)}; \mathcal{U}_n(p-1)) - \rho_K].$$

It then suffices to show that

$$\mu^{U_n(n+An^{1/2})}\left\{ \sup_{|s-n| \le An^{1/2}} \left| \sum_{t+1}^{s} Z_p \right| \ge \tfrac{1}{8}\varepsilon n^{1/2} \right\} \le \varepsilon.$$

By (3.8) (with $\delta = \infty$), the Z_p are bounded by $C_{14}(K)$ and $\mu^{U_n(p)}\{Z_p\} = 0$. The Z_p are not independent, but we have that for $p_1 \ne p_2$ the joint distribution of $(\mathcal{U}_n(p_1-1) - X_{p_1}^{(n)}) \cap [-K, K]^d$ and $(\mathcal{U}_n(p_2-1) - X_{p_2}^{(n)}) \cap [-K, K]^d$ converges to the distribution of two independent Poisson fields on $[-K, K]^d$ as $n \to \infty$. From this it follows that $EZ_{p_1}Z_{p_2} \to 0$ and

$$\frac{1}{n}E\left(\sum_{t+1}^{s} Z_p \right)^2 \to 0$$

as $n \to \infty$, uniformly in $n - An^{1/2} \le s \le n + An^{1/2}$. Therefore, for fixed $\eta > 0$,

$$n^{-1/2} \sum_{t+1}^{t+j\eta An^{1/2}} Z_p \to 0$$

in $\mu^{U_n(n+An^{1/2})}$-measure, uniformly for $j = 1, 2, \ldots, \lfloor 2/\eta \rfloor$. This proves

$$(4.20) \qquad\qquad n^{-1/2} \sup_{|s-n| \le An^{1/2}} \left| \sum_{t+1}^{s} Z_p \right| \to 0$$

in $\mu^{U_n(n+An^{1/2})}$-measure because

$$\sup_{t+j\eta An^{1/2} \le s \le t+(j+1)\eta An^{1/2}} \left| \sum_{t+1}^{s} Z_p - \sum_{t+1}^{t+j\eta An^{1/2}} Z_p \right| \le \eta An^{1/2} C_{14}(K).$$

Inequality (4.18) and hence (4.14) now follow from (4.20). □

PROOF OF COROLLARY 1. Note that

$$\mu^{U_n(n+1)}\{M(X_1^{(n)}, \ldots, X_{n+1}^{(n)})\} - \mu^{U_n(n)}\{M(X_1^{(n)}, \ldots, X_n^{(n)})\} - \rho$$
$$= \mu^{U_n(n+1)}\{D(\{X_{n+1}^{(n)}\}, \mathcal{U}_n(n))\} - \rho$$
$$= \mu^{U_n(n+1)}\{\Box_K(X_{n+1}^{(n)}) + D_K(X_{n+1}^{(n)}; \mathcal{U}_n(n))\} - \rho_K + \rho_K - \rho.$$

It therefore follows from (4.17), (4.19) and (4.13) that

$$\mu^{U_n(n+1)}\{M(X_1^{(n)}, \ldots, X_{n+1}^{(n)})\} - \mu^{U_n(n)}\{M(X_1^{(n)}, \ldots, X_n^{(n)})\} \to \rho \qquad (n \to \infty).$$

In the special case when $\psi(x) = x^\alpha$, rescaling to the unit cube shows that (with $X_1^{(1)}, X_2^{(1)}, \ldots$ i.i.d. and uniform on $[0,1]^d$)

$$(4.21) \quad n^{\alpha/d}[E\{M(X_1^{(1)}, \ldots, X_{n+1}^{(1)}; \alpha)\} - E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; \alpha)\}] \to \rho$$
$$(n \to \infty).$$

This implies the following asymptotic equivalences as $n \to \infty$:

$$(4.22) \quad E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; \alpha)\} \sim \sum_1^n \rho k^{-\alpha/d} \sim \frac{\rho d}{d - \alpha} n^{1-\alpha/d} \qquad \text{if } \alpha < d,$$

$$(4.23) \quad E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; d)\} \sim \sum_1^n \rho k^{-1} \sim \rho \log n,$$

$$(4.24) \quad E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; \alpha)\} \sim -\sum_n^\infty \rho k^{-\alpha/d} \sim \frac{\rho d}{d - \alpha} n^{1-\alpha/d} \quad \text{if } \alpha > d,$$

provided

$$(4.25) \qquad E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; \alpha)\} \to 0 \qquad (n \to \infty)$$

in case $\alpha > d$.

