The Annals of Probability
1996, Vol. 24, No. 4, 1903–1925

THE RANDOM MINIMAL SPANNING TREE IN HIGH DIMENSIONS

BY MATHEW D. PENROSE

University of Durham

For the minimal spanning tree on n independent uniform points in the d-dimensional unit cube, the proportionate number of points of degree k is known to converge to a limit $\alpha_{k,d}$ as $n \to \infty$. We show that $\alpha_{k,d}$ converges to a limit α_k as $d \to \infty$ for each k. The limit α_k arose in earlier work by Aldous, as the asymptotic proportionate number of vertices of degree k in the minimum-weight spanning tree on k vertices, when the edge weights are taken to be independent, identically distributed random variables. We give a graphical alternative to Aldous's characterization of the α_k.

1. Introduction. The minimal spanning tree (MST) on a set of n random points in $\mathbf{R}^d$ (say, the measurements of d variables on each of n objects) has been used by statisticians as a means of imposing a structure on the observations. The structure is summarized by the lengths of the edges of the MST and the degrees of the vertices. A probabilistic literature on this structure has been growing, especially for the case where the n points are independent and identically distributed, and this is the setting of the present paper.

Consider the Euclidean MST on a set of points $\eta_1, \dots, \eta_n$ in $\mathbf{R}^d$, i.i.d. with common density f. Let $V_k(n)$ be the number of vertices in the tree with degree k. Steele, Shepp and Eddy [23] showed that $V_k(n)/n$ converges almost surely to a number $\alpha_{k,d}$ that depends on d, but not otherwise on f. Aldous and Steele [2] showed that $(\alpha_{k,d},\, k \geq 1)$ can be interpreted as the probability mass function of the degree of 0 in a sort of "MST" on a homogeneous Poisson process in $\mathbf{R}^d$. This characterisation is elegant, but does not lead immediately to any simple formula for $\alpha_{k,d}$. According to [23], "it is unreasonable to expect any determination of $\alpha_{k,d}$ for large k and d." One aim of the present paper is to make some progress in this direction by describing the behaviour of $\alpha_{k,d}$ in the limit $d \to \infty$. In physics terminology, this is a "mean-field limit."

One may also be interested in the lengths of the edges of the MST. For example, much work has been done on the sums of various powers of these lengths. For a discussion and references, see [22] and [13]. Our second main result is concerned with the asymptotic empirical distribution of these lengths in high dimensions. It is consistent with the result of [6] on the high-dimensional behavior for the growth rate of the total length of the MST.

Received July 1995; revised December 1995.

AMS 1991 *subject classifications*. Primary 60D05; secondary 05C05, 90C27.

Key words and phrases. Geometric probability, minimal spanning tree, vertex degrees, continuum percolation, invasion percolation.

Our limits for high dimension are described below in terms of a mean-field weighted tree. This limiting system (described somewhat differently) has arisen in [1], which is concerned with large-n asymptotics for the MST on n points when the $\binom{n}{2}$ "interpoint distances" are taken to be *independent*, identically distributed random variables. A further reason for this study is to demonstrate an explicit link between the Euclidean MST, which generally tends to be less tractable, and the MST for independent edge weights, in the spirit of [4].

One concrete application of our results is in the study of the multivariate nonparametric two-sample problem: given samples of size n_1 and n_2 from d-dimensional distributions with densities f_1 and f_2, respectively, suppose one wishes to know if it is reasonable to assume the distributions are the same. The *multivariate runs test* of Friedman and Rafsky [9] goes as follows: construct the MST on the pooled sample, and let R denote the number of edges in this MST which have one endpoint from sample 1 and one endpoint from sample 2. If R is small, reject the null hypothesis H_0 that $f_1 \equiv f_2$. Using (14) of [9] or otherwise, one can obtain a distribution-free large-sample asymptotic formula for the variance under H_0 of R, in terms of the numbers $\alpha_{k,d}$ (we omit the details). Other nonparametric test statistics based on the MST were proposed by Friedman and Rafsky [10], and similar formulae for their asymptotic variance can be derived.

Another statistical application of the MST is given in [19]. The minimal spanning tree also has applications in computer science, the physical sciences and in biology. See, for example, [5], [12], [8] and other papers referred to in [23].

In the next section we describe our main results in detail. In Section 3 we describe detailed properties of the limiting mean-field tree. Sections 4 and 5 contain the proofs of the main results.

There are connections between the MST on random points and continuum percolation. More details of this relationship are given in [17], which contains complementary (but distinct) results on the high-dimensional behavior of continuum percolation. Section 6 describes some analogues to the results of the present paper, which are likely to be true for invasion percolation, a lattice growth model which has some similarities to the construction of the "MST" on the points of the Poisson process.

2. The main results.

The MST on the random points $\eta_1, \dots, \eta_n$ is the tree with vertex set $\{\eta_1, \dots, \eta_n\}$ and edges $e_1, \dots, e_{n-1}$ chosen to minimize $\sum_i |e_i|$, where $|\cdot|$ denotes Euclidean length. Recall that $V_k(n)$ denotes the number of vertices of degree k in this tree.

PROPOSITION 1 ([23]). *There exist numbers $\alpha_{k,d}$ depending on d but not otherwise on f, such that*

$$\lim_{n \to \infty} V_k(n)/n = \alpha_{k,d} \quad a.s.$$

For a countably infinite connected graph $\mathscr{G}$ with distinct finite weights defined on its edges, the following analogue of the MST has been proposed by Aldous and Steele [2]. An initial vertex x of $\mathscr{G}$ is specified, and an increasing sequence (t_n) of trees (in $\mathscr{G}$) is then generated as follows (we identify each tree with its set of edges). Let e_1 be the edge with minimal weight out of those edges of $\mathscr{G}$ with one end at x, and let t_1 be the single-edge tree $\{e_1\}$. Then recursively, given t_n, let e_{n+1} be the edge of $\mathscr{G}$ of minimal weight with just one endpoint in t_n, and set $t_{n+1} = t_n \cup \{e_{n+1}\}$. Set $t_\infty(x, \mathscr{G}) = \bigcup_{n=1}^{\infty} t_n$, the union of all the trees t_n generated in this way starting from x.

Consider now the subgraph of $\mathscr{G}$ obtained by including every edge (x, y) that is either in $t_\infty(x, \mathscr{G})$ or in $t_\infty(y, \mathscr{G})$ (or both). The component of this subgraph containing x is a tree, which we denote $g(\mathscr{G}, x)$. See Lemma 1 of [2].

Let $\mathscr{P}_d$ denote a homogeneous Poisson process in $\mathbf{R}^d$ of density 1, with a point added at 0. The construction above was applied in [2] to the complete graph with the points of $\mathscr{P}_d$ as its vertex set and edge weights given by the Euclidean interpoint distances; we shall denote this graph $\mathscr{P}_d$ as well. Define

$$D_t(d) = \operatorname{card}\{X \in \mathscr{P}_d : |X| \le t, (0, X) \in g(\mathscr{P}_d, 0)\}, \tag{1}$$

where card denotes cardinality. Set $D(d) = \lim_{t\to\infty} D_t(d)$, the degree of 0 in $g(\mathscr{P}_d, 0)$.

Proposition 2 ([2]). *When f is the uniform density on the unit cube in $\mathbf{R}^d$, $E[V_k(n)/n] \to P[D(d) = k]$ as $n \to \infty$.*

It is immediate from Propositions 1 and 2 that $\alpha_{k,d} = P[D(d) = k]$.

We turn now to the empirical distribution of the lengths of the edges of the MST on n points. For $t \ge 0$ let the random variable $F_n(t; f)$ denote the proportion of edges e of the MST on $\{\eta_1, \dots, \eta_n\}$ with $|e| \le t$.

Proposition 3. *For $t \ge 0$,*

$$F_n(n^{-1/d}; f) \to_{L^2} \tfrac{1}{2} \int_{R^d} E\big[D_{tf(x)^{1/d}}(d)\big] f(x)\, dx \quad \text{as } n \to \infty, \tag{2}$$

where $\to_{L^2}$ denotes convergence in mean square. In particular, in the special case where f is the uniform density on the unit cube, $F_n(n^{-1/d}t; f) \to_{L^2} E[D_t(d)]/2$ as $n \to \infty$.

The proof of this result is an extension of the methods of [2], and is given in the Appendix.

The above results indicate why $\alpha_{k,d}$ and $ED_t(d)$ might be of interest. However, these quantities seem to be hard to evaluate; thus we consider their limiting behavior for large d. To describe the limits, consider the complete rooted $\mathbf{N}$-ary tree $\mathscr{T}$, where $\mathbf{N} = \{1, 2, 3, \dots\}$. We identify the vertex set of $\mathscr{T}$ with the set $\bigcup_{n=0}^{\infty} \mathbf{N}^n$ of finite sequences (words) of natural numbers (including the empty word). If $w \in \mathbf{N}^n$ and $v \in \mathbf{N}^m$ are words of length n and m, respectively, let $w * v$ denote the concatenation of w and v, a word of length $n + m$ (e.g., $244 * 43 = 24443$). We make $\mathscr{T}$ into a tree by including an edge between w and $w * i$, for each word $w \in \bigcup_{n=0}^{\infty} \mathbf{N}^n$ and $i \in \mathbf{N}$ (viewed as a one-letter word). The root of $\mathscr{T}$ is the empty word, denoted $\emptyset$. For each word w other than $\emptyset$, let $e(w)$ denote the (unique) edge of $\mathscr{T}$ which has w as its endpoint further from the root.

We give random weights to the edges of $\mathscr{T}$ as follows. On a suitable probability space let $\mathscr{P}(w)$, $w \in \bigcup_{n=0}^{\infty} \mathbf{N}^n$, be independent homogeneous Poisson processes of rate 1 on $(0, \infty)$. For each $w \in \bigcup_{n=0}^{\infty} \mathbf{N}^n$, label the arrival times of $\mathscr{P}(w)$ in increasing order as $\Gamma_1(w), \Gamma_2(w), \Gamma_3(w), \dots$. For each $i \in \mathbf{N}$, give the edge $e(w * i)$ from w to $w * i$ the weight $W(w * i) := \Gamma_i(w)$. With probability 1, the weights $W(w)$, $w \in \bigcup_{n=1}^{\infty} \mathbf{N}^n$, are all distinct.

Our limits are described in terms of the weighted tree $g(\mathscr{T}, \emptyset)$, defined by applying the construction above to the randomly weighted graph $\mathscr{T}$. For $t \geq 0$, let $D_t(\infty)$ denote the total number of edges from $\emptyset$ in $g(\mathscr{T}, \emptyset)$ whose weights are at most t. That is,

$$D_t(\infty) = \operatorname{card}\{i \in \mathbf{N}: e(i) \in g(\mathscr{T}, \emptyset), W(i) \leq t\}. \tag{3}$$

Set $D(\infty) + 1 := \lim_{t \to \infty} D_t(\infty)$, the degree of the root $\emptyset$ in $g(\mathscr{T}, \emptyset)$. So $D(\infty) + 1$ is the number of one-letter words i such that $i \in t_\infty(\emptyset, \mathscr{T})$ or $\emptyset \in t_\infty(i, \mathscr{T})$ (or both).

