The Annals of Probability
1995, Vol. 23, No. 3, 1024–1060

PERCOLATION OF ARBITRARY WORDS IN $\{0,1\}^{\mathbb{N}}$

BY ITAI BENJAMINI[1] AND HARRY KESTEN[2]

Cornell University

Let $\mathscr{G}$ be a (possibly directed) locally finite graph with countably infinite vertex set $\mathscr{V}$. Let $\{X(v): v \in \mathscr{V}\}$ be an i.i.d. family of random variables with $P\{X(v) = 1\} = 1 - P\{X(v) = 0\} = p$. Finally, let $\xi = (\xi_1, \xi_2, \dots)$ be a generic element of $\{0,1\}^{\mathbb{N}}$; such a ξ is called a *word*. We say that the word ξ is *seen from the vertex* v if there exists a self–avoiding path $(v, v_1, v_2, \dots)$ on $\mathscr{G}$ starting at v and such that $X(v_i) = \xi_i$ for $i \geq 1$. The traditional problem in (site) percolation is whether $P\{(1,1,1,\dots)$ is seen from $v\} > 0$. So-called AB–percolation occurs if $P\{(1,0,1,0,1,0,\dots)$ is seen from $v\} > 0$. Here we investigate (a) whether $P\{$all words are seen from $v\} > 0$ (for a fixed v) and (b) whether $P\{$ all words are seen from some $v\} = 1$. We show that both answers are positive if $\mathscr{G} = \mathbb{Z}^d$, or even $\mathbb{Z}^d_+$ with all edges oriented in the "positive direction," when d is sufficiently large. We show that on the oriented $\mathbb{Z}^3_+$ the answer to (a) is negative, but we do not know the answer to (b) on $\mathbb{Z}^3_+$. Various graphs $\mathscr{G}$ are constructed (almost all of them trees) for which the set of words ξ which can be seen from a given v (or from some v) is large, even though it is w.p.1 not the set of all words.

1. Introduction. The traditional problem of (site) percolation on a countably infinite, possibly directed, graph $\mathscr{G}$ can be formulated as follows. Let $\mathscr{V}$ be the vertex set of $\mathscr{G}$; assume $\mathscr{V}$ is countably infinite. Assign independently to each $v \in \mathscr{V}$ a random variable $X(v)$, which takes the value 1 or 0 with probability p or $1-p$, respectively. Denote by P_p the corresponding measure on the space $\Omega := \{0,1\}^{\mathscr{V}}$, in which the configurations $\{X(v): v \in \mathscr{V}\}$ take their value. Hence P_p is just the product over $\mathscr{V}$ of the measures which assign masses p and $1-p$ to the points 1 and 0, respectively. We denote a typical point of Ω by ω, and sometimes write $X(v,\omega)$ for $X(v)$ to indicate the dependence of $X(v)$ on the configuration. The event $\{X(v) = 1\}$ ($\{X(v) = 0\}$) is often interpreted as "v is occupied" ("v is vacant," respectively). A *path* on $\mathscr{G}$ is a sequence $(v_1, v_2, \dots)$ of vertices in $\mathscr{V}$ such that $v_i \neq v_j$ for $i \neq j$ and such that there is an edge of $\mathscr{G}$ between v_i and v_{i+1}; in the directed case, this edge has to be oriented from v_i to v_{i+1}. On a number of occasions we shall take for $\mathscr{G}$ a rooted tree. Throughout we consider such a tree as being oriented away from the root; that is, in any path $(v_1, v_2, \dots)$ on such a tree, v_{i+1} has to be one step further away from the root than v_i. Note also that by our definition a path on any graph is automatically self–avoiding throughout this paper. Now

Received October 1993; revised October 1994.
[1]Supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.
[2]Supported by the NSF through a grant to Cornell University.
AMS 1991 *subject classification*. Primary 60K35.
Key words and phrases. Percolation, oriented percolation, words, graphs, trees.

define the so–called percolation probability:

$$\theta(p,v) = P_p\{\exists \text{ infinite path } (v, v_1, v_2, \dots) \text{ which starts at } v \text{ and has all } v_i \text{ with } i \geq 1 \text{ occupied}\}. \tag{1.1}$$

The first question in percolation theory was "When is $\theta(p,v) > 0$?" In many cases of interest it is easy to see from the FKG inequality that the answer is independent of v [compare Kesten (1982), Section 4.1]. Furthermore, it is also known in many cases that $\theta(p,v) > 0$ is equivalent to

$$\rho(p) := P_p\{\exists \text{ some path } (v_1, v_2, \dots) \text{ on } \mathscr{G} \text{ with all vertices occupied}\} = 1. \tag{1.2}$$

Mai and Halley (1980) introduced AB–percolation as a model for certain physical phenomena. Basically this investigates whether

$$P_p\{\exists \text{ some path } (v_1, v_2, \dots) \text{ on } \mathscr{G} \text{ with } X(v_{2i-1}) = 1, X(v_{2i}) = 0 \text{ for all } i \geq 1\}$$

equals 1 or not. For some graphs such alternating sequences of ones and zeroes do occur (for certain p), and for others they occur only with probability zero (no matter what p is); see Wierman (1989) and its references and Wierman and Appel (1987).

In this paper we study the occurrence or nonoccurrence of paths $(v_1, v_2, \dots)$ with $X(v_i) = \xi_i$, $i \geq 1$, for any prescribed sequence $\{\xi_i\}_{i\geq 1} \in \{0,1\}^{\mathbb{N}}$. (The above-mentioned cases correspond to $\xi_i \equiv 1$ for all i and to $\xi_{2i-1} = 1$, $\xi_{2i} = 0$, $i \geq 1$, respectively.) In particular we wish to know when the collection of sequences ξ which do occur is large in some sense. To make this more precise we introduce some notation. Let

$$\Xi = \{0,1\}^{\mathbb{N}}.$$

A generic element of Ξ is denoted by $\xi = (\xi_1, \xi_2, \dots)$ and is called a *word*. We write $\mathbb{1}$ for the special word $(1, 1, \dots)$. We say that the word ξ is *seen from the vertex v in the configuration ω* if there exists a path $(v, v_1, v_2, \dots)$ on $\mathscr{G}$, starting from v and such that $X(v_i, \omega) = \xi_i$, $i \geq 1$. We shall also say on occasion that the finite sequence $(\xi_1, \dots, \xi_n)$ is seen along the path $(v, v_1, v_2, \dots, v_m)$ with $n \leq m \leq \infty$, if $X(v_i) = \xi_i$, $1 \leq i \leq n$. Note that $X(v)$, at the initial point v, plays no role here; the ξ_i's have to equal $X(v_i)$ for $i \geq 1$ only. We write

$$S(v) = S(v,\omega) = \text{collection of words which are seen from } v \text{ (in the configuration } \omega).$$