Now, first, if $\alpha < d$, (4.22) shows that

$$[(n+1)^{\alpha/d} - n^{\alpha/d}]E\{M(X_1^{(1)}, \ldots, X_{n+1}^{(1)}; \alpha)\} \to \frac{\rho \alpha}{d - \alpha} \qquad (n \to \infty),$$

which together with (4.21) proves (1.9) with $\overline{\rho} = \rho d/(d - \alpha)$ in this case. Equation (1.8) is just (4.22).

Next, if $\alpha = d$, then (4.23), together with (1.2) shows that $\rho = 0$. Then, by (4.21) and (1.2),

$$(n+1)E\{M(X_1^{(1)}, \ldots, X_{n+1}^{(1)}; d)\} - nE\{M(X_1^{(1)}, \ldots, X_n^{(1)}; d)\}$$

$$= o(1) + E\{M(X_1^{(1)}, \ldots, X_{n+1}^{(1)}; d)\} \to c(d, d).$$

This proves (1.9) when $\alpha = d$ while (1.8) is immediate from (1.2).

Finally, when $\alpha > d$, we must first prove (4.25). This follows from Lemmas 3 and 6. Indeed, if $\overline{\delta} = \overline{\delta}(x)$ is defined as in (3.17), then by Lemma 3, for any MST T on $\mathscr{U}_n(n) = \{X_1^{(n)}, \ldots, X_n^{(n)}\}$, it holds that

$$\sum_{\substack{e \in T, \text{ one endpoint} \\ \text{of } e \text{ in } C(x)}} |e|^\alpha \leq d^{\alpha/2}(\overline{\delta}(x) + 3)^\alpha |C(x) \cap \mathscr{U}_n(n)|.$$

Therefore, the expectation of the left-hand side with respect to μ^{U_n} is at most

$$d^{\alpha/2}[\mu^{U_n}\{(\overline{\delta}(x) + 3)^{2\alpha}\}\mu^{U_n}\{|C(x) \cap \mathscr{U}_n(n)|^2\}]^{1/2} \leq C_{15}$$

for some C_{15}, independent of x, n. Summing over $x = v(1), v(2), \ldots,$ $v((\lfloor n^{1/d} \rfloor + 1)^d)$ proves that for $\psi(x) = x^\alpha$,

$$\mu^{U_n}\{M(X_1^{(n)}, \ldots, X_n^{(n)}; \psi)\} \leq C_{15}(\lfloor n^{1/d} \rfloor + 1)^d.$$

Rescaling again to the unit cube shows that

$$E\{M(X_1^{(1)}, \ldots, X_n^{(1)}; \alpha)\} \leq C_{16} n^{1-\alpha/d}.$$

Thus (4.25) and (4.24) follow. We leave it to the reader to derive (1.8) and (1.9) for $\alpha > d$.

The strict positivity of $\overline{\rho}$ follows in all cases by the argument of Steele [(1988), end of Section 4]. □

LEMMA 14. *If along some subsequence* $n^{-1/2}(\widetilde{M}_n - \mu^{P_n}(\widetilde{M}_n))$ *converges in distribution to F with characteristic function $\widehat{F}$, then along the same subsequence* $n^{-1/2}(M(X_1^{(n)}, \ldots, X_n^{(n)}) - \mu^P(\widetilde{M}_n))$ *converges in distribution to G with characteristic function*

$$\widehat{G}(\theta) = \int \exp(i\theta x)\,dG(x) = \widehat{F}(\theta) \cdot \exp(\tfrac{1}{2}\theta^2 \rho^2).$$