We now state our two main results.

THEOREM 1. *For each $k \in \mathbf{N}$, setting $\alpha_k = P[D(\infty) + 1 = k]$, we have*

$$\lim_{d \to \infty} \alpha_{k,d} = \alpha_k. \tag{4}$$

As we shall see in the next section, the structure of $g(\mathscr{T}, \emptyset)$ is equivalent to that described in different terms by Aldous [1], and $D(\infty)$ is the same as in [1] (this is why 1 was added in the definition). Aldous derives a formula for the α_k and evaluates it numerically; we have $\alpha_1 = 0.408$, $\alpha_2 = 0.324$, $\alpha_3 = 0.171$, $\alpha_4 = 0.022$, $\alpha_5 = 0.006$ and $\alpha_6 = 0.001$.

Let $v_d(s)$ denote the volume of a Euclidean ball of radius s in $\mathbf{R}^d$, and let $r_d(t)$ denote the radius of a Euclidean ball of volume t in $\mathbf{R}^d$. That is,

$$v_d(s) = \pi_d s^d, \qquad v_d(r_d(t)) = t, \tag{5}$$

where $\pi_d := \pi^{d/2}/\Gamma((d/2) + 1)$, the volume of the unit ball in $\mathbf{R}^d$. It turns out to be natural to measure lengths on the "volume scale," that is, to transform edge lengths by the monotone function v_d.

Let $\psi(t)$ denote the extinction probability for a Galton–Watson branching process [denoted $(G_n, n \geq 0)$] with a Poisson(t) offspring distribution and with $G_0 = 1$, that is, the smallest solution to

$$\psi(t) = \exp(t(\psi(t) - 1)), \tag{6}$$

and let $\bar{\psi}(t) = 1 - \psi(t)$.

THEOREM 2. *For each* $t \geq 0$,

$$\lim_{d \to \infty} E\left[D_{r_d(t)}(d)\right] = E\left[D_t(\infty)\right] \tag{7}$$

$$= \int_0^t \left(1 - \left(\bar{\psi}(s)\right)^2\right) ds. \tag{8}$$

Here is a brief explanation of why the graph $\mathcal{T}$ might be expected to arise in the large-d limit. Define points Y_w of $\mathcal{P}_d$ for each $w \in \bigcup_{n=0}^{\infty} \mathbf{N}^n$ as follows. Set $Y_\varnothing = 0$, and let Y_i be the ith nearest neighbor of 0 in $\mathcal{P}_d$. Let Y_{ij} be the jth nearest neighbor of Y_i, not counting 0, and let Y_{ijk} be the kth nearest neighbor of Y_{ij}, not counting its "parent" Y_i, and so on. For each word w and $i \in \mathbf{N}$, set $X_{w*i} = Y_{w*i} - Y_w$. Then, for any finite collection $\mathcal{W}$ of words w, for large d the vectors $(X_w,\ w \in \mathcal{W})$ are very likely to be approximately (i) mutually orthogonal and (ii) all the same length. This means there is unlikely to be much "interference" between different vectors of the form X_w.

Also, $\{v_d(|Y_i|),\ i \geq 1\}$ are the arrival times of a Poisson process of rate 1 on $(0, \infty)$. If we identify each Y_w with the vertex w of $\mathcal{T}$ and give a weight $W^d(w) := v_d(|X_w|)$ to the edge $e(w)$ of $\mathcal{T}$, it can be shown that the finite-dimensional distributions of $\{W^d(w)\}$ converge to those of $\{W(w)\}$ given by independent Poisson processes as above. This suggests that the weighted graph $g(\mathcal{P}_d, 0)$ [with each edge e given weight $v_d(|e|)$] might converge weakly (in an appropriate topology on weighted rooted trees) to $g(\mathcal{T}, \varnothing)$. This conjectured convergence would immediately give us (4) and (7), which are concerned with particular aspects of these trees. The main source of difficulty in proving this sort of result is the long-range dependence in the MST.

This interpretation can be used to understand (and presumably to rederive) other results on large-d limits, found in [21] and [15]. For example, Theorem 4.1 of [21] is concerned with the large-d limit of the probability that 0 is the rth nearest neighbor of its sth nearest neighbor in $\mathcal{P}_d$, and the above discussion suggests that this limit ought to be the probability that there are $s - 1$ points of $\mathcal{P}(r)$ in the interval $(0, \Gamma_r(\varnothing))$. By considering the pooled Poisson process $\mathcal{P}(\varnothing) \cup \mathcal{P}(r)$, it is easily verified that this probability is indeed the limit in Theorem 4.1 of [21].

3. The limiting tree. The graph $g(\mathcal{T}, \varnothing)$ is closely related to a tree described via a tree-valued Markov process in [1], as we shall now show. For any vertex w of $\mathcal{T}$, let $\mathcal{T}_w$ be the subtree of $\mathcal{T}$ rooted at w (i.e., the full subgraph on $\{w * v\colon v \in \bigcup_{n=0}^{\infty} \mathbf{N}^n\}$). Clearly, $\mathcal{T}_w$ is isomorphic to $\mathcal{T}$.

For $s > 0$, let $\mathcal{G}_s$ denote the tree obtained from $\mathcal{T}$ by deleting all edges of weight strictly greater than s, then taking the component of the resulting graph that includes the root. Let $\mathcal{G}_s^w$ be the tree obtained in the same way from $\mathcal{T}_w$. Since the number of edges $(w, w * i)$ of weight at most s from a given vertex w is Poisson with mean s, the graph $\mathcal{G}_s$ is the family tree of a Galton–Watson branching process with Poisson(s) offspring distribution, with one progenitor. The graphs ($\mathcal{G}_s$, $s \geq 0$) (with labels and weights ignored) form a realization of the tree-valued Markov process described in [1], page 388, and also denoted ($\mathcal{G}_s$) there. This is because, given $\mathcal{G}_s$, each vertex w of $\mathcal{G}_s$ has probability ds of acquiring a new arrival [say $\Gamma_i(w)$, where $i = i(w, s)$ here] before time $s + ds$. Such an arrival will add an extra edge $e(w * i)$ to $\mathcal{G}_{s+ds}$, together with the whole of $\mathcal{G}_s^{w*i}$, which is an independent copy of $\mathcal{G}_s$.

As in [1], define the random variable L by

$$L = \inf\{s: |\mathcal{G}_s| = \infty\} = \sup\{W(w): e(w) \in t_\infty(\emptyset, \mathcal{T})\}, \tag{9}$$

where the second equality follows from the definition of $t_\infty(\emptyset, \mathcal{T})$. Then $\mathcal{G}_L$ is infinite by the second definition of L, and thus

$$P[L \leq s] = P[|\mathcal{G}_s| = \infty] = 1 - \psi(s) = \bar{\psi}(s). \tag{10}$$

Thus $P[1 < L < \infty] = 1$.

As argued in [1], page 389, with probability 1, $\mathcal{G}_L$ is infinite but $\mathcal{G}_{L-}$ (the union of all $\mathcal{G}_s$, $s < L$) is finite. Therefore, since the edge weights $W(w)$ are a.s. distinct, there is a single (random) edge, $e(w_0)$ say, of $\mathcal{G}_L$, with weight $W(w_0) = L$. Moreover, $\mathcal{G}_{L-} \subset t_\infty(\emptyset, \mathcal{T}) \subset \mathcal{G}_L$ and $e(w_0) \in t_\infty(\emptyset, \mathcal{T})$.

For each vertex w of $\mathcal{T}$, define L_w analogously to L on $\mathcal{T}_w$ by

$$L_w = \sup\{W(w * v): e(w * v) \in t_\infty(w, \mathcal{T}_w)\}. \tag{11}$$

Clearly, L_w has the same distribution as L, and the supremum is achieved at a unique edge of $\mathcal{T}_w$.

Since $\mathcal{G}_{L-} \subset t_\infty(\emptyset, \mathcal{T}) \subset \mathcal{G}_L$, the rooted trees $\mathcal{G}_L$ and $t_\infty(\emptyset, \mathcal{T})$, with the branch containing w_0 deleted (i.e., with the first edge of the path from $\emptyset$ to w_0 deleted, along with the component of the resulting graph that does not contain $\emptyset$), are the same tree, denoted $\mathcal{H}_L^0$ in [1] (and here).

To see which edges of weight greater than L are included in $g(\mathcal{T}, \emptyset)$, let $\mathcal{H}_s^0$ denote the component containing $\emptyset$ of the graph $g(\mathcal{T}, \emptyset)$ with all edges of weight greater than s removed and with the branch containing w_0 removed. Then $\mathcal{H}_L^0$ is as defined above. For $s > L$, if $w \in \mathcal{H}_{s-}^0$ and $W(w * i) = s$, then $e(w * i) \notin t_\infty(w, \mathcal{T})$, so $e(w * i)$ will be included in $g(\mathcal{T}, \emptyset)$ if and only if $w \in t_\infty(w * i, \mathcal{T})$, which happens if and only if $L_{w*i} > s$. Thus, if $w \in \mathcal{H}_{s-}^0$ and $W(w * i) = s$,

$$\mathcal{H}_s^0 = \begin{cases} \mathcal{H}_{s-}^0, & \text{if } |\mathcal{G}_s^{w*i}| = \infty, \\ \mathcal{H}_{s-}^0 \cup \{e(w * i)\} \cup \mathcal{G}_s^{w*i}, & \text{if } |\mathcal{G}_s^{w*i}| < \infty. \end{cases}$$

But if labels (except at the root) are ignored, this is equivalent to the transition mechanism after time L of the tree-valued process denoted ($\mathcal{H}_s^0$) in [1]; thus the processes denoted $\mathcal{H}_s^0$ here and in [1] are equivalent. Therefore, the degree of $\emptyset$ in $g(\mathcal{T}, \emptyset)$ is the same as $D(\infty) + 1$ in [1], and the calculation in [1], page 395, of its distribution is also valid for the present interpretation. Also, the limit α_k of Theorem 1 is the same as the large-n limit established in [1] for the proportion of vertices of degree k in the minimal-weight spanning tree on the complete graph on n points when all $\binom{n}{2}$ edge weights are taken to be independent and uniform on $[0, 1]$.

We now prove (8) of Theorem 2.