More generally, for distinct vertices $v_1, \dots, v_k$,

$$S(v_1, v_2, \dots, v_k, \omega) = \bigcup_{1}^{k} S(v_i, \omega) = \text{collection of words which are seen from at least one } v_j, \quad 1 \leq j \leq k.$$

We also consider the collection of words seen from some vertex in $\mathscr{G}$,

$$S_\infty = S_\infty(\omega) := \bigcup_{v \in \mathscr{V}} S(v, \omega).$$

Clearly the largest these random sets can be is all of Ξ. Dekking (1989) seems to have been the first one to investigate when $S(v)$ can equal all of Ξ. Specifically, Dekking asked when all words can be seen from the root of a regular tree. Our principal results of a positive nature state that, somewhat surprisingly, this also occurs when $\mathscr{G} = \mathbb{Z}^d$ (undirected) for large d, or even when $\mathscr{G} = \mathbb{Z}_+^d$ with all edges of $\mathbb{Z}_+^d$ oriented in the positive direction (that is, there is a directed edge from v only to the vertices $v + e_j$, $1 \le j \le d$, where e_j is the jth coordinate vector).

Before we can tackle these questions we derive some auxiliary results in Section 2 which show that various events of interest are measurable. Let

$$\mathscr{B} = \sigma\text{-field generated by } \{X(v,\cdot)\colon\ v \in \mathscr{V}\}$$

($\mathscr{B}$ is a σ-field of subsets of Ω). Then the most important of the measurability results is as follows.

PROPOSITION 2. *If $\mathscr{G}$ is locally finite, then each of the events $\{S(v) = \Xi\}$, $\{S(v_1, \dots, v_k) = \Xi\}$ and $\{S_\infty = \Xi\}$ belongs to $\mathscr{B}$.*

(This is nontrivial only for $\{S_\infty = \Xi\}$.)

Even though all the problems below are meaningful for any value of p, and some of the proofs even go through (with minor modifications) for general p, *we restrict ourselves to $p = \frac{1}{2}$*. This seems a natural restriction—it creates a symmetry between the zeroes and ones and therefore rules out that a certain word is not seen for the trivial reasons that its frequency of ones or zeroes is too high. To simplify the notation we shall therefore simply write P instead of $P_{1/2}$. *Throughout the paper all probabilities concerning the $X(v)$ are therefore calculated at $p = \frac{1}{2}$.*

Here is our most specific result.

THEOREM 1. *Let $\mathscr{G} = \mathbb{Z}_+^d$ with all edges oriented in the positive direction, as described above. Then for $d \ge 10$,*

$$P\{S_\infty = \Xi\} = 1, \tag{1.3}$$

and for $d \ge 40$,

$$P\{S(v) = \Xi \text{ for some } v\} = 1. \tag{1.4}$$

Thus, in $\mathbb{Z}_+^d$ with $d \ge 40$ there is a vertex from which one sees all words, and there is a strictly positive probability that one sees all words from the origin. Note that we do not claim that 10 and 40 are the best bounds on d for (1.3) and (1.4), respectively. Presumably, (1.3) holds even for some $d < 10$, but our proof is not sharp enough to decide this. A similar comment applies to the restrictions we put on d in other places.

Since it is more difficult to see a word on the oriented $\mathbb{Z}_+^d$ than on the undirected $\mathbb{Z}^d$, the following corollary is immediate.

COROLLARY 1. *If $\mathscr{G} = \mathbb{Z}^d$ with undirected edges, then* (1.3) *remains valid for $d \geq 10$ and* (1.4) *remains valid for $d \geq 40$.*

In the unoriented case we can say more about the set on which one sees all words: For large d this can be a rather narrow "tube." Specifically we shall indicate in Section 4 a proof of the following theorem.

THEOREM 2. *For $d \geq 132$ and $\mathscr{G} = \mathbb{Z}^d$ (unoriented),*

$$P\{\exists \textit{ path } (0, v_1, v_2, \ldots) \textit{ in } [-1,1]^2 \times \mathbb{Z}^{d-2}, \textit{ starting from the origin and with } \textstyle\sum_{\ell=3}^{d} v_i(\ell) \textit{ increasing in } i \textit{ and such that one sees all words from the origin in the tube } \{w\colon |w - v_i| \leq 2 \textit{ for some } i\}\} > 0. \tag{1.5}$$

Here and in the sequel $v(j)$ stands for the jth coordinate of v, $|v| = \sum_1^d v(\ell)$ and $f(i)$ is increasing means $f(i+1) \geq f(i)$.

We do not know whether (1.5) remains valid when $\mathscr{G}$ is $\mathbb{Z}_+^d$ with positively oriented edges and d large.