PROOF. Couple $\mathscr{P}_n$ and $\mathscr{U}_n(m)$ by first choosing an infinite sequence $X_1^{(n)}, X_2^{(n)}, \ldots$ of i.i.d. uniform random variables on Q_n and, independently of the $X_i^{(n)}$, a Poisson variable $N = N_n$ with mean n. Then $(X_1^{(n)}, \ldots, X_m^{(n)})$ has the distribution $\mu^{U_n(m)}$ and $(X_1^{(n)}, \ldots, X_N^{(n)})$ has the distribution μ^{P_n}. In particular, $\widetilde{M}_n$ has the same distribution as $M(X_1^{(n)}, \ldots, X_N^{(n)})$, $n^{-1/2}(N_n - n) \Rightarrow N(0, 1)$ and, by Lemma 13,

$$(4.26) \qquad n^{-1/2}[M(X_1^{(n)}, \ldots, X_N^{(n)}) - M(X_1^{(n)}, \ldots, X_n^{(n)}) - (N_n - n)\rho] \to 0$$

in probability. Thus, if $n^{-1/2}(\widetilde{M}_n - \mu^{P_n}\{\widetilde{M}_n\})$ is tight, then $n^{-1/2}(M(X_1^{(n)}, \ldots, X_n^{(n)}) - \mu^{P_n}\{\widetilde{M}_n\})$ is tight. Therefore, if along some subsequence $n^{-1/2}(\widetilde{M}_n - \mu^{P_n}\{\widetilde{M}_n\})$ converges in distribution to F with characteristic function $\widehat{F}$, then along the same subsequence, by (4.26), $n^{-1/2}(M(X_1^{(n)}, \ldots, X_n^{(n)}) - \mu^{P_n}\{\widetilde{M}_n\} + (N_n - n)\rho)$ converges in distribution to F. Since N_n is independent of the $X_i^{(n)}$, along the same subsequence, $n^{-1/2}(M(X_1^{(n)}, \ldots, X_n^{(n)}) - \mu^{P_n}\{\widetilde{M}_n\})$ converges in distribution to G with characteristic function $\widehat{G}$, where $\widehat{G}(\theta)\exp(-\tfrac{1}{2}\theta^2\rho^2) = \widehat{F}(\theta)$. □

PROOF OF THEOREM 2. At first we prove the theorem in the Poissonized version which says that $n^{-1/2}(\widetilde{M}_n - \mu^{P_n}\{\widetilde{M}_n\})$ converges in distribution to a normal distribution with mean 0 and variance τ^2 for some $\tau > 0$. By virtue of the representation (4.2) of $\widetilde{M}_n - \mu^{P_n}\{\widetilde{M}_n\}$ as a sum of martingale differences and Theorem 2.3 in McLeish (1974), it suffices to verify the following relations (4.27)–(4.30):

$$(4.27) \qquad\qquad \frac{1}{n}\sum_k \Delta_k^2 \to \tau^2 \quad \text{in } \mu^{P_n}\text{-measure},$$

where

$$(4.28) \quad \tau^2 = \lim_{L \to \infty} \tau_L^2,$$

$$\tau_L^2 = \mu^P\big[\mu^P\{D(C(\mathbf{0}) \cap \mathscr{P}, ([-L, L+1]^d \setminus C(\mathbf{0})) \cap \mathscr{P}) \mid \mathscr{F}\}$$
$$- \mu^P\{D(C(\mathbf{0}) \cap \mathscr{P}, ([-L, L+1]^d \setminus C(\mathbf{0})) \cap \mathscr{P}) \mid \mathscr{F}'\}\big]^2,$$

$$\mathscr{F} = \sigma\{C(x) \cap \mathscr{P}: \ x \in \mathbb{Z}^d, C(x) \cap [-L, L+1]^d \neq \varnothing,$$
$$x \text{ strictly precedes } \mathbf{0} \text{ in the lexicographical order or } x = \mathbf{0}\}$$