PROPOSITION 4.

$$E[D_t(\infty)] = \int_0^t \Big(1 - \big(\bar{\psi}(s)\big)^2\Big)\, ds. \tag{12}$$

PROOF. We view $\mathscr{P}(\varnothing)$ as a *marked* Poisson process in which each arrival $\Gamma_i(\varnothing)$ is marked by the value of L_i given by (11) (so the marks are independent copies of L). For each $i \in \mathbf{N}$, define

$$L_{-i} = \inf_{j \neq i}\Big(\max\big(\Gamma_j(\varnothing), L_j\big)\Big) \tag{13}$$

(note that L is given by the same formula with the infimum taken over *all* j). The statement "$e(i) \in t_\infty(0, \mathscr{T})$" means that, starting from 0, one cannot continue indefinitely adding edges of length less than $W(i)$, that is, $L_{-i} > W_i$. Similarly, $e(i) \in t_\infty(i, \mathscr{T})$ if and only if $L_i > W(i)$. Thus $e(i) \in g(\mathscr{T}, \varnothing)$ if and only if $\max(L_{-i}, L_i) > W(i)$, and so

$$D_t(\infty) = \sum_{i:\, \Gamma_i(\varnothing) \le t} \mathbf{1}_{\{\max(L_i, L_{-i}) > W(i)\}}. \tag{14}$$

If L' is an independent copy of L, then, by (10), for $s > 0$,

$$P[\max(L, L') > s] = 1 - \bar{\psi}(s)^2. \tag{15}$$

The formula (12) for $E[D_t(\infty)]$ is now a routine application of Palm theory for the marked Poisson process. A proof from first principles goes as follows. Condition on the number N_t of points of $\mathscr{P}(\varnothing)$ in the interval $(0, t)$. Given that $N_t = k$, this point process consists of k unordered points, denoted $U_1, \ldots, U_k$ say, independent and uniform on $(0, t]$. Let L'_i denote the mark at U_i, and let $L'_{-i} = \inf_{j:\, i \neq j \le k} \max(U_j, L'_j)$ (when $k = 1$ take $L'_{-i} = +\infty$). Then, by (14) and exchangeability, for $k \ge 1$,

$$\begin{aligned} E[D_t(\infty) | N_t = k] &= E \sum_{i=1}^{k} \mathbf{1}_{\{\max(L'_i, L'_{-i}) > U_i\}} \\ &= k \int_0^t P[\max(L'_1, L'_{-1}) > s]\, t^{-1}\, ds. \end{aligned}$$

Since L'_{-1} and L'_1 are independent for $k > 1$, with L'_1 having the same distribution as L, and L'_{-1} having the conditional distribution of L given $N(t) = k - 1$, we have that

$$\begin{aligned} E[D_t(\infty)] &= \sum_{k=1}^{\infty} (e^{-t} t^k / k!) k \int_0^t P[\max(L', L) > s | N_t = k - 1]\, t^{-1}\, ds \\ &= \int_0^t P[\max(L', L) > s]\, ds, \end{aligned}$$

and (12) now follows from (15). □

In connection with this proof, note that, by (15) and symmetry,

$$(1/2)\left(1 - \overline{\psi}(s)^2\right) = P[L > \max(s, L')], \tag{16}$$

which is the (Palm) probability, given that there is an arrival of $\mathscr{P}(\varnothing)$ at time s, that the corresponding edge of $\mathscr{T}$ is part of the path from 0 to w_0 [defined after (10) above]. Setting J to be the weight of the first edge from $\varnothing$ in this path, we obtain $P[J < t] = \int_0^t (1 - \overline{\psi}(s)^2)\, ds/2$, which was obtained by other means in (33) of [1].

4. Preliminaries. Before proving Theorem 1 we introduce some notation. For $r > 0$ and $x \in \mathbf{R}^d$, define

$$B_r(x) = \{y \in \mathbf{R}^d : |y - x| \le r\}, \tag{17}$$

the closed ball of radius r centred at x, and set $B_r = B_r(0)$.

For any point process $\mathbf{X}$ in $\mathbf{R}^d$, define an r-path in $\mathbf{X}$ to be a (possibly infinite) sequence (X_n) of distinct points of $\mathbf{X}$ with $|X_n - X_{n+1}| < r$ for each n. For x and y in $\mathbf{R}^d$, let $\{x \leftrightarrow_r y \text{ in } \mathbf{X}\}$ denote the event that either there is a finite r-path $(X_1, \dots, X_m)$ in $\mathbf{X}$ with $|x - X_1| < r$ and $|y - X_m| < r$, or $|x - y| < r$. Let $\{x \leftrightarrow_r \infty \text{ in } \mathbf{X}\}$ denote the event that there is an infinite r-path (X_n) in $\mathbf{X}$ with $|x - X_1| < r$ and $|X_n| \to \infty$ as $n \to \infty$. For any sets $A \subset \mathbf{R}^d$, $B \subset \mathbf{R}^d$, define the event

$$\{A \leftrightarrow_r y \text{ in } \mathbf{X}\} := \bigcup_{x \in A} \{x \leftrightarrow_r y \text{ in } \mathbf{X}\},$$

and define $\{A \leftrightarrow_r \infty \text{ in } \mathbf{X}\}$ and $\{A \leftrightarrow_r B \text{ in } \mathbf{X}\}$ likewise.

The next two lemmas are slight modifications of Lemmas 3 and 4 of [17]. The proofs are virtually unchanged, and are omitted.

LEMMA 1. *Let $U(d)$ be a uniform random variable on the d-dimensional unit ball B_1. Then*

$$\lim_{d \to \infty} P[|U(d)| > 3/4] = 1 \tag{18}$$

and

$$\lim_{d \to \infty} \left(\sup\{P[|U(d) - x| \le 1.1] : x \in \mathbf{R}^d, |x| \ge 3/4\}\right) = 0. \tag{19}$$

LEMMA 2. *Define $L: \mathbf{R}^d \to \mathbf{R}^2$ by $L(x_1, x_2, \dots, x_d) = \sqrt{d}\,(x_1, x_2)$. The two-dimensional random vector $L(U(d))$ converges in distribution to the bivariate normal with mean 0 and the identity matrix I as covariance matrix.*

We shall use the technique of comparison with oriented site percolation on the lattice

$$\mathscr{L} := \{(i, j) \in \mathbf{Z}^2 : i \ge 0, |j| \le i, (i + j)/2 \in \mathbf{Z}\}, \tag{20}$$

with oriented edges from (i, j) to $(i + 1, j \pm 1)$. For $p \in (0, 1)$, let $\theta_{\mathscr{L}}(p)$ denote the probability that, for oriented site percolation on $\mathscr{L}$ with parameter p (the probability each site is occupied), there is an infinite path of occupied sites from $(0,0)$. By a contour argument (see [7], page 86), there exist constants $c > 0$ and $q_0 > 0$ such that

$$1 - \theta_{\mathscr{L}}(1-q) \le \sum_{n \ge 4} 3^{n-1} q^{n/4} \le cq \quad \text{for all } q < q_0. \tag{21}$$

We shall use procedures in which states of $\mathscr{L}$ are examined in turn in the order $(0,0), (1,-1), (1,1), (2,-2), (2,0), \dots$. Each site will be deemed "occupied" or "vacant" by a random mechanism that may depend on the earlier sites but gives conditional probability of "occupied," given the status of the earlier sites, uniformly exceeding some value p. The probability of an infinite path from $(0,0)$ of occupied sites will then exceed $\theta_{\mathscr{L}}(p)$. See, for example, [11], Lemma 1.

For $(i,j) \in \mathscr{L}$, let $B_{1/2}(i,j)$ be the closed disk of diameter 1 in $\mathbf{R}^2$ centered at (i,j), and set

$$A_{i,j} = L^{-1}\big(B_{1/2}(i,j)\big), \tag{22}$$

with L as in Lemma 2. The $A_{i,j}$ are disjoint regions of $\mathbf{R}^d$.

Let $\mathscr{P}_d^t$ denote a homogeneous Poisson process of rate t/π_d on $\mathbf{R}^d$. Note that $\mathscr{P}_d^1 \ne \mathscr{P}_d$. The point process $\mathscr{P}_d^t$ arises naturally as a result of scaling; see below. Our first use of oriented percolation is to obtain the following uniform bound.

LEMMA 3. *There exist constants $c_1 > 0$, $c_2 > 0$ such that, for some $t_0 > 0$ and $d_0 \in \mathbf{N}$,*

$$1 - P\big[0 \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^t\big] \le c_1 \exp(-c_2 t) \quad \text{for all } d \ge d_0,\ t \ge t_0.$$

PROOF. It follows from Lemma 2 that there is a constant $\eta > 0$ and some d_0 such that, for $d \ge d_0$,

$$\inf_{x \in A_{0,0}} \{P[x + U(d) \in A_{1,-1}]\} > \eta \quad \text{and} \quad P\big[U(d) \in A_{0,0}\big] > \eta.$$

Put another way, the first of these inequalities says that, for $d \ge d_0$ and $x \in A_{i,j}$, the proportion of the ball $B_1(x)$ lying in $A_{i+1,j\pm1}$ is at least η, and therefore, for $t > 0$ and $d \ge d_0$,

$$\int_{B_1(x) \cap A_{i+1,j\pm1}} (t/\pi_d)\, dy \ge \eta t \quad \text{and} \quad \int_{B_1 \cap A_{0,0}} (t/\pi_d)\, dy \ge \eta t. \tag{23}$$

Let $\bar{A}_{ij}$ be the union of those $A_{i'j'}$ with $(i',j') \in \mathscr{L}$ and (i',j') preceding (i,j) in the ordering described above. That is, either $i' < i$ or $i' = i$ and $j' < j$. Let $\mathscr{F}_{i,j}$ be the σ-field generated by the positions of those particles of $\mathscr{P}_d^t$ lying in $\bar{A}_{i,j}$, and define the event

$$E_{ij} = \Big\{\exists\, X \in \mathscr{P}_d^t \cap A_{ij} \text{ with } 0 \leftrightarrow_1 X \text{ in } \mathscr{P}_d^t \cap \bar{A}_{ij}\Big\}.$$

By (23), $P[E_{0,0}^c] \le e^{-\eta t}$. For $(i,j) \in \mathscr{L} \setminus \{(0,0)\}$, the events $E_{i-1,j\pm1}$ are in $\mathscr{F}_{i,j}$, and it follows from (23) that

$$P\big[E_{ij}^c | \mathscr{F}_{ij}\big] \le e^{-\eta t} \quad \text{on } E_{i-1,j-1} \cup E_{i-1,j+1}. \tag{24}$$

If $E_{i,j}$ occurs for infinitely many $(i,j) \in \mathscr{L}$, then $0 \leftrightarrow_1 \infty$ in $\mathscr{P}_d^t$. The result now follows from (24) by comparison with oriented percolation and (21). □

The following application of Lemma 3 says that edges of large length (measured on the volume scale) are unlikely to contribute to the degree $D(d)$ of 0 in $g(\mathscr{P}_d, 0)$, regardless of d.