It also follows from (1.3) (and Remark 1 in Section 2) that for each k, there is a smallest d for which on $\mathscr{G} = \mathbb{Z}_+^d$ with positively oriented edges there exists a k-tuple $(v_1, v_2, \ldots, v_k)$ with

$$P\{S(v_1, \ldots, v_k) = \Xi\} > 0. \tag{1.6}$$

Denote this smallest d by $d(k)$. Then by (1.4), $d(k) \leq d(1) \leq 40$. We know little also about these $d(k)$, but we do prove in Corollary 3 in Section 6 that

$$d(1) \geq 4, \tag{1.7}$$

or equivalently, that on the positively oriented $\mathbb{Z}_+^3$ one does not see all words from one point. We prove this together with some similar results for the regular 3–tree. If $\mathscr{G}$ = oriented regular 3–tree, then one does see all words from two points, but not from a single point, that is,

$$P\{S(v) = \Xi \text{ for some } v\} = 0, \tag{1.8a}$$

but

$$P\{S(v_1, v_2) = \Xi \text{ for some pair } v_1, v_2\} = 1. \tag{1.8b}$$

We prove (1.8) in Section 6. In addition this section contains the construction of trees on which, for given k,

$$P\{S(v_1, \ldots, v_k) = \Xi \text{ for some } v_1, \ldots, v_k\} = 0, \tag{1.9a}$$

but

$$P\{S(v_1, \ldots, v_{k+1}) = \Xi \text{ for some } (k+1)\text{–tuple } v_1, \ldots, v_{k+1}\} = 1. \tag{1.9b}$$

Thus there are examples which separate the events in (1.9a) and (1.9b).

Even if one does not see all words, that is, if $P\{S_\infty = \Xi\} = 0$, it is possible that $S(v_1, \ldots, v_k)$ or S_∞ consists of "almost all" words in the following sense. Let $\mu = \prod_1^\infty \mu_i$ be the product measure on Ξ in which each factor μ_i on $\{0,1\}$ is given by

$$(1.10) \qquad \mu_i(\{0\}) = \mu_i(\{1\}) = \tfrac{1}{2}.$$

(Thus under μ, the ξ_i are i.i.d. symmetric binomial variables.) For $\mathscr{G} = \mathbb{Z}^d$ or $\mathbb{Z}_+^d$ or a locally finite tree, it is not hard to see (compare proof of Proposition 3) that for each ξ,

$$(1.11) \qquad \rho(\xi) := P\{\xi \text{ is seen from some } v\}$$

equals 0 or 1. We shall say that ξ *percolates* if $\rho(\xi) = 1$. We shall see in Section 2 that $\xi \to \rho(\xi)$ is measurable with respect to the standard σ–field $\mathscr{F}$ on Ξ generated by the cyclinder sets. We shall also see that for many graphs $\rho(\xi)$ does not depend on any finite number of coordinates $\xi_1, \ldots, \xi_n$. On such graphs ρ is a tail variable, and by Kolmogorov's zero–one law,

$$(1.12) \qquad \mu\{\xi\colon\ \rho(\xi) = 1\} \text{ equals 0 or 1.}$$

In the former (latter) case almost no word (respectively, almost all words) percolate. We note that, by Fubini's theorem applied to the indicator function of the set $\Lambda := \{(\xi, \omega)\colon\ \xi$ is seen for some v in the configuration $\omega\}$, we have

$$(1.13) \qquad \mu\{\xi\colon\ \rho(\xi) = 1\} = \begin{cases} 0 \\ 1 \end{cases} \text{according as } P\{\mu(S_\infty) = 1\} = \begin{cases} 0 \\ 1 \end{cases}.$$

[Note that Λ, and hence $\rho(\cdot)$ and S_∞, are measurable with respect to the appropriate σ–algebras; see Proposition 1.] Another way to formulate the preceding dichotomy is by picking ξ at random according to μ and then to see whether for this random ξ, $\rho(\xi)$ equals 0 or 1. If $\mu\{\xi\colon\ \rho(\xi) = 1\} = 0$, then "the random word does not percolate," while if $\mu\{\xi\colon\ \rho(\xi) = 1\} = 1$, then "the random word percolates (almost surely)." It seems difficult to decide which of these two cases prevails. The problem is most striking for graphs for which the critical probability for site percolation equals $\frac{1}{2}$. For instance, if $\mathscr{G}$ is the triangular lattice, then at $p = \frac{1}{2}$ ordinary percolation does not occur, that is, the word $\mathbb{1}$ is not seen in $\mathscr{G}$ [Kesten (1982), Section 3.3], but AB–percolation does occur [Wierman and Appel (1987)]. Another example is bond percolation on $\mathbb{Z}^2$. For bond percolation on $\mathbb{Z}^d$ one considers a family $X(e)$ of i.i.d. random variables, where e runs through the *edge*–set of $\mathbb{Z}^d$. Again we take $P\{X(e) = 0\} = P\{X(e) = 1\} = \frac{1}{2}$. The meaning of "$\xi$ is seen from v" now should be that there exists a self–avoiding path $(v_0 = v, v_1, \ldots)$ on $\mathbb{Z}^d$, starting at v, such that $X(\text{edge between } v_{i-1} \text{ and } v_i) = \xi_i$, $i \geq 1$. There is some ambiguity, though, in whether we should require the path $(v_0, v_1, \ldots)$ to be self–avoiding in the sense described above, or whether the path merely should not use the same edge more than once (the latter is a weaker restriction). Here we adopt the former requirement. Thus, when we discuss bond percolation on

$\mathbb{Z}^d$, we want words to be seen along paths which do not contain the same *vertex* more than once.

Both for site percolation on the triangular lattice and bond percolation on $\mathbb{Z}^2$ we know [cf. Kesten (1982), Section 3.3] that the word $\mathbb{1} = (1, 1, \dots)$ is w.p.1 not in S_∞, so that

$$P\{S_\infty = \Xi\} = 0. \tag{1.14}$$

This leads to the following problem.

OPEN PROBLEM 1. For site percolation on the triangular lattice, or bond percolation on $\mathbb{Z}^2$, does the random word percolate?

We note that there is some hope that the answer to this problem is affirmative, because at $p = \frac{1}{2}$ the word $\mathbb{1}$ is the least likely to percolate in the sense that for any $\xi \in \Xi$ and any v,

$$P\{\xi \text{ is seen from } v\} \geq P\{ \text{ the word } \mathbb{1} \text{ is seen from } v\}. \tag{1.15}$$

This follows by a simple modification of the proof of Proposition 3.1 in Wierman (1989), which proves (1.15) for $\xi = (1, 0, 1, 0, \dots)$. We note in passing that this implies that if $p_c(\mathbb{Z}^d, \text{site}) :=$ critical probability for site percolation on $\mathbb{Z}^d < \frac{1}{2}$, then

$$P\{\mu(S_\infty) = 1\} = 1. \tag{1.16}$$

By Campanino and Russo (1985) this therefore holds on $\mathbb{Z}^d$ with $d \geq 3$, but formally it does not imply that

$$P\{S_\infty = \Xi\} = 1. \tag{1.17}$$

As a special case, we have the following problem:

OPEN PROBLEM 2. Does (1.17) hold when $\mathscr{G} = \mathbb{Z}^3$?