and

$$\mathscr{F}' = \sigma\{C(x) \cap \mathscr{P}: \ x \in \mathbb{Z}^d, \ C(x) \cap [-L, L+1]^d \neq \varnothing,$$
$$x \text{ strictly precedes } \mathbf{0} \text{ in the lexicographical order}\};$$

$$(4.29) \qquad n^{-1/2} \max_k |\Delta_k| \to 0 \quad \text{in } \mu^{P_n}\text{-measure}$$

and

$$(4.30) \qquad n^{-1}\mu^{P_n}\Big\{\max_k |\Delta_k|^2\Big\} \text{ is bounded in } n$$

[the existence of the limit in (4.28) is part of what needs to be proven].

First let us deal with the easy relations (4.29) and (4.30). Equation (4.29) holds, because

$$\mu^{P_n}\Big\{n^{-1/2} \max_k |\Delta_k| \geq \varepsilon\Big\} \leq \mu^{P_n}\Big\{\sum_k |\Delta_k|^3\Big\}/(\varepsilon n^{1/2})^3 \leq n C_{17} C(3)/(\varepsilon n^{1/2})^3$$

by (4.3) and (3.12). For (4.30) we observe that

$$\mu^{P_n}\Big\{\max_k |\Delta_k|^2\Big\} \leq \sum_k \mu^{P_n}\{\Delta_k^2\} \leq n C_{17} C(2),$$

again by (4.3) and (3.12).

Finally we prove that for fixed L,

$$(4.31) \qquad \frac{1}{n}\sum_k \Delta_{k, L}^2 \to \tau_L^2 \quad \text{in } \mu^{P_n}\text{-measure}.$$

We claim that (4.31) implies (4.27). To see this, note first that (4.28) holds, as in the argument for (4.13), by the eventual monotonicity in L of $\mu^P\{D(C(\mathbf{0}) \cap \mathscr{P}, ([-L, L+1]^d \setminus C(\mathbf{0})) \cap \mathscr{P})|\mathscr{F}\}$. We may therefore choose $\overline{L}$ is so large that for all $L \geq \overline{L}$,

$$(4.32) \qquad\qquad |\tau^2 - \tau_L^2| < \varepsilon/3.$$

Now, for given $\varepsilon > 0$, we can [by (4.5)] choose $L = L(\varepsilon^2) \geq \overline{L}$ so that in addition to (4.32), for $n \geq n_3(L)$,

$$(4.33) \qquad \frac{1}{n} \sum_k \mu^{P_n} \{ |\Delta_k^2 - \Delta_{k,L}^2| \} < \varepsilon^2.$$

Finally, by (4.31), we can obtain an $n_4(L, \varepsilon) \geq n_3$ so that for $n \geq n_4$,

$$\mu^{P_n} \left\{ \left| \frac{1}{n} \sum_k \Delta_{k,L}^2 - \tau_L^2 \right| > \frac{\varepsilon}{3} \right\} < \varepsilon.$$

Then, for $n \geq n_4$, by (4.33) and (4.32),

$$\mu^{P_n} \left\{ \left| \frac{1}{n} \sum_k \Delta_k^2 - \tau^2 \right| > \varepsilon \right\}$$

$$\leq \mu^{P_n} \left\{ \left| \frac{1}{n} \sum_k \Delta_k^2 - \frac{1}{n} \sum_k \Delta_{k,L}^2 \right| > \frac{\varepsilon}{3} \right\} + \mu^{P_n} \left\{ \left| \frac{1}{n} \sum_k \Delta_{k,L}^2 - \tau_L^2 \right| > \frac{\varepsilon}{3} \right\}$$

$$\leq \frac{\varepsilon^2}{\varepsilon/3} + \varepsilon.$$

Thus (4.31) indeed implies (4.27).