LEMMA 4. *There exists $d_0 > 0$ such that*

$$\lim_{t\to\infty} \sup_{d \ge d_0} P\left[D_{r_d(t)}(d) < D(d)\right] = 0. \tag{25}$$

PROOF. The probability on the left-hand side of (25) is bounded above by the expression

$$\begin{aligned} &E\left[\operatorname{card}\{X \in \mathscr{P}_d : |X| > r_d(t), (0, X) \in g(\mathscr{P}_d, 0)\}\right] \\ &\quad = \int_{y \in R^d : |y| > r_d(t)} P\left[(0, y) \in g(\mathscr{P}_d \cup \{y\}, 0)\right] dy \\ &\quad \le \int_{y : |y| > r_d(t)} P\left[\{0 \leftrightarrow_{|y|} \infty \text{ in } \mathscr{P}_d, y \leftrightarrow_{|y|} \infty \text{ in } \mathscr{P}_d\}^c\right] dy \\ &\quad \le \int_{y : |y| > r_d(t)} 2P\left[\{0 \leftrightarrow_{|y|} \infty \text{ in } \mathscr{P}_d\}^c\right] dy \\ &\quad = 2\int_t^\infty P\left[\{0 \leftrightarrow_{r_d(s)} \infty \text{ in } \mathscr{P}_d\}^c\right] ds. \end{aligned} \tag{26}$$

The image of $\mathscr{P}_d$, under the scaling transformation $x \to x/r_d(t)$, is a Poisson process on $\mathbf{R}^d$ of rate $(r_d(t))^d$, that is, of rate t/π_d by (5). With this scaled Poisson process denoted $\mathscr{P}_d^t$, the upper bound in (26) is equal to

$$2\int_t^\infty P\left[\{0 \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^s\}^c\right] ds,$$

which converges to 0 uniformly in d by Lemma 3. □

The next lemma is based on the following idea. Given a point $x \in \mathbf{R}$, viewed as "generation 0," let the points of $\mathscr{P}_d^t$ in $B_1(x)$ be thought of as its "offspring," a Poisson number of points with mean t. Each offspring would itself have a Poisson(t) number of offspring and so on, were it not for the overlap of the balls centered at x and at its offspring. However, this effect becomes negligible for large d, and the sequence of offspring, grandchildren and so on of x resembles a branching process with a single progenitor at x. The result here says the same is true if one considers an initial set of k progenitors, provided they are not too close together.

For $K > 0$ and $d, \nu \in \mathbf{N}$, let $\mathscr{U}(d, \nu, K)$ denote the set of subsets $\mathbf{u} \subset \mathbf{R}^d$, of cardinality ν, such that $|u| > 3/4$ and $|L(u)| \le K$ for each $u \in \mathbf{u}$ and such that $|u - u'| > 1$ for all distinct u, u' in $\mathbf{u}$.

LEMMA 5. *Let* $t > 0$, $\nu \in \mathbf{N}$ *and* $K > 0$. *Let* $(\rho(d), d \geq 1)$ *be a sequence of nonnegative numbers with* $\limsup_{d \to \infty} \rho(d) \leq 1$. *Then, with* $\mathscr{P}_d^t$ *as in Lemma* 3 *and* $\psi(t)$ *as in* (6),

$$\sup_{\mathbf{u} \in \mathscr{U}(d,\nu,K)} P\big[\mathbf{u} \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^t \setminus B_{\rho(d)}\big] \leq 1 - \psi(t)^\nu \quad \text{for all } d, \tag{27}$$

while, conversely,

$$\lim_{d \to \infty} \inf_{\mathbf{u} \in \mathscr{U}(d,\nu,K)} P\big[\mathbf{u} \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^t \setminus B_{\rho(d)}\big] = 1 - \psi(t)^\nu. \tag{28}$$

REMARKS. In [17] we proved the special case of this result with $\nu = 1$, $\rho(d) = 0$. The proof here is based on a similar argument. The result is also true with $K = \infty$, but the proof is then more involved.

PROOF OF LEMMA 5. The first inequality (27) follows from the continuum FKG inequality (see [14]) and Proposition 1 of [17].

To prove (28), it suffices [by (27)] to prove that, for $\varepsilon > 0$,

$$\inf_{\mathbf{u} \in \mathscr{U}(d,\nu,K)} P\big[\mathbf{u} \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^t \setminus B_{\rho(d)}\big] \geq 1 - \psi(t)^\nu - 2\varepsilon, \qquad d \text{ large}. \tag{29}$$

Assume $t > 1$ (else there is nothing to prove). Define the lattice $\mathscr{L}$ and the regions $A_{i,j}$ as in (20) and (22). Choose $\delta \in (0, \varepsilon/3)$ such that $\theta_{\mathscr{L}}(1 - 3\delta) > 1 - \varepsilon$; this can be done by (21).

Let $(Z_n^d,\ n = 0, 1, 2, \dots)$ be a branching random walk (BRW) in $\mathbf{R}^d$, in which each particle gives birth to a Poisson number of offspring with mean t and the positions of the offspring of a particle at x are uniformly distributed over the ball $B_1(x)$. According to context, we shall regard Z_n^d either as a random subset of $\mathbf{R}^d$ or as the corresponding point measure: for $A \subset \mathbf{R}^d$, let $Z_n^d(A)$ denote the number of particles of the nth generation of this BRW in A.

Let (Z_n^∞) denote a BRW in $\mathbf{R}^2$, also with a Poisson(t) offspring distribution and the offspring of a particle at x having a normal distribution with mean x and variance matrix I. For $A \subset \mathbf{R}^2$, let $Z_n^\infty(A)$ denote the number of particles of the nth generation of this BRW in A.

By Lemma 2, the image under L of (Z_n^d) has approximately the distribution of (Z_n^∞). Since $t > 1$, by the proof of Lemma 2 of [16], there exist $m > 0$ and $k_1 > 0$ such that, for sufficiently large d,

$$P\Big[\big\{Z_{k_1}^d(A_{1,1}) > m\big\} \cap \big\{Z_{k_1}^d(A_{1,-1}) > m\big\}\Big] > 1 - \delta \quad \text{if } Z_0^d(A_{0,0}) \geq m. \tag{30}$$

Write P_x for probability referring to the BRW Z_n^∞ (or Z_n^d) with Z_0^∞ consisting of a single progenitor at x. The proof of Lemma 3 of [16] can be used to show that there exists $k_0 > 0$ such that, for any $x \in \mathbf{R}^2$ with $|x| \leq K$,

$$P_x\Big[\big\{Z_{k_0}^\infty(L(A_{1,1})) < m\big\} \cup \big\{Z_{k_0}^\infty(L(A_{1,-1})) < m\big\}\Big] < (1 + \delta)^{1/\nu}\psi(t). \tag{31}$$

The argument in [16] refers to a BRW with bounded offspring distribution. To use it here, truncate the two-dimensional BRW Z_n^∞ by removing any child at a distance greater than M, say, from its parent. This gives a BRW with a bounded offspring distribution, and, by the choice of M, the survival probability of the underlying branching process for this truncated BRW can be taken to be close to $1 - \psi(t)$ (by the continuity of ψ). The argument in [16] applies directly to the truncated BRW, and if the statement (31) is true for the truncated BRW it is also true for (Z_n^∞) because it dominates the truncated BRW.

By (31), for large enough d,

$$P\left[\left\{Z_{k_0}^d(A_{1,1}) \geq m\right\} \cap \left\{Z_{k_0}^d(A_{1,-1}) \geq m\right\}\right] > 1 - \psi(t)^\nu - \delta \quad \text{if } Z_0^d \in \mathscr{U}(d,\nu,K). \tag{32}$$

Let $(G_n,\ n \geq 0)$ denote a Galton–Watson process with a Poisson(t) offspring distribution. We can (and do) assume that k_0 is so large that

$$P\left[G_{k_0} > 0 | G_0 = \nu\right] < 1 - \psi(t)^\nu + \varepsilon. \tag{33}$$

Choose R to be so large that

$$P_0\left[\sum_{n=0}^{\max(k_1, k_0)} Z_n^\infty(B_{R-1}^c) > 0\right] < \delta/m. \tag{34}$$

This implies that, for large enough d, if Z_0^d consists of m points in $A_{i,j}$, then the probability that any of the first k_1 or the first k_0 generations of (Z_n^d) lies outside $L^{-1}(B_R(i,j))$ is at most δ, where $B_R(i,j)$ is the two-dimensional disk of radius R centered at (i,j).

Choose k_2 so large that, for the Galton–Watson process (G_n),

$$\max\left(P\left[\sum_{n=0}^{k_1} G_n > k_2 | G_0 = m\right], P\left[\sum_{n=0}^{k_0} G_n > k_2 | G_0 = \nu\right]\right) < \delta. \tag{35}$$

We now describe an algorithm consisting of a sequence of steps, indexed by the sites (i,j) of $\mathscr{L}$, taken in the same order as in the proof of Lemma 3. In step (i,j), first define finite sets S_{ij} and ζ_{ij} in $\mathbf{R}^d$; we set $S_{00} = \varnothing$, the empty set, and $\zeta_{00} = \mathbf{u}$, and define subsequent S_{ij} and ζ_{ij} later on. The set S_{ij} includes all points which have already been "examined" before step (i,j); roughly, the "examination" of a point corresponds to the observation of all points of $\mathscr{P}_d^t$ in its 1-neighborhood (by which we mean the translate of B_1 centered at that point). Using $Z_0^d = \zeta_{ij}$ as the set of progenitors in generation 0, run the BRW Z_n^d, $n = 0, 1, \dots, k(i)$, for $k(i)$ generations, where we set $k(i) = k_1$ (defined above) for $i \geq 1$, and $k(0) = k_0$ as defined above. Order the particles of the BRW as follows: the particles of an earlier generation precede those of a later one, siblings are ranked in order of increasing modulus and particles in the same generation with distinct parents inherit the ordering of their parents (a procedure analogous to the class system in British society). Points of ζ_{00} are ordered by modulus.

Modify the BRW as follows. Consider successively each particle X after generation 0 of the BRW, in the ordering given above (starting with the particle of smallest modulus of generation 1). Remove X (along with its descendants) if it lies in the 1-neighborhood of any point of S_{ij}, or in the 1-neighborhood of any unremoved particle of the BRW that precedes X in the ordering, or if X lies inside $B_{\rho(d)}$ or outside $L^{-1}(B_R(i,j))$. Finally, if there are more than k_2 remaining particles in generations $0,1,2,\dots,k(i)$, remove all but the first k_2 remaining particles in the ordering.

Let σ_{ij} denote the set of all unremoved particles in generations $0,1,2,\dots,k(i)$. By construction, this set has cardinality at most k_2. The set σ_{ij} includes the set of points examined in the course of step (i,j) and also unexamined points from the final generation, some of which may be examined later on.