On the other side, it is a priori not clear that there exists any graph on which (1.14) holds, but still the random word percolates, or

$$P\{\mu(S_\infty) = 1, \text{ but } S_\infty \neq \Xi\} = 1. \tag{1.18}$$

In Section 7 we construct a tree for which (1.18) holds.

This separates percolation of the random word from percolation of all words. On the other side one may also wish to separate percolation of the random word from standard percolation, that is, percolation of the word $\mathbb{1}$. In Section 8 we construct a graph $\mathscr{G}$ for which

$$P\{\mu(S_\infty) = 1, \text{ but } \mathbb{1} \notin S_\infty\} = 1. \tag{1.19}$$

Basically, Open Problem 1 asks whether (1.19) occurs for bond percolation on $\mathbb{Z}^2$ or site percolation on the triangular lattice.

Finally, we discuss in Section 5 what happens when $\mathscr{G}$ is taken to be the random family tree of a Bienaymé–Galton–Watson process [see Harris (1963), Section 6.2, or Jagers (1975), Section 1.2] for a description of such trees; they are rooted ordered trees, sometimes also called planted plane trees. Let the offspring distribution be given by

$$p_k = P\{\text{a given individual has } k \text{ children}\} \tag{1.20}$$

and have mean and generating function

$$m = \sum_{1}^{\infty} kp_k \quad \text{and} \quad f(z) = \sum_{0}^{\infty} p_k z^k, \tag{1.21}$$

respectively. The following result holds [when $X(v)$ and ξ_i take the values 0 or 1 again].

THEOREM 3. *Let $\mathscr{G}$ be the random family tree of a Bienaymé–Galton–Watson process. If $m > 2$, then $\mathscr{G}$ has with probability 1 the property*

$$P\{S(v_1, \dots, v_k) = \Xi \text{ for some } k \text{ and } v_1, \dots, v_k\} = 1. \tag{1.22}$$

We are not aware of any closely related articles, other than the ones cited above. Somewhat loosely related are Dekking (1991), Dekking and Pakes (1991), Evans (1992), Lyons (1992) and Menshikov and Zuyev (1992). Dekking and Dekking and Pakes give the necessary and sufficient condition for a random Bienaymé–Galton–Watson tree to contain a full binary or a full b–ary subtree. Evans investigated for a set $B \subset \Xi^{(b)} := \{0, 1, \dots, b-1\}^{\mathbb{N}}$ and $\mathscr{G}$ a regular c–ary tree when

$$P\{\text{some word in } B \text{ is seen}\} = P\{S_\infty \cap B \neq \varnothing\}$$

is zero or strictly positive. This is generalized by Lyons when $\mathscr{G}$ is a more general tree. Menshikov and Zuyev investigate so–called ρ–percolation on $\mathbb{Z}^d$ or on homogeneous trees. They consider, in the case when $X(v)$ and ξ_i again take values in $\{0, 1\}$, whether one sees any words whose frequency of ones is at least ρ. Thus, if

$$B(\rho) = \left\{\xi\colon \liminf \frac{1}{n} \sum_{1}^{n} \xi_i \geq \rho\right\},$$

then they discuss, for $\mathscr{G} = \mathbb{Z}^d$ or a homogeneous tree, for which values of p

$$P_p\{S_\infty \cap B(\rho) \neq \varnothing\} > 0.$$

We summarize the different percolation versions which we considered in Figure 1. The implications shown are all trivial except for the one marked $(*)$. This implication follows from (1.15) on $\mathscr{G} = \mathbb{Z}^d$, $\mathbb{Z}_+^d$ with positive orientation or a locally finite tree. If $P\{\mathbb{1}$ is seen from $v\} > 0$, then by (1.15), $P\{\xi$ is seen from $v\} > 0$ for all ξ. On the above graphs Proposition 3 shows that $\rho(\xi) = P\{\xi$ is seen somewhere on $\mathscr{G}\} = 1$ for all ξ. The letters in parentheses refer to the examples (listed below) which show that in general the arrow cannot be reversed:

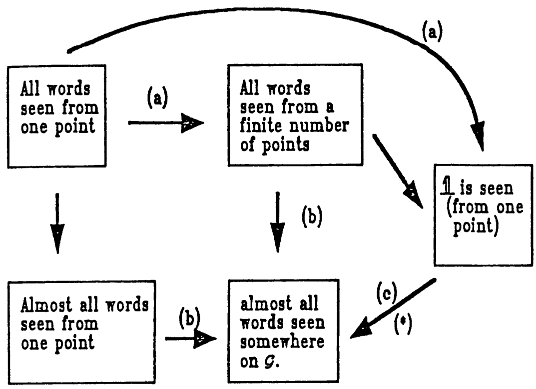

FIG. 1. (a) *The converse fails when $\mathscr{G}$ = regular 3–tree; see Section* 6. (b) *The converse fails for $\mathscr{G}$ equal to the trees in Section* 7 [*see* (7.5)–(7.7) *and Remark* 1]. (c) *The converse fails for the graph of Section* 8.

2. Measurability and zero–one laws. We first prove a simple measurability result about the set of pairs of words and configurations such that the former is seen in the latter (somewhere, or from a specific point). This will guarantee that (1.12) and (1.13) are meaningful and justify the application of Fubini's theorem to go from one to the other. After that we use Baire's category theorem to prove the measurability of the events $\{S(v) = \Xi\}$ and $\{S_\infty = \Xi\}$ in Ω.