To see (4.31) we merely note that $\Delta_{k,L}$ and $\Delta_{j,L}$ are independent as soon as $\Pi[(v(k))_i - L, (v(k))_i + L + 1] \cap \Pi[(v(j))_i - L, (v(j))_i + L + 1] = \phi$ and that for all $v(k) \in [L, n^{1/d} - L - 1]^d$, $\mu^{P_n} \{ \Delta_{k,L}^2 \} = \tau_L^2$ as defined above, provided we order the $v(k)$ lexicographically (as we may). Since the contribution of the terms with $v(k) \notin [L, n^{1/d} - L - 1]^d$ to (4.31) is at most $C_{18} n^{-1/d}$, (4.31) follows from a simple variance estimate and Chebychev's inequality.

As stated above, the relations (4.27)–(4.30) prove that

$$\frac{\widetilde{M}_n - \mu^{P_n} \{ \widetilde{M}_n \}}{n^{1/2}} \Rightarrow N(0, \tau^2).$$

By Lemma 14 we then have

$$(4.34) \qquad \frac{M(X_1^{(n)}, \dots, X_n^{(n)}) - \mu^{P_n} \{ \widetilde{M}_n \}}{n^{1/2}} \Rightarrow N(0, \sigma^2),$$

where $\tau^2 = \sigma^2 + \rho^2$. Below we will show that $\sigma^2 > 0$. This then completes the proof of the Poissonized version of Theorem 2.

To prove the original version of Theorem 2, we observe first that

$$(4.35) \qquad n^{-1/2} (\mu^{P_n} \{ \widetilde{M}_n \} - \mu^{U_n} \{ M(\mathscr{U}_n) \}) \to 0,$$

because [see (4.26)]

$$(4.36) \qquad n^{-1/2} (\widetilde{M}_n - M(\mathscr{U}_n)) \Rightarrow N(0, \rho^2)$$

and (4.15) shows that (when we use the coupling of Lemma 14, that is, $N = N_n$, a Poisson variable with mean n)

$$\frac{1}{n}E\{|\widetilde{M}_n - M(\mathscr{U}_n)|^2\}$$

$$\leq \frac{1}{n}E\bigg\{ \sum_{j=n+1}^{\infty} I[N=j]\bigg| \sum_{p=n+1}^{j} D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\bigg|^2$$

$$+ \sum_{j=0}^{n-1} I[N=j]\bigg| \sum_{p=j+1}^{n} D(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\bigg|^2\bigg\}$$

$$\leq \frac{1}{n}E\bigg[\sum_{j=n+1}^{\infty} I[N=j](j-n) \sum_{p=n+1}^{j} E\{D^2(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\}$$

$$+ \sum_{j=0}^{n} I[N=j](n-j) \sum_{p=j+1}^{n} E\{D^2(\{X_p^{(n)}\}, \mathscr{U}_n(p-1))\}\bigg]$$

<div align="right">(by Schwarz's inequality)</div>

$$\leq \frac{1}{n}C_{19}E\{(N_n - n)^2\} \leq C_{20}.$$

Therefore, the left-hand side of (4.35) is uniformly integrable and hence, by (4.36), (4.35) holds. We may therefore replace $\mu^{P_n}\{\widetilde{M}_n\}$ by $\mu^{U_n}\{M(\mathscr{U}_n)\}$ in (4.34).

Finally we show $\sigma^2 > 0$, by proving that $M(\mathscr{U}_n)$ cannot be concentrated on an interval of length $o(n^{1/2})$. This is done by a block argument which is similar to that of Avram and Bertsimas [(1993), Proposition 5]. We just describe our block for $d = 2$ and leave the case $d \geq 3$ to the reader. For $d = 2$ we consider a 13×13 square $W \subset Q_n$. For each 1×1 square along the boundary of W (see Figure 4) we require that

(4.37)   there is at least one point of $\mathscr{U}_n$ in the 1×1 square.

We also require that

(4.38) $\mathscr{U}_n$ contains exactly one point v_i in each of B_i, $i = 1, 2$,

where B_i are 1×1 squares as in Figure 4. Finally we require that

(4.39)   apart from v_1, v_2 and the points in the 1×1 squares along
                 the boundary of W, there are no points of $\mathscr{U}_n$ in W.