Step (i,j) is deemed to be "successful" if (i) $Z^d_{k(i)}(A_{i+1,j+1}) \geq m$ and $Z^d_{k(i)}(A_{i+1,j-1}) \geq m$; (ii) no particle has cause to be removed; and (iii) no particle of the final generation $k(i)$ lies within a distance of $3/4$ of its parent or siblings.

We need to initialize step (i,j) by defining the initial sets S_{ij} and ζ_{ij} for $i > 0$ (we defined $S_{00} = \varnothing$ and $\zeta_{00} = \mathbf{u}$ earlier on). These definitions will depend on the outcomes of earlier steps. Define S_{ij} to be the union of $\{0\}$ and all sets $\sigma_{i',j'}$ with (i',j') preceding (i,j) in the ordering on $\mathscr{L}$. Assume there is an oriented path in $\mathscr{L}$ from $(0,0)$ to $(i-1,j-1)$ or to $(i-1,j+1)$ of sites i',j' for which steps (i',j') were all successful [if not, nothing happens in step (i,j), and σ_{ij} is empty]. If there is a successful path to $(i-1,j-1)$, let ζ_{ij} be the set of m points of smallest modulus which are both in $A_{i,j}$ and in the last generation of the (successful) BRW run at step $(i-1,j-1)$. If there is a successful path to $(i-1,j+1)$ but not to $(i-1,j-1)$, define ζ_{ij} similarly using the last generation from step $(i-1,j+1)$. Note that ζ_{ij} consists of m points in the region A_{ij}, each of which is distant at least $3/4$ from all other points in S_{ij}.

For sufficiently large d, the probability that step $(0,0)$ is "successful" exceeds $1-\psi(t)^{\nu}-3\delta$, by (32), (34), (35) and Lemma 1. For each (i,j) other than $(0,0)$ in $\mathscr{L}$, the probability that step (i,j) satisfies conditions (i)–(iii) to be "successful," given that it is attempted at all, exceeds $1-3\delta$. Indeed, (i) is likely to hold by (30). Also, the rule that each step has particles outside $L^{-1}(B_R(i,j))$ discarded means that only a finite number of previous steps [namely, those (i',j') with $|(i',j')-(i,j)| \leq 2R$] can possibly affect step (i,j), and all such (i',j') satisfy $\operatorname{card}(\sigma_{i',j'}) \leq k_2$. By this fact, (34), (35) and Lemma 1, the probability that (ii) or (iii) fails can be shown to be small for large d.

Let S_∞ denote the set of all points created (and not removed) during the course of the algorithm, together with the initial set of points $\mathbf{u}$. Then, for all $x,y \in S_\infty$, $x \leftrightarrow_1 y$ in S_∞. The algorithm is equivalent to the observation of the points of $\mathscr{P}^t_d$ in successive disjoint regions of $\mathbf{R}^d \setminus B_{\rho(d)}$ (see [17], Section 4 for more details), so $S_\infty \setminus \mathbf{u}$ may be viewed as a subset of $\mathscr{P}^t_d \setminus B_{\rho(d)}$, and $\operatorname{card}(S_\infty) = \infty$ implies $\mathbf{u} \leftrightarrow_1 \infty$ in $\mathscr{P}^t_d \setminus B_{\rho(d)}$. A comparison with oriented percolation with parameter $1-3\delta$ shows that the probability that there are infinitely many successful steps in the algorithm exceeds $(1-\psi(t)^{\nu}-3\delta)(1-\varepsilon)$; therefore, we have (29). $\square$

Let $\mathscr{U}'(d,\nu,K)$ denote the set of pairs $(\mathbf{u},x)$ with $\mathbf{u}\in\mathscr{U}(d,\nu,K)$ and $\mathbf{u}\cup\{x\}\in\mathscr{U}(d,\nu+1,K)$. The next result says that, for such a $(\mathbf{u},x)$, if there is a 1-path from $\mathbf{u}$ to x there is also likely to be an infinite 1-path from $\mathbf{u}$, when d is large.

LEMMA 6. *Let $t>0$, $\nu\in\mathbf{N}$ and $K>0$. Let $(\rho(d),\,d\ge 1)$ be a sequence of positive numbers with $\limsup_{d\to\infty}\rho(d)\le 1$. Let $E_{\mathbf{u},x}$ (resp. $E_{\mathbf{u},\infty}$) denote the event that $\mathbf{u}\leftrightarrow_1 x$ (resp. $\mathbf{u}\leftrightarrow_1\infty$) in $\mathscr{P}_d^t\setminus B_{\rho(d)}$. Then*

$$\lim_{d\to\infty}\ \sup_{(\mathbf{u},x)\in\mathscr{U}'(d,\nu,K)} P[E_{\mathbf{u},x}\setminus E_{\mathbf{u},\infty}]=0. \tag{36}$$

PROOF. Let $\varepsilon>0$. Consider the same algorithm as in the proof of Lemma 5, again viewed as generating a random subset S_∞ of $\mathscr{P}_d^t\setminus B_{\rho(d)}$. Step $(0,0)$ of this algorithm is a BRW running for k_0 steps, here denoted $(Z_0^d,Z_1^d,\dots,Z_{k_0}^d)$, with $Z_0^d=\mathbf{u}$. Let E_0 be the event that step $(0,0)$ is "successful" in the sense of the earlier proof. Then

$$\begin{aligned}&P[E_{\mathbf{u},x}\setminus E_{\mathbf{u},\infty}]\\&\quad\le P\big[E_{\mathbf{u},x}\setminus\{Z_{k_0}^d\ne\varnothing\}\big]+P\big[\{Z_{k_0}^d\ne\varnothing\}\setminus E_0\big]+P[E_0\setminus E_{\mathbf{u},\infty}].\end{aligned}\tag{37}$$

It suffices to show that each of the three terms on the right-hand side of (37) is bounded by 2ε for large d, uniformly on $(\mathbf{u},x)\in\mathscr{U}'(d,\nu,K)$. First, by Lemma 1, for d large,

$$P\big[E_{\mathbf{u},x}\setminus\{Z_{k_0}^d\ne\varnothing\}\big]\le P\left[x\in\bigcup_{n=0}^{k_0-1}\ \bigcup_{y\in Z_n^d}B_1(y)\right]<\varepsilon. \tag{38}$$

As in the proof of Lemma 5, by (32), (34), (35) and Lemma 1, for large d we have $P[E_0]\ge 1-\psi(t)^\nu-\varepsilon$; therefore, by the choice of k_0 to satisfy (33),

$$P\big[\{Z_{k_0}^d\ne\varnothing\}\setminus E_0\big]=P\big[Z_{k_0}^d\ne\varnothing\big]-P[E_0]<2\varepsilon. \tag{39}$$

The comparison with oriented percolation in the proof of Lemma 5 shows that, for large d, if step $(0,0)$ is successful, then, with probability exceeding $1-\varepsilon$, there are infinitely many successful steps, in which case $E_{\mathbf{u},\infty}$ occurs; thus $P[E_0\setminus E_{\mathbf{u},\infty}]<\varepsilon$, which completes the proof. □

5. Proof of main results. List the points of $\mathscr{P}_d\setminus\{0\}$ in order of increasing modulus as $Y_1,Y_2,Y_3,\dots$. Fix $k\in\mathbf{N}$ and numbers $0<t_1<t_2<\cdots<t_k\le s<\infty$. Let $\Theta_1,\dots,\Theta_k$ be d-dimensional random variables, independent of one another and of $\mathscr{P}_d$, uniformly distributed on the unit sphere in $\mathbf{R}^d$. Define the d-dimensional variables $Y_i'=r_d(t_i)\Theta_i$ and the point processes

$$\mathscr{Y}_i=\{Y_1',\dots,Y_i'\},\qquad 1\le i\le k. \tag{40}$$

Define the modified point process

$$\mathscr{P}_d'=\{0\}\cup\mathscr{Y}_k\cup\big(\mathscr{P}_d\setminus B_{r_d(s)}\big). \tag{41}$$

The conditional distribution of $\mathscr{P}_d$, given that $\operatorname{card}\{X \in \mathscr{P}_d\colon\ 0 < |X| \le r_d(s)\} = k$ and that $|Y_i| = r_d(t_i)$ for $1 \le i \le k$, is that of $\mathscr{P}'_d$.

For $1 \le i \le k$, define the events

$$E_i = \{0 \leftrightarrow_{r_d(t_i)} \infty \text{ in } \mathscr{P}'_d\}, \qquad F_i = \{Y'_i \leftrightarrow_{r_d(t_i)} \infty \text{ in } P'_d\} \tag{42}$$

and

$$H_i = \{0 \leftrightarrow_{r_d(t_i)} Y'_i \text{ in } P'_d\}. \tag{43}$$

If $E_i \cap F_i$ occurs, then $(0, Y'_i)$ is an edge neither of $t_\infty(0, \mathscr{P}_d)$ nor of $t_\infty(Y'_i, \mathscr{P}_d)$; in other words, $E_i \cap F_i \subset \{(0, Y'_i) \notin g(\mathscr{P}'_d, 0)\}$, and we now show that the probability that the inclusion is strict vanishes as $d \to \infty$. If $(0, Y'_i) \notin g(\mathscr{P}'_d, 0)$ but $E_i \cap F_i$ does not hold, then necessarily H_i holds; therefore,

$$\begin{aligned} P\big[\{(0,Y'_i) \notin g(\mathscr{P}'_d,0)\} \setminus (E_i \cap F_i)\big] &\le P[H_i \setminus E_i] + P[H_i \setminus F_i] \\ &\le P\big[\{\mathscr{Y}_{i-1} \leftrightarrow_{r_d(t_i)} (\mathscr{Y}_k \setminus \mathscr{Y}_{i-1}) \text{ in } \mathscr{P}_d \setminus B_{r_d(s)}\} \\ &\qquad \setminus \{\mathscr{Y}_{i-1} \leftrightarrow_{r_d(t_i)} \infty \text{ in } \mathscr{P}_d \setminus B_{r_d(s)}\}\big] \\ &\quad + P\big[\{Y'_i \leftrightarrow_{r_d(t_i)} (\mathscr{Y}_k \setminus \{Y'_i\}) \text{ in } \mathscr{P}_d \setminus B_{r_d(s)}\} \\ &\qquad \setminus \{Y'_i \leftrightarrow_{r_d(t_i)} \infty \text{ in } \mathscr{P}_d \setminus B_{r_d(s)}\}\big]. \end{aligned} \tag{44}$$