Let $\mathscr{B}$ and $\mathscr{F}$ be the σ–fields in Ω and Ξ, respectively, generated by the cylinder sets. Define

$$\Lambda(v) = \{(\xi, \omega) \in \Xi \times \Omega\colon\ \xi \text{ is seen from } v \text{ in the configuration } \omega\}$$

and

$$\Lambda = \{(\xi, \omega) \in \Xi \times \Omega\colon\ \xi \text{ is seen somewhere on } \mathscr{G} \text{ in the configuration } \omega\}.$$

PROPOSITION 1. *Let $\mathscr{G}$ be locally finite. Then, for each v, $\Lambda(v) \in \mathscr{F} \times \mathscr{B}$. Also $\Lambda \in \mathscr{F} \times \mathscr{B}$.*

PROOF. Clearly

$$\Lambda = \bigcup_{v \in \mathscr{V}} \Lambda(v),$$

so that we only need to prove the first statement. Fix v and an $(\eta_1, \dots, \eta_n) \in \{0,1\}^n$. Define

$$\begin{aligned} B_n(\eta_1, \dots, \eta_n) &= B_n(\eta_1, \dots, \eta_n; v) \\ &= \{\omega \in \Omega\colon\ \exists \text{ path } (v, v_1, \dots, v_n) \text{ of } n+1 \\ &\qquad \text{vertices on } \mathscr{G}, \text{ which starts at } v \text{ and} \\ &\qquad \text{has } X(v_i) = \eta_i \text{ for } 1 \le i \le n\}. \end{aligned} \tag{2.1}$$

Also let

$$C_n(\eta_1,\ldots,\eta_n)=\{\xi\in\Xi:\ \xi_i=\eta_i,\ 1\le i\le n\} \tag{2.2}$$

["the cylinder with base $(\eta_1,\ldots,\eta_n)$"]. Then

$$\Lambda(v)=\bigcap_n\ \bigcup_{(\eta_1,\ldots,\eta_n)} C_n(\eta_1,\ldots,\eta_n)\times B_n(\eta_1,\ldots,\eta_n). \tag{2.3}$$

Clearly $C_n(\eta_1,\ldots,\eta_n)\in\mathscr{F}$. Also $B_n(\eta_1,\ldots,\eta_n)\in\mathscr{B}$, because there exist at most finitely many paths $(v,v_1,\ldots,v_n)$ on $\mathscr{G}$ which start at v. The required result is immediate from (2.3). □

We turn to the proof of Proposition 2 stated in the Introduction.

PROOF OF PROPOSITION 2. The fact that $\{S(v)=\Xi\}\in\mathscr{B}$ follows from

$$\{S(v)=\Xi\}=\bigcap_n\ \bigcap_{(\eta_1,\ldots,\eta_n)} B_n(\eta_1,\ldots,\eta_n)$$

in the notation of (2.1). A similar argument works for $\{S(v_1,\ldots,v_k)=\Xi\}$ and we therefore only need to prove that

$$\{S_\infty=\Xi\}\in\mathscr{B}. \tag{2.4}$$

To do this we apply the Baire category theorem [see Rudin (1987), Theorem 5.6]. For fixed ω, the set $S(v)=\{\xi:\ \xi$ is seen from $v\}$ is closed in Ξ in the product topology, which is also the topology induced by the metric

$$d(\xi',\xi'')=\sum_1^\infty 2^{-k}|\xi'_k-\xi''_k|. \tag{2.5}$$

Indeed, if $\xi^{(n)}\in S(v)$, $n=1,2,\ldots$, and $\xi^{(n)}\to\xi$ as $n\to\infty$, then $\xi_i^{(n)}=\xi_i$ for n sufficiently large. Moreover, because $\mathscr{G}$ is locally finite, there must for each k be a path $(v,v_1,v_2,\ldots,v_k)$ with distinct vertices such that $X(v_i)=\xi_i^{(n)}=\xi_i$, $1\le i\le k$, for infinitely many n. A standard selection argument then shows the existence of an infinite path $(v,v_1,v_2,\ldots)$ such that $X(v_i)=\xi_i$, $i\ge 1$, so that $\xi\in S(v)$. Thus $S(v)$ is closed, as claimed.

Now if

$$S_\infty=\bigcup_v S(v)=\Xi,$$

then, since Ξ with the metric (2.5) is complete, one of the $S(v)$ must contain an open set, and in particular some $C_n(\eta_1,\ldots,\eta_n)$. In the opposite direction, $C_n(\eta_1,\ldots,\eta_n)\subset S(v)$ shows that each word ξ which begins with $\xi_i=\eta_i$, $1\le i\le n$, is seen from v, no matter what $\xi_{n+1},\xi_{n+2},\ldots$ are. However, then every word is seen from some vertex. Thus

$$\{S_\infty=\Xi\}=\bigcup_v\ \bigcup_{n,\eta_1,\ldots,\eta_n}\{C_n(\eta_1,\ldots,\eta_n)\subset S(v)\}; \tag{2.6}$$

(2.4) now follows easily. □

REMARK 1. Let $\mathscr{G}$ be locally finite. If all words are seen somewhere on $\mathscr{G}$, then from some point v a whole cylinder $C_n(\eta_1, \dots, \eta_n)$ is seen. [This is the content of (2.6).] In turn this implies that all words are seen from the finite set

$$\{w\colon \exists \text{ a path } (v, v_1, \dots, v_n = w) \text{ of } (n+1) \text{ vertices from } v \text{ to } w\}$$

[since all continuations of $(\eta_1, \dots, \eta_n)$ have to be seen from this set]. Thus

$$\begin{aligned}\{S_\infty = \Xi\} &= \{\text{all words are seen somewhere on } \mathscr{G}\}\\ &\subset \{\text{for some finite set } \{v_1, \dots, v_k\}\\ &\qquad \text{all words are seen from } v_1, \dots, v_k\}\\ &= \bigcup_{k, v_1, \dots, v_k} \{S(v_1, \dots, v_k) = \Xi\}.\end{aligned}$$

The analogous statement with "all words" replaced by "almost all words" is false. That is, it is possible that almost all words are seen on $\mathscr{G}$, but from no finite set does one see almost all words. An example of this is provided in Theorem 5 [see (7.6) and (7.7)].

We end this section with some simple zero–one laws for the function ρ of (1.11).