We claim that, if $\mathscr{U}_n \cap W$ has these properties, then in any MST T on $\mathscr{U}_n$, there is only one edge incident to $v_2 \in B_2$, and this edge connects v_2 to v_1. Indeed, for any $v \in \mathscr{U}_n \setminus \{v_1, v_2\}$, the line segment $\langle v, v_2 \rangle$ must intersect a 1×1 square on the boundary of W. Choose a $\mathscr{U}_n$-point w_1 in this square. Next choose a $\mathscr{U}_n$-point w_2 in the adjacent boundary square. Keep choosing $\mathscr{U}_n$-points w_j in successive adjacent boundary squares until we choose a $\mathscr{U}_n$-point w_m in B_3 (see Figure 4). Then $\pi = (v, w_1, \ldots, w_m, v_1, v_2)$ is a path from
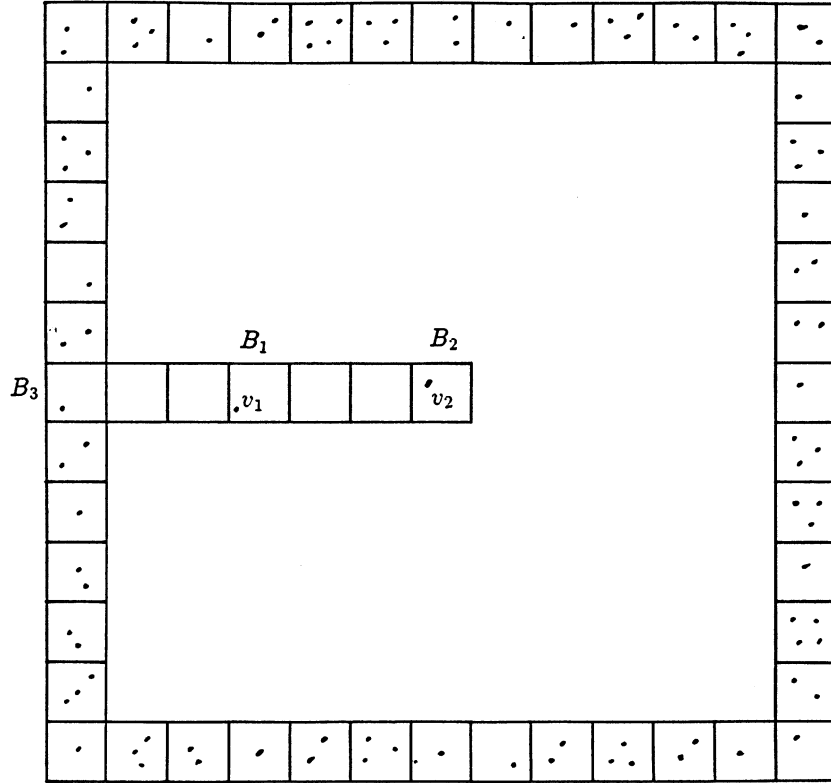
FIG. 4. *If $\mathscr{U}_n$ contains this 13×13 square configuration, then in any MST on $\mathscr{U}_n$ the only edge incident to v_2 is the edge (v_1, v_2).*

v to v_2 in $\mathscr{U}_n$ such that $|(v, w_1)| < |(v, v_2)|$, $|(w_{i-1}, w_i)| < |(v, v_2)|$, $2 \le i \le m$, $|(w_m, v_1)| < |(v, v_2)|$ and $|(v_1, v_2)| < |(v, v_2)|$. Therefore, by Lemma 2, (v, v_2) cannot be an edge in any MST T on $\mathscr{U}_n$. This proves our claim that the only edge in T incident to v_2 is the edge (v_1, v_2).