We shall prove that this vanishes by using Lemma 6. First we introduce more notation; define the numbers

$$\rho_i(d) = r_d(s)/r_d(t_i), \tag{45}$$

the vectors

$$W^i_j := Y'_i/r_d(t_i) = (r_d(t_m)/r_d(t_i))\Theta_j \tag{46}$$

and the sets

$$\mathbf{u}^i_j = \{W^i_1, \dots, W^i_j\}. \tag{47}$$

Since the image of $\mathscr{P}'_d$ under $x \mapsto x/r_d(t_i)$ is a Poisson process $\mathscr{P}^{t_i}_d$ of rate t_i/π_d, the upper bound in (44) is equal to

$$\begin{aligned} &P\big[\{\mathbf{u}^i_{i-1} \leftrightarrow_1 (\mathbf{u}^i_k \setminus \mathbf{u}^i_{i-1}) \text{ in } \mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}\} \setminus \{\mathbf{u}^i_{i-1} \leftrightarrow_1 \infty \text{ in } \mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}\}\big] \\ &\quad + P\big[\{W^i_i \leftrightarrow_1 (\mathbf{u}^i_k \setminus \{W^i_i\}) \text{ in } \mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}\} \setminus \{W^i_i \leftrightarrow_1 \infty \text{ in } \mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}\}\big]. \end{aligned} \tag{48}$$

As $d \to \infty$, all ratios $r_d(t_i)/r_d(t_j)$ converge to 1. In particular, $\rho_i(d) \to 1$. Also, Lemmas 1 and 2 still hold with the vector $U(d)$ taken to be uniform on the unit *sphere* $\{x \in \mathbf{R}^d\colon |x| = 1\}$. Therefore, for $K \in (0, \infty)$,

$$P\big[\mathbf{u}^i_k \in \mathscr{U}(d,k,K)\big] \to (P[|Z| < K])^k, \tag{49}$$

where Z is a two-dimensional standard normal. This limit can be made arbitrarily close to 1 by choice of K. Therefore, by Lemma 6, the expression in (48) converges to 0 as $d \to \infty$, and so

$$\lim_{d\to\infty} P\big[\{(0,Y_i') \notin g(\mathscr{P}_d', 0)\} \triangle (E_i \cap F_i)\big] = 0, \tag{50}$$

where $\triangle$ denotes symmetric difference.

Clearly, $E_1 \subset E_2 \subset E_3 \subset \cdots$. Define the random variable

$$I = \max\{i \le k : E_i \text{ does not hold}\} \tag{51}$$

and the random set

$$S = \{i \in \{I+1, I+2, \dots, k\} : F_i \text{ occurs}\}. \tag{52}$$

By (50), except on an event of small probability, the set $\{1,2,\dots,k\} \setminus S$ is the set of $i \le k$ for which $(0, Y_i')$ is an edge of $g(\mathscr{P}_d', 0)$.

Let $i \in \{1,2,\dots,k\}$ and $\sigma \subset \{i+1, i+2, \dots, k\}$ with cardinality denoted $|\sigma|$. Write σ^c for the set $\{i+1, i+2, \dots, k\} \setminus \sigma$. Then

$$\begin{aligned} P[I = i, S = \sigma] &= P\left[E_{i+1} \cap E_i^c \cap \left(\bigcap_{j\in\sigma} F_j\right) \cap \left(\bigcap_{j\in\sigma^c} F_j^c\right)\right] \\ &= P\left[E_{i+1} \cap \left(\bigcap_{j\in\sigma} F_j\right) \cap \left(\bigcap_{j\in\sigma^c} F_j^c\right)\right] - P\left[E_i \cap \left(\bigcap_{j\in\sigma} F_j\right) \cap \left(\bigcap_{j\in\sigma^c} F_j^c\right)\right]. \end{aligned} \tag{53}$$

By the inclusion–exclusion formula, the second of the two terms in (53) is given by

$$P\left[E_i \cap \left(\bigcap_{j\in\sigma} F_j\right) \cap \left(\bigcap_{j\in\sigma^c} F_j^c\right)\right] = \sum_{\tau\subset\sigma^c} (-1)^{|\tau|} P\left[E_i \cap \left(\bigcap_{j\in\sigma\cup\tau} F_j\right)\right]. \tag{54}$$

Consider one of the terms in this sum. By the FKG inequality, writing Θ for the vector $(\Theta_1, \dots, \Theta_k)$,

$$P\left[E_i \cap \left(\bigcap_{j\in\sigma\cup\tau} F_j\right) \Big| \Theta\right] \ge P[E_i|\Theta] \prod_{j\in\sigma\cup\tau} P[F_j|\Theta]. \tag{55}$$

By using the definition of E_i and applying the map $x \mapsto x/r_d(t_i)$ to $\mathbf{R}^d$, we have

$$\begin{aligned} P[E_i|\Theta] &\ge P\big[\mathscr{Y}_{i-1} \leftrightarrow_{r_d(t_i)} \infty \text{ in } \mathscr{P}_d \setminus B_{r_d(s)} |\Theta\big] \\ &= P\big[\mathbf{u}_{i-1}^i \leftrightarrow_1 \infty \text{ in } \mathscr{P}_d^{t_i} \setminus B_{\rho_i(d)} |\Theta\big]. \end{aligned} \tag{56}$$

Therefore, Lemma 5 gives us, for all $K > 0$,

$$\liminf_{d\to\infty} \inf_{\{\Theta:\, \mathbf{u}_k^i \in \mathscr{U}(d,k,K)\}} P[E_i|\Theta] \ge 1 - \psi(t_i)^{i-1}, \tag{57}$$

and, similarly,

$$\liminf_{d\to\infty} \inf_{\{\Theta:\, \mathbf{u}_k^i \in \mathscr{U}(d,k,K)\}} P[F_j|\Theta] \ge 1 - \psi(t_j) = \bar{\psi}(t_j). \tag{58}$$

Therefore, by (55), integrating over (Θ) and using (49),

$$\liminf_{d\to\infty} P\left[E_i \cap \left(\bigcap_{j\in\sigma\cup\tau} F_j\right)\right] \geq \left(1-\psi(t_i)^{i-1}\right)\prod_{j\in\sigma\cup\tau}\bar{\psi}(t_j). \tag{59}$$

To get an upper bound on $P[E_i \cap (\bigcap_{j\in\sigma\cup\tau} F_j)]$, first rescale $\mathbf{R}^d$ by a factor of $1/r_d(t_i)$ as in (56) to obtain

$$\begin{aligned} &P\left[E_i \cap \left(\bigcap_{j\in\sigma\cup\tau} F_j\right)\right] \\ &\quad\leq P\Bigg[\Bigg[\left\{\mathbf{u}^i_{i-1} \leftrightarrow_1 \infty \text{ in } \mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}\right\} \\ &\qquad\cap\left(\bigcap_{j\in\sigma\cup\tau}\left\{W^i_j \leftrightarrow_{r_d(t_j)/r_d(t_i)} \infty \text{ in } \mathscr{P}^{t_i}_d\setminus B_{\rho_i(d)}\right\}\right)\Bigg] \cup H'\Bigg], \end{aligned} \tag{60}$$

where H' is the event that $\mathbf{u}^i_{i-1} \leftrightarrow_1 \mathbf{u}^i_k \setminus \mathbf{u}^i_{i-1}$ in $\mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}$, or, for some $j \in \sigma\cup\tau$ and $m \in \{1,2,\dots,k\}\setminus\{j\}$, $W^i_j \leftrightarrow_1 W^i_m$ in $\mathscr{P}^{t_i}_d \setminus B_{\rho_i(d)}$.

Let $((G^j_n,\ n\geq 0),\ j\in\{i\}\cup\sigma\cup\tau)$ be independent Galton–Watson branching processes, such that each $(G^j_n,\ n\geq 0)$ has a Poisson(t_j) offspring distribution, with initial values $G^i_0 = i-1$ and $G^j_0 = 1$ for each $j\in\sigma\cup\tau$. Let $\varepsilon>0$. Take k_0 so that

$$P\left[\left\{G^i_{k_0} > 0\right\} \cap \left(\bigcap_{j\in\sigma\cup\tau}\left\{G^j_{k_0}>0\right\}\right)\right] < \left(1-\psi(t_i)^{i-1}\right)\prod_{j\in\sigma}\bar{\psi}(t_j) + \varepsilon. \tag{61}$$

For $j\in\{i\}\cup\sigma$, define "generations" $\mathscr{G}^j_1,\dots,\mathscr{G}^j_{k_0}$ of $\mathscr{P}^{t_i}_d\setminus B_{\rho_i(d)}$ as follows: initially set

$$\mathscr{G}^i_0 = \mathbf{u}^i_{i-1}, \qquad \mathscr{G}^j_0 = \left\{W^i_j\right\}, \qquad j\in\sigma\cup\tau.$$

Then, recursively for $n=1,2,\dots,k_0$, let $\mathscr{G}^j_n$ be the set of points of $\mathscr{P}^{t_i}_d\setminus B_{\rho_i(d)}$ which lie in $B_{r_d(t_j)/r_d(t_i)}(y)$ for some $y\in\mathscr{G}^j_{n-1}$, but which are not in any of $\mathscr{G}^j_0,\dots,\mathscr{G}^j_{n-1}$.

If $\mathbf{u}^i_{i-1}\leftrightarrow_1\infty$ in $\mathscr{P}^{t_i}_d\setminus B_{\rho_i(d)}$, it must be the case that $\mathscr{G}^i_{k_0}$ is nonempty; similarly, for each $j\in\sigma\cup\tau$, if $W^i_j \leftrightarrow_{r_d(t_j)/r_d(t_i)}\infty$ in $\mathscr{P}^{t_i}_d\setminus B_{\rho_i(d)}$, it must be the case that $\mathscr{G}^j_{k_0}$ is nonempty. Thus, by (60),

$$P\left[E_i\cap\left(\bigcap_{j\in\sigma\cup\tau}F_j\right)\right] \leq P\left[\bigcap_{j\in\{i\}\cup\sigma\cup\tau}\left\{G^j_{k_0}\neq\varnothing\right\}\right] + P[H''], \tag{62}$$

where we define H'' to be the event that the connection in the definition of H occurs before the k_0th generation, that is,

$$\begin{aligned} H'' := &\left\{\left(\mathbf{u}^i_k\setminus\mathbf{u}^i_{i-1}\right)\cap\left(\bigcup_{m=0}^{k_0}\bigcup_{y\in\mathscr{G}^i_m} B_1(y)\right)\neq\varnothing\right\} \\ &\cup\bigcup_{j\in\sigma\cup\tau}\left\{\left(\mathbf{u}^i_k\setminus\left\{W^i_j\right\}\right)\cap\left(\bigcup_{m=0}^{k_0}\bigcup_{y\in G^j_m} B_{r_d(t_j)/r_d(t_i)}(y)\right)\neq\varnothing\right\}. \end{aligned}$$

We can generate a sequence of sets with the same distribution as the generations $(\mathscr{G}_n^j,\ n \le k_0)$ by branching random walks. For $j \in \{i\} \cup \sigma \cup \tau$, let $(Z_n^{d,j},\ n = 0, 1, 2, \dots, k_0)$ be a BRW in $\mathbf{R}^d$, in which each particle gives birth to a Poisson number of offspring with mean t_j and the offspring of a particle at x are uniformly distributed over the ball $B_{r_d(t_j)/r_d(t_i)}(x)$. Let the initial value $Z_0^{d,j}$ of this BRW be given by the set $\mathbf{u}_i^{i-1}$ for $j = i$ and by the single point $\{W_j^i\}$ for $j \in \sigma \cup \tau$. These BRW's are to be run independently up to generation k_0.