PROPOSITION 3. *If $\mathscr{G}$ is locally finite, then $\rho(\cdot)$ is $\mathscr{F}$–measurable. If $\mathscr{G} = \mathbb{Z}^d$ (unoriented) or $\mathbb{Z}_+^d$ (with positive orientation) or a locally finite tree, then $\rho(\cdot)$ only takes the values* 0 *and* 1 *and* (1.12) *holds.*

PROOF. Note that $\rho(\xi)$ is the probability of the ξ–section of Λ, that is, of the set $\{\omega\colon (\xi, \omega) \in \Lambda\}$. Since $\Lambda \in \mathscr{F} \times \mathscr{B}$, by Proposition 2, the $\mathscr{F}$–measurability of $\rho(\cdot)$ is standard [see Rudin (1987), Theorem 8.6]. In the same way we see that $\xi \to P\{\xi \text{ is seen from } v\}$ is $\mathscr{F}$–measurable for each fixed v. If this probability vanishes for all v, then clearly $\rho(\xi) = 0$. We next want to argue that if

$$P\{\xi \text{ is seen from } v\} > 0 \quad \text{for some } v, \tag{2.7}$$

then $\rho(\xi) = 1$. For $\mathscr{G} = \mathbb{Z}^d$ or the oriented $\mathbb{Z}_+^d$ or $\mathscr{G}$ a homogeneous tree, in which all vertices play the same role, we can argue as in Harris (1960) or in the proof of Theorem 4 below. For such $\mathscr{G}$ the probability in (2.7) is the same for all v and if it is strictly positive, then the ergodic theorem shows that ξ is seen from infinitely many v w.p.1, so that $\rho(\xi) = 1$.

For an inhomogeneous tree the proof only works for $p = \frac{1}{2}$. It does, however, apply equally well whether we consider an unoriented tree or a rooted tree oriented away from the root. Assume that $\mathscr{G}$ is a tree and that the probability in (2.7) is greater than 0 for $v = v_0$. Let $\mathscr{V}_n$ be the collection of vertices which can be reached from v_0 by a path containing at most $(n+1)$ vertices ($\mathscr{V}_0 = \{v_0\}$). If $\mathscr{B}_n$ is the σ–field generated by $\{X(v)\colon v \in \mathscr{V}_n\}$, then by the martingale convergence theorem [see Corollary 5.22 in Breiman (1968)]

$$P\{\xi \text{ is seen from } v_0 \mid \mathscr{B}_n\} \to 1 \quad \text{a.e. on the event } \{\xi \text{ is seen from } v_0\}.$$

In particular, for any $\varepsilon > 0$ and n large enough, there must exist a choice for $\eta(v)$, $v \in \mathscr{V}_n$, such that

$$P\{\xi \text{ is seen from } v_0 \mid X(v) = \eta(v),\ v \in \mathscr{B}_n\} \geq 1 - \varepsilon. \tag{2.8}$$

However, because $\mathscr{G}$ is a tree, the probability on the left is at most

$$P\{\exists \text{ path } (v, v_1, \dots) \text{ with } v \in \mathscr{V}_n/\mathscr{V}_{n-1} \text{ and } v_i \notin \mathscr{V}_n \text{ for } i \geq 1,\ X(v_i) = \xi_{n+i},\ i \geq 1\}. \tag{2.9}$$

Moreover, because $\mathscr{G}$ is a tree *and* $p = \frac{1}{2}$, the last probability is independent of ξ. To see this define

$$Y(w) = \begin{cases} X(w), & \text{if } w \in \mathscr{V}_k \setminus \mathscr{V}_{k-1} \text{ and } \xi_k = 1, \\ 1 - X(w), & \text{if } w \in \mathscr{V}_k \setminus \mathscr{V}_{k-1} \text{ and } \xi_k = 0. \end{cases}$$

Then the families $\{X(w)\colon w \in \mathscr{V}\}$ and $\{Y(w)\colon w \in \mathscr{V}\}$ have the same distribution. Moreover, since any path $(v, v_1, \dots)$ with $v \in \mathscr{V}_n/\mathscr{V}_{n-1}$ and $v_i \notin \mathscr{V}_n$ for $i \geq 1$ must have $v_i \in \mathscr{V}_{n+i}/\mathscr{V}_{n+i-1}$, $i \geq 1$, it follows that

$$\{X(v_i) = \xi_{n+i},\ i \geq 1\} = \{Y(v_i) = 1,\ i \geq 1\}.$$

Thus the probability in (2.9) equals, independently of ξ,

$$P\{\exists \text{ path } (v, v_1, \dots) \text{ with } v \in \mathscr{V}_n/\mathscr{V}_{n-1} \text{ and } v_i \notin \mathscr{V}_n \text{ for } i \geq 1,\ X(v_i) = 1,\ i \geq 1\}.$$

Since the probability in (2.9) is at least as large as the left-hand side of (2.8), it follows that

$$\begin{aligned} \rho(\xi) &= P\{\xi \text{ is seen from some } v \text{ in } \mathscr{G}\} \\ &\geq P\{\exists \text{ path } (v, v_1, \dots) \text{ with } v \in \mathscr{V}_n/\mathscr{V}_{n-1} \text{ and} \\ &\qquad v_i \notin \mathscr{V}_n, X(v_i) = 1,\ i \geq 1\} \\ &= \text{probability in (2.9)} \geq 1 - \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this establishes our claim that $\rho(\xi) = 1$. This shows in particular that for our $\mathscr{G}$'s, $\rho(\xi)$ can only take the values 0 or 1.

We now return to (2.7). Assume that (2.7) holds for a given ξ. Then for any ξ' which differs from ξ in finitely many coordinates only, also $P\{\xi' \text{ is seen from } v\} > 0$ and by the preceding argument $\rho(\xi') = 1$. Thus $\rho(\xi) = 1$ implies $\rho(\xi') = 1$ whenever ξ and ξ' differ in finitely many coordinates only. Similarly $\rho(\xi) = 0$ implies $\rho(\xi') = 0$ for such ξ and ξ'. Thus $\rho(\xi)$ is independent of $\xi_1, \dots, \xi_n$. (1.12) now follows from the Kolmogorov zero–one law [Breiman (1968), Theorem 3.12]. □

3. Oriented percolation on $\mathbb{Z}_+^d$. This section is devoted to the proof of Theorem 1. Much of this section follows suggestions of R. Durrett and G. Grimmett, for which we are grateful. As stated before we restrict ourselves to $p = \frac{1}{2}$, but most of this section can be done for general $p \in (0,1)$ with only minor modifications. Throughout this section we take $\mathscr{G} = \mathbb{Z}_+^d$ with all edges oriented in the positive direction. The random variables $X(v)$ and ξ_i take values in $\{0,1\}$ and $p = \frac{1}{2}$. In this case,

$$P\{\xi \text{ is seen from } v\} \tag{3.1}$$

is trivially independent of v.