It follows from this claim that, if $\mathscr{U}_n \cap W$ satisfies (4.37)–(4.39), then

$$M(\mathscr{U}_n) = M(\mathscr{U}_n \setminus \{v_2\}) + \psi(|(v_1, v_2)|).$$

As in Avram and Bertsimas (1993), this proves that $\mathrm{Var}(M(\mathscr{U}_n)) \ge C_{21}n$ for some constant $C_{21} > 0$. However, without a proof of uniform integrability of $n^{-1}[M(\mathscr{U}_n) - \mu^{U_n}\{M(\mathscr{U}_n)\}]^2$, this does not guarantee $\sigma^2 > 0$. For $\psi(x) = |x|$, this uniform integrability is in fact contained in Talagrand's much more detailed and deeper tail estimates for $M(\mathscr{U}_n)$ [see Theorem 11.3.2 in Talagrand (1995)]. Alternatively, we can use the following well known argument. Choose $M = \lfloor C_{22}n \rfloor$ disjoint 13×13 squares $W_1, \ldots, W_M$ in Q_n. Let $W_{i_1}, \ldots, W_{i_s}$ be the random collection of those squares for which $\mathscr{U}_n \cap W_{ij}$ has the properties (4.37)–(4.39) (after a suitable translation). Let $v_{1,j}$ and $v_{2,j}$ be the analogues in W_{ij} of v_1 and v_2, respectively. Then conditioned on the index set $\{i_1, \ldots, i_s\}$

and $\mathcal{U}_n \setminus \left[\bigcup_{j=1}^s \{v_{2,j}\} \right]$, the $\psi(|(v_{1,j}, v_{2,j})|)$ are independent and

$$\frac{1}{\sqrt{s}} \sum_{j=1}^s \psi(|(v_{1,j}, v_{2,j})|)$$

is asymptotically normal with some mean, but with a variance $\nu^2 \geq C_{23}$ for some constant $C_{23} > 0$. In addition, it is easy to see that there exists some $C_{24} > 0$ such that

$$\mu^{U_n}\{s \geq C_{24} n\} \to 1$$

as $n \to \infty$. Standard arguments now show that

$$\sup_a \mu^{U_n}\{a \leq n^{-1/2} M(\mathcal{U}_n) \leq a + \varepsilon\} \to 0$$

as $\varepsilon \downarrow 0$, uniformly in n. Together with (4.34) this implies $\sigma^2 > 0$. This completes the proof of Theorem 2. $\square$

REFERENCES

ALDOUS, D. and STEELE, J. M. (1992). Asymptotics for Euclidean minimal spanning trees on random points. *Probab. Theory Related Fields* **92** 247–258.

ALEXANDER, K. S. (1994). Rates of convergence of means for distance-minimizing subadditive Euclidean functionals. *Ann. Appl. Probab.* **4** 902–922.

ALEXANDER, K. S. (1996). The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees. *Ann. Appl. Probab.* **6** 466–494.

AVRAM, F. and BERTSIMAS, D. (1993). On central limit theorems in geometric probability. *Ann. Appl. Probab.* **3** 1033–1046.

CHARTRAND, G. and LESNIAK, L. (1986). *Graphs and Digraphs*. Wadsworth, Belmont, CA.

JANSON, S. (1995). The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. Preprint.

KRUSKAL, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Amer. Math. Soc.* **7** 48–50.

LÉVY, P. (1937). *Théorie de l'Addition des Variables Aléatoires*. Gauthier–Villars, Paris.

MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628.

RAMEY, D. B. (1982). A non-parametric test of bimodality with applications to cluster analysis. Ph.D. dissertation, Yale Univ.

REDMOND, C. and YUKICH, J. E. (1994). Limit theorems and rates of convergence for Euclidean functionals. *Ann. Appl. Probab.* **4** 1057–1073.

STEELE, J. M. (1988). Growth rates of Euclidean minimal spanning trees with power weighted edges. *Ann. Probab.* **16** 1767–1787.

TALAGRAND, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. Publ. Math. vol. 81, Inst. des Mautes Etudes Scientifiques.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
E-MAIL: kesten@math.cornell.edu

DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE 0511
E-MAIL: matleesc@leonis.nus.sg