If these BRW's $(Z_n^{d,j})$ are modified by removing some of the particles, in a similar manner to the modification of the BRW's in the proof of Lemma 5, they generate sets with the same distributions as the generations $\mathscr{G}_n^j$ described above, we omit the details. Thus each $\mathscr{G}_n^j$ has the distribution of a subset of $Z_n^{d,j}$, and so, by (62), we have

$$P\left[E_i \cap \left(\bigcap_{j \in \sigma \cup \tau} F_j\right)\right] \le P\left[\bigcap_{j \in \{i\} \cup \sigma \cup \tau} \left\{Z_{k_0}^{d,j} \ne \emptyset\right\}\right] + P[H'''], \tag{63}$$

where we define the event H''' in the same manner as H'', but with each set of the form G_n^j replaced by the set of points of $Z_n^{d,j}$.

Since each BRW $(Z_n^{d,j})$ runs for a fixed number of generations, the probability approaches 0 that it visits any given ball of radius close to 1, by Lemma 1; thus $P[H'''] \to 0$ as $d \to \infty$. Also, the population sizes of these BRW's are independent simple branching processes, so, by (61), for large enough d,

$$P\left[E_i \cap \left(\bigcap_{j \in \sigma \cup \tau} F_j\right)\right] < \left(1 - \psi(t_i)^{i-1}\right) \prod_{j \in \sigma \cup \tau} \bar{\psi}(t_j) + 2\varepsilon. \tag{64}$$

Since $\varepsilon > 0$ is arbitrary, (59) and (64) together yield

$$\lim_{d \to \infty} P\left[E_i \cap \left(\bigcap_{j \in \sigma \cup \tau} F_j\right)\right] = \left(1 - \psi(t_i)^{i-1}\right) \prod_{j \in \sigma \cup \tau} \bar{\psi}(t_j). \tag{65}$$

By applying (65) to each term in (54), we obtain

$$\begin{aligned}
&\lim_{d \to \infty} P\left[E_i \cap \left(\bigcap_{j \in \sigma} F_j\right) \cap \left(\bigcap_{j \in \sigma^c} F_j^c\right)\right] \\
&\qquad = \sum_{\tau \subset \sigma^c} (-1)^{|\tau|} \left(1 - \psi(t_i)^{i-1}\right) \prod_{j \in \sigma \cup \tau} \bar{\psi}(t_j) \\
&\qquad = \left(1 - \psi(t_i)^{i-1}\right) \left(\prod_{j \in \sigma} \bar{\psi}(t_j)\right) \prod_{j \in \sigma^c} \left(1 - \bar{\psi}(t_j)\right).
\end{aligned} \tag{66}$$

By arguing similarly for the other term in (53) and using the fact that $1 - \bar{\psi}(t) = \psi(t)$, we may conclude

$$\lim_{d \to \infty} P[I = i, S = \sigma] = \left(\psi(t_i)^{i-1} - \psi(t_{i+1})^i\right) \prod_{j \in \sigma} \bar{\psi}(t_j) \prod_{j \in \sigma^c} \psi(t_j). \tag{67}$$

Let $Q_{i,\sigma}$ denote the event that $e(1),\dots,e(i)$ are the edges incident on $\varnothing$ of $t_\infty(\varnothing,\mathscr{T})$ and $\{e(j)\colon i+1\le j\le k,\ j\notin\sigma\}$ are the remaining edges up to $e(k)$ incident on $\varnothing$ in $g(\mathscr{T},\varnothing)$. Write $\mathbf{t}$ for $(t_1,\dots,t_k)$ and write $\{\Gamma_1^k(\varnothing)=\mathbf{t}\}$ for the event that, for $1\le j\le k$, the weight $\Gamma_j(\varnothing)$ of edge $e(j)$ is equal to t_j.

Recall from Section 3 that L_i is the maximum edge weight in the graph $t_\infty(i,\mathscr{T}_i)$ and the L_i are independent and identically distributed. By (10) and (67),

$$\begin{aligned}\lim_{d\to\infty} & P[I=i,S=\sigma]\\ &= P\Bigg[\Big\{\min_{1\le j\le i-1} L_j>t_i\Big\}\cap\Big\{\min_{1\le j\le i} L_j<t_{i+1}\Big\}\\ &\qquad\cap\Big(\bigcap_{j\in\sigma}\{L_j<t_j\}\Big)\cap\Big(\bigcap_{j\in\sigma^c}\{L_j>t_j\}\Big)\Bigg]\\ &= P\big[Q_{i,\sigma}|\Gamma_1^k(\varnothing)=\mathbf{t}\big].\end{aligned}\tag{68}$$

Recall from (1) [resp. (3)] that $D_t(d)$ [resp. $D_t(\infty)$] is the degree of 0 (resp. $\varnothing$) in the graph $g(\mathscr{P}_d,0)$ [resp. $g(\mathscr{T},\varnothing)$] with all edges of length (weight) greater than t removed.

The variable $\mathrm{card}(\mathscr{P}_d\setminus\{0\})\cap B_{r_d(s)}$ is Poisson with mean s; conditional on its taking the value k, the distribution of $(v_d(|Y_1|),\dots,v_d(|Y_k|))$ is uniform on $\Delta_k(s):=\{\mathbf{t}=(t_1,\dots,t_k)\colon 0<t_1<\cdots<t_k\le s\}$. Therefore, by the remark following the definition (41) of $\mathscr{P}_d'$,

$$\begin{aligned}P&\big[D_{r_d(s)}(d)=m\big]\\ &=\sum_{k=m}^{\infty}\frac{e^{-s}s^k}{k!}\int_{\Delta_k(s)}\frac{d\mathbf{t}}{k!}\sum_{\{(i,\sigma)\colon k-i-|\sigma|=m\}}P[I=i,S=\sigma]\\ &\to\sum_{k=m}^{\infty}\frac{e^{-s}s^k}{k!}\int_{\Delta_k(s)}\frac{d\mathbf{t}}{k!}\sum_{\{(i,\sigma)\colon k-i-|\sigma|=m\}}P\big[Q_{i,\sigma}|\Gamma_1^k(\varnothing)=\mathbf{t}\big]\\ &=P\big[D_s(\infty)=m\big].\end{aligned}\tag{69}$$

Thus $D_{r_d(s)}$ converges weakly to $D_s(\infty)$ as $d\to\infty$. Also, $D_{r_d(s)}\le\mathrm{card}(\mathscr{P}_d\cap B_{r_d(s)})$, a Poisson variable with mean s; thus $P[D_{r_d(s)}\ge k]$ is uniformly bounded by $\sum_{m=k}^{\infty}e^{-s}s^m/m!$, which is summable in k. Therefore, $E[D_{r_d(s)}(d)]\to E[D_s(\infty)]$, and Theorem 2 is proved.

The proof of Theorem 1 is completed by taking $s\to\infty$ in (69), using Lemma 4 and a routine argument.

6. Invasion percolation. In invasion percolation on the integer lattice $\mathbf{Z}^d$ (made into a graph by including bonds between all nearest-neighbor pairs), independent random weights, uniformly distributed on $[0,1]$, are assigned to the bonds of the lattice. A random sequence of subgraphs (C_n) of $\mathbf{Z}^d$ (the "invaded cluster at time n") is defined as follows. Initially, $C_0=\{0\}$.

Given C_{n-1}, define C_n by adding the edge of $\mathbf{Z}^d$ with smallest weight out of those edges not in C_n with at least one end at a vertex of C_n. If the added edge has an end in $\mathbf{Z}^d \setminus C_{n-1}$, it is denoted a *breakout bond*; otherwise, the edge added is denoted a *backfill bond*; the terminology is from [3]. Let $C_\infty = \bigcup_{n=1}^{\infty} C_n$, and let C'_∞ denote the subgraph of C_∞ obtained by deleting backfill bonds. As pointed out by Alexander [3], this model is related to the MST; indeed, C'_∞ is precisely $t_\infty(0, \mathbf{Z}^d)$, as given by the algorithm in Section 2, applied to the graph $\mathbf{Z}^d$ with the given edge weights.

Let $\beta_{k,d}$ denote the probability that 0 has degree k in C_∞, and let $\overline{D}_s(d)$ denote the number of edges of weight at most s in C_∞. It seems likely that the methods of the present paper can be adapted to prove the following statements about weak convergence of the structure of C_∞ to that of $t_\infty(\emptyset, \mathscr{T})$; let $D^-(\infty) + 1$ denote the degree of 0 in that graph, and let D_s^- denote the number of $i \in \mathbf{N}$ with $e(i) \in t_\infty(\emptyset, \mathscr{T})$ and $w(i) \le s$.

CONJECTURE 1. *For each* $k \in \mathbf{N}$, $\lim_{d \to \infty} \beta_{k,d} = P[D^-(\infty) + 1 = k]$. *Also*, $\lim_{d\to\infty} E[\overline{D}_{2ds}(d)] = E[D_s^-(\infty)]$ *for each* $s > 0$.

To prove this, one might first consider C'_∞. Arrange the $2d$ sites of $\mathbf{Z}^d$ adjacent to 0 in order of increasing associated weight as $Y_1, Y_2, \dots,$ and let R_i denote the weight of edge $(0, Y_i)$. It is a standard result in extreme value theory (see, e.g., [18], Proposition 3.21) that the point process $\{2dR_i,\ i \ge 1\}$ converges in distribution to a homogeneous Poisson process of rate 1 on $(0, \infty)$.

The edges from 0 of C'_∞ are the I edges of smallest weight for some (random) I; for fixed $k \in \mathbf{N}$, $I \ge k$ if and only if there is no infinite path from 0 in $\mathbf{Z}^d$ of edges of weight less than R_i. By a similar argument to Lemma 5, it should be possible to prove

$$\lim_{d\to\infty} P[I \ge k | 2dR_k = s] = \psi(s)^{k-1},$$

and this is the probability that $e(k)$ is an edge of $t_\infty(\emptyset, \mathscr{T})$, given that $\Gamma_\nu(\emptyset) = s$. Finally, it should be possible to prove that $\lim_{d\to\infty} P[\exists$ backfill bond from 0$] = 0$.