Define the *lifetime* of a word $\xi = (\xi_1, \xi_2, \dots)$ as

$$\tau(\xi) = \max\{k\colon \exists \text{ oriented path } (\mathbf{0}, v_1, \dots, v_k) \text{ with } X(v_i) = \xi_i\}.$$

Thus $\tau(\xi)$ is the length of the largest initial segment of ξ that is seen from the origin and $\{\tau(\xi) = \infty\}$ is the event that ξ itself is seen from the origin. The same argument as used to prove the independence of (2.9) from ξ shows that the distribution of $\tau(\xi)$ is independent of ξ. We merely have to redefine $\mathscr{V}_n$ as

$$\mathscr{V}_n = \left\{v = (v(1), \dots, v(d)) \in \mathbb{Z}_+^d\colon \sum_1^d v(i) = n\right\}$$

and to observe that for any oriented path $(v_1, v_2, \dots)$, $v_j \in \mathscr{V}_n$ implies $v_{j+i} \in \mathscr{V}_{n+i}$. Similarly, the probability in (3.1) is the same for all ξ.

The following lemma contains the principal estimate of this section.

LEMMA 1. *For each ξ,*

$$P\{\tau(\xi) = m\} \le \frac{z^{2m+4}}{1-z} \quad \text{with } z = \mu 2^{-1/2\lfloor d/2 \rfloor + 1/4}, \tag{3.2}$$

where μ is the so-called connective constant for $\mathbb{Z}^2$ [see Hammersley (1961) *or Madras and Slade* (1993), *Section* 1.2].

PROOF. As shown above we may take $\xi = \mathbb{1}$. Abbreviate $\tau(\mathbb{1})$ to τ and for $k + \ell \ge 1$, $k \ge 0$, $\ell \ge 0$ define $\kappa(k, \ell)$ as the indicator function of the event

$$\begin{aligned} \Big\{ &\exists \text{ path } (\mathbf{0}, v_1, \dots, v_n) \text{ on } \mathbb{Z}_+^d \text{ from } \mathbf{0} \text{ to some } v_n = \\ &(v_n(1), \dots, v_n(d)) \text{ with } X(v_i) = 1,\ 1 \le i \le n, \text{ and} \\ &\sum_{1 \le j \le d/2} v_n(j) = k,\ \sum_{d/2 < j \le d} v_n(j) = \ell \Big\}. \end{aligned} \tag{3.3}$$

If $(v_0 = \mathbf{0},\ v_1, \dots, v_n)$ is an oriented path on $\mathbb{Z}_+^d$, then $v_{i+1} = v_i + e_j$ for some $1 \le j \le d$ (where e_j is the jth coordinate vector). In particular, $\sum_{\ell=1}^d v_i(\ell) = i$ and the event in (3.3) forces $n = k + \ell$. Note also that if $\tau = m$, then there exists some path $(\mathbf{0}, v_1, \dots, v_m)$ on $\mathbb{Z}_+^d$ with $X(v_i) = 1$ for $1 \le i \le m$, but no such path $(\mathbf{0}, v_1, \dots, v_{m+1})$ exists. If, for $v \in \mathbb{Z}_+^d$, we now define

$$\pi(v) = \left(\sum_{1 \le j \le d/2} v(j), \sum_{d/2 < j \le d} v(j) \right) \in \mathbb{Z}_+^2,$$

then $(\mathbf{0}, \pi(v_1), \dots, \pi(v_m))$ is an oriented path on $\mathbb{Z}_+^2$ with $\kappa(\pi(v_i)) = 1$, $0 \le i \le m$, but no such path of length $(m+1)$ exists on $\{\tau = m\}$. This will allow us to use the standard contour argument for oriented two–dimensional site percolation to estimate (3.2).

Let $\mathscr{C}$ be the "cluster of the origin" for the κ–process; that is, $(k, \ell) \in \mathscr{C}$ if and only if there exists an oriented path $(\mathbf{0}, (k_1, \ell_1), \dots, (k_m, \ell_m) = (k, \ell))$ from $\mathbf{0}$ to (k, ℓ) on $\mathbb{Z}_+^2$ with $\kappa((k_i, \ell_i)) = 1$, $1 \le i \le m$. From the preceding observations we see that actually

$$\mathscr{C} = \{(k, \ell) \in \mathbb{Z}_+^2 \colon \kappa(k, \ell) = 1\} \tag{3.4}$$

and

$$\{\tau = m\} = \{\mathscr{C} \text{ contains some point } (k, \ell) \text{ with } k + \ell = m, \text{ but no points } (k, \ell) \text{ with } k + \ell = m + 1\}. \tag{3.5}$$

We can view $\mathscr{C}$ as a connected set of $\mathbb{Z}^2$, ignoring the orientation. Then it is well known [see Kesten (1982), Corollary 2.2, or Durrett (1988), Section 5a] that if $\mathscr{C}$ is finite, then it is "surrounded by a dual circuit." Specifically this means that there exists a path $\pi^* = (w_1^*, w_2^*, \dots, w_p^*)$ on the unoriented graph $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$ with the following properties:

(3.6) $w_1^*, \dots, w_p^*$ are distinct vertices of $\mathbb{Z}^2 + \left(\frac{1}{2}, \frac{1}{2}\right)$ and w_p^* is adjacent to w_1^* (thus π^* may be called a self–avoiding circuit);

(3.7) if f_i^* is the edge between w_i^* and w_{i+1}^*, for $1 \le i < p$, and between w_p^* and w_1^* for $i = p$, then each f_i^* bisects an edge f_i of $\mathbb{Z}^2, 1 \le i \le p$, and if the endpoints of f_i are a_i and b_i, then for each $1 \le i \le p$ one vertex of the pair $\{a_i, b_i\}$ belongs to $\mathscr{C}$, and the other vertex of the pair $\{a_i, b_i\}$ does not belong to $\mathscr{C}$;

(3.8) the circuit $(w_1^*, \dots, w_p^*)$ separates $\mathscr{C}$ from ∞; that is, any path on $\mathbb{Z}^2$ from some vertex in $\mathscr{C}$ to ∞ must cross one of the edges f_i^*, $1 \le i \le p$.