The limits in the above conjecture are given by the formulae

$$E[D_t^-(\infty)] = \int_0^t \psi(s)\, ds \tag{70}$$

and

$$P[D^-(\infty) = k] = \int_0^1 \left(e^{-h(u)} (h(u))^k / k!\right) du, \tag{71}$$

where we set

$$h(u) = \left(\frac{1-u}{u}\right) \log\left(\frac{1}{1-u}\right).$$

We omit the details of the proofs, but (70) follows from Palm theory for the marked Poisson process, and (71) follows from the fact that, given $L = s$, the conditional distribution of $D^{-}(\infty)$ is Poisson with mean $s\psi(s)$; see [1], Lemma 10.

APPENDIX

PROOF OF PROPOSITION 3. First we prove (2) in the special case of the uniform distribution on the unit cube, with probability density function denoted f_U. Let $M_{n,i}$ denote the number of edges in the MST on $\{\eta_1, \dots, \eta_n\}$ of length at most $n^{-1/d}t$ with η_i as an endpoint. Then

$$F_n(n^{-1/d}t; f_U) = (1/2)(n-1)^{-1} \sum_{i=1}^{n} M_{n,i}. \tag{72}$$

Taking expectations, using the exchangeability of $\eta_1, \dots, \eta_n$ and applying Proposition 9 of [2], we obtain

$$EF_n(n^{-1/d}t; f_U) = \frac{n}{2(n-1)} EM_{n,1} \to (1/2) ED_t(d). \tag{73}$$

By (72) and exchangeability,

$$\begin{aligned} &\operatorname{Var}(F_n(n^{-1/d}t; f_U)) \\ &\quad = \frac{1}{4(n-1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(E[M_{n,i} M_{n,j}] - E[M_{n,1}]^2 \right) \\ &\quad = \frac{n}{4(n-1)^2} \operatorname{Var}(M_{n,1}) + \frac{n}{4(n-1)} \operatorname{Cov}[M_{n,1}, M_{n,2}]. \end{aligned} \tag{74}$$

By Lemma 2 of [2], the degrees of the edges of the MST are uniformly bounded by a constant depending only on d, so the first term on the right-hand side of (74) vanishes as $n \to \infty$. We now show that the second term also vanishes.

Let $\mathscr{M}$ denote the space of locally finite subsets of $\mathbf{R}^d$, with the (metrizable) topology of vague convergence as given in [2] or [18], Proposition 3.13. A point process is a random element of $\mathscr{M}$. Consider now the random element $(\mathscr{N}_{n,1}, \mathscr{N}_{n,2})$ of the product space $\mathscr{M} \times \mathscr{M}$, where, for $i = 1, 2$, we define $\mathscr{N}_{n,i} = \{n^{1/d}(\eta_j - \eta_i), 1 \le j \le n\}$, the rescaled empirical point process of the n points centered at η_i. Let $\mathscr{P}_{d,1}$ and $\mathscr{P}_{d,2}$ denote independent copies of the Poisson process (with a point added at 0) $\mathscr{P}_d$. By an easy generalization of Lemma 8 of [2],

$$(\mathscr{N}_{n,1}, \mathscr{N}_{n,2}) \to_d (\mathscr{P}_{d,1}, \mathscr{P}_{d,2}) \quad \text{as } n \to \infty, \tag{75}$$

where $\to_d$ denotes weak convergence in $\mathscr{M} \times \mathscr{M}$. Using the Skorohod representation theorem, one may take versions of $(\mathscr{N}_{n,1}, \mathscr{N}_{n,2})$ and $(\mathscr{P}_{d,1}, \mathscr{P}_{d,2})$ such that $(\mathscr{N}_{n,1}, \mathscr{N}_{n,2})$ converges almost surely to $(\mathscr{P}_{d,1}, \mathscr{P}_{d,2})$.

Let $C(L)$ denote the cube $[-L, L]^d$. By a simple adaptation of the proof of Proposition 9 of [2], given any increasing sequence n of positive integers, there is a subsequence (n_k) such that, as $n \to \infty$ along the subsequence, for each rational L_1 and L_2,

$$\begin{aligned}&\left(\sum_{\xi \in \mathscr{N}_{n,1} \cap C(L_1)} d(\xi, g(\mathscr{N}_{n,1})), \sum_{\xi \in \mathscr{N}_{n,2} \cap C(L_2)} d(\xi, g(\mathscr{N}_{n,2}))\right)\\ &\quad \to \left(\sum_{\xi \in \mathscr{P}_{d,1} \cap C(L_1)} d(\xi, g(\mathscr{P}_{d,1}, \xi)),\right.\\ &\qquad \left.\sum_{\xi \in \mathscr{P}_{d,2} \cap C(L_2)} d(\xi, g(\mathscr{P}_{d,2}, \xi))\right) \quad \text{a.s.,}\end{aligned} \tag{76}$$

where $d(\xi, \mathscr{G})$ denotes the degree of ξ in a graph $\mathscr{G}$ and $g(\mathscr{N}_{n,i})$ is the minimal spanning tree on the finite point process $\mathscr{N}_{n,i}$. By Lemma 6(c) of [2], it follows from (76) that with these realizations the graphs $g(\mathscr{N}_{n_k,1})$ and $g(\mathscr{N}_{n_k,2})$ converge a.s. to $g(\mathscr{P}_{d,1}, 0)$ and $g(\mathscr{P}_{d,2}, 0)$, respectively. Therefore,

$$M_{n_k,1} M_{n_k,2} \to D_t(d,1) D_t(d,2) \quad \text{a.s.,}$$

where, for $i = 1, 2$, we set $D_t(d, i)$ to be the number of edges of length at most t from 0 in the graph $g(\mathscr{P}_{d,i}, 0)$. Since the quantities $M_{n_k,1}$ and $M_{n_k,2}$ are uniformly bounded, it follows that $E[M_{n_k,1} M_{n_k,2}] \to ED_t(d)^2$, so that the second term on the right-hand side of (74) converges to 0. Therefore, $F_n(n^{-1/d}t; f_U)$ converges in mean square to $E[D_t(d)]/2$, as asserted.

The proof of (2) for general f is similar. With probability 1, X_1 lies at a Lebesgue point x of f with $f(x) > 0$; see, for example, [20], Theorem 7.7. For any such x, the point process $\{n^{1/d} f(x)^{1/d}(\eta_i - x), 2 \le i \le n\}$ converges in distribution to a Poisson process of rate 1 on $\mathbf{R}^d$; see, for example, [18], Proposition 3.21. Therefore, $E[M_{n,1} | X_1 = x] \to E[D_{f(x)^{1/d}t}]$ as $n \to \infty$, and so

$$E[M_{n,1}] \to \int_{R^d} E\left[D_{f(x)^{1/d}t}\right] f(x)\, dx.$$

Also, for distinct Lebesgue points x, y of f with $f(x) > 0$ and $f(y) > 0$, the point processes $\{n^{1/d} f(x)^{1/d}(\eta_i - x), 3 \le i \le n\}$ and $\{n^{1/d} f(y)^{1/d}(\eta_i - x), 3 \le i \le n\}$ converge in distribution to independent Poisson processes of rate 1 on $\mathbf{R}^d$. Therefore,

$$E[M_{n,1} M_{n,2} | X_1 = x, X_2 = y] \to E\left[D_{f(x)^{1/d}t}\right] E\left[D_{f(y)^{1/d}t}\right].$$

Integrating over possible values of X_1 and X_2, we find that $\mathrm{Cov}(M_{n,1}, M_{n,2}) \to 0$, so that the expression in (74) vanishes as before. □

Acknowledgment. The author thanks Sungchul Lee for some useful comments and questions.

REFERENCES

[1] ALDOUS, D. (1990). A random tree model associated with random graphs. *Random Structures and Algorithms* **1** 383–402.

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

[3] ALEXANDER, K. (1995). Percolation and minimal spanning forests in infinite graphs. *Ann. Probab.* **23** 87–104.

[4] AVRAM, F. and BERTSIMAS, D. (1992). The minimum spanning tree constant in geometrical probability and under the independent model: a unified approach. *Ann. Appl. Probab.* **2** 113–130.

[5] BENTLEY, J. L. and FRIEDMAN, J. H. (1978). Fast algorithms for constructing minimal spanning trees in coordinate spaces. *IEEE Trans. Comput.* **27** 97–105.

[6] BERTSIMAS, D. J. and VAN RYZIN, G. (1990). An asymptotic determination of the minimum spanning tree and minimum matching constants in geometrical probability. *Oper. Res. Lett.* **9** 223–231.

[7] DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation.* Wadsworth and Brooks Cole, Pacific Grove, CA.

[8] DUSSERT, C., RASIGNI, M., PALMARI, J., RASIGNI, G., LLEBARIA, A. and MARTY, F. (1987). Minimal spanning tree analysis of biological structures. *J. Theoret. Biol.* **125** 317–323.

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

[10] FRIEDMAN, J. H. and RAFSKY, L. C. (1983). Graph-theoretic measures of multivariate association and prediction. *Ann. Statist.* **11** 377–391.

[11] GRIMMETT, G. R. and MARSTRAND, J. M. (1990). The supercritical phase of percolation is well behaved. *Proc. Roy. Soc. London Ser. A* **430** 439–457.

[12] KANG, A. N. C., LEE, R. C. T., CHANG, C.-L. and CHANG, S.-K. (1977). Storage reduction through minimal spanning trees and spanning forests. *IEEE Trans. Comput.* **26** 425–434.

[13] KESTEN, H. and LEE, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **6** 495–527.

[14] MEESTER, R. W. J. and ROY, R. (1995). *Continuum Percolation.* Cambridge Univ. Press.

[15] NEWMAN, C. M., RINOTT, Y. and TVERSKY, A. (1983). Nearest neighbors and Voronoi regions in certain point processes. *Adv. in Appl. Probab.* **15** 726–751.

[16] PENROSE, M. D. (1993). On the spread-out limit for bond and continuum percolation. *Ann. Appl. Probab.* **3** 253–256.

[17] PENROSE, M. D. (1996). Continuum percolation and Euclidean minimal spanning trees in high dimensions. *Ann. Appl. Probab.* **6** 528–544.

[18] RESNICK, S. I. (1987). *Extreme Values, Point Processes and Regular Variation.* Springer, New York.

[19] ROHLF, F. J. (1975). Generalization of the gap test for the detection of multivariate outliers. *Biometrics* **31** 93–101.

[20] RUDIN, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York.

[21] SCHILLING, M. F. (1986). Mutual and shared neighbor probabilities: finite- and infinite-dimensional results. *Adv. in Appl. Probab.* **18** 388–405.

[22] STEELE, J. M. (1993). Probability and problems in Euclidean combinatorial optimization. *Statist. Sci.* **8** 48–56.

[23] STEELE, J. M., SHEPP, L. A. and EDDY, W. F. (1987). On the number of leaves of a Euclidean minimal spanning tree. *J. Appl. Probab.* **24** 809–826.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF DURHAM
SOUTH ROAD
DURHAM DH1 3LE
UNITED KINGDOM
E-MAIL: mathew.penrose@durham.ac.uk