On the event (3.5), $\mathscr{C}$ contains $\mathbf{0}$ as well as some point (k, ℓ) with $k + \ell = m$. This, together with the property (3.8), shows that the circuit π^* must satisfy

$$p \ge 2(k + 1 + \ell + 1) = 2k + 2\ell + 4 = 2m + 4. \tag{3.9}$$

Let us denote the endpoints in $\mathscr{C}$ (outside $\mathscr{C}$) of the edge of f_i in (3.7) by a_i (b_i, respectively). By (3.4) we must have $\kappa(b_i) = 0$. We are especially interested in those (a_i, b_i) for which

$$(3.10) \qquad b_i = a_i + (1,0) \quad \text{or} \quad b_i = a_i + (0,1),$$

for reasons which will become clear [see (3.12)]. We shall need that

(3.11) the number of $i \le p$ for which (3.10) holds equals $p/2 \ge m+2$.

Basically the argument for (3.11) is given in Liggett [(1985), proof of Theorem 3.19] or in Durrett [(1988), proof of (ii), page 79]. Let c^* be the curve obtained by successively traversing the edges $f_1^*, f_2^*, \dots, f_p^*$. Then c^* is a Jordan curve, because the w_i^* are distinct. Consider a horizontal line $\{(x, r)\colon x \in \mathbb{R}\}$, with r an integer, which intersects c^*. As one moves along this line from $x = -\infty$ to $x = +\infty$, one goes from the exterior of c^* to the exterior of c^*. There are therefore as many intervals from (s, r) to $(s+1, r)$, $s \in \mathbb{Z}$, on this line with (s, r) in the exterior of c^* and $(s+1, r)$ in the interior of c^*, as there are such intervals with (s, r) in the interior and $(s+1, r)$ in the exterior. If the interval from (s, r) to $(s+1, r)$ goes from the interior of c^* to its exterior, then the segment (s, r) to $(s+1, r)$ is one of the edges from a_i to $b_i = a_i + (1,0)$, which satisfies (3.10). The intervals which go from the exterior to the interior of c^* as one goes from (s, r) to $(s+1, r)$ are edges f_i with $b_i = a_i - (1,0)$, and do not satisfy (3.10). By varying r we find all the horizontal edges f_i. Therefore, there are as many horizontal edges f_i for which (3.10) holds as there are horizontal edges with $b_i = a_i - (1,0)$. Similarly for vertical edges f_i. Since the total number of edges f_i equals p, (3.11) follows.

Next let $c^* = (w_1^*, \dots, w_p^*)$ be a fixed self–avoiding dual circuit. We wish to estimate the probability that c^* is the outer boundary of $\mathscr{C}$ as described by (3.6)–(3.8). The circuit c^* determines the dual edges f_i^* as well as the edges f_i which intersect f_i^*. The edge f_i has endpoint a_i and b_i with $a_i \in$ interior of c^*, $b_i \in$ exterior of c^*. Thus c^* also determines the collection of pairs $(a_{i_q}, b_{i_q}) \in \mathbb{Z}^2 \times \mathbb{Z}^2$, $1 \le q \le p/2$, which satisfy (3.10). Thus, we can write

$$P\{\mathscr{C} \text{ has } c^* \text{ as outer boundary}\}$$
$$= P\{\mathscr{C} \text{ has } c^* \text{ as outer boundary}, \kappa(a_{i_q}) = 1, \kappa(b_{i_q}) = 0 \text{ for } 1 \le q \le p/2\}.$$

We shall now prove that

$$(3.12) \qquad P\{\mathscr{C} \text{ has } c^* \text{ as outer boundary}\} \le 2^{-p/2\lfloor d/2\rfloor + p/4}.$$

Let us first fix the values of $X(v)$ for v in the collection $\mathscr{D}$, say, of all v with $\pi(v) \in$ interior of c^*. Note that $\kappa(a_{i_q}) = 1$ is possible only if there exists an oriented path $(\mathbf{0}, v_1, \dots, v_n)$ with $X(v_i) = 1$, $1 \le i \le n$, and $\pi(v_n) = a_{i_q}$. In this case $\kappa(\pi(v_i)) = 1$, $\pi(v_i) \in \mathscr{C}$ and hence $\pi(v_i)$ must lie in interior of c^*. There may be several choices for the path $(\mathbf{0}, v_1, \dots, v_n)$ but we arbitrarily fix one of these choices and denote its endpoint by u_q [so that $\pi(u_q) = a_{i_q}$,

$X(u_q) = 1]$. The event $A := \{\kappa(a_{i_q}) = 1,\ 1 \le q \le p/2\}$ depends therefore only on $\{X(v)\colon\ v \in \mathscr{D}\}$. We now estimate

$$P\{\kappa(b_{i_q}) = 0,\ 1 \le q \le b/2 \mid X(v),\ v \in \mathscr{D}\} \tag{3.13}$$

on the event A. We define the set of vertices $\mathscr{E}_q$ as the set $\{u_q + e_j\colon\ 1 \le j \le d/2\}$ in case $b_{i_q} = a_{i_q} + (1,0)$, and $\mathscr{E}_q$ as $\{u_q + e_j\colon\ d/2 < j \le d\}$ in case $b_{i_q} = a_{i_q} + (0,1)$. Now on A, $\kappa(b_{i_q}) = 0$ can occur only if

$$X(u_q + e_j) = 0 \quad \text{for } u_q + e_j \in \mathscr{E}_q. \tag{3.14}$$

Indeed, given the path $(\mathbf{0}, v_1, \dots, v_n = u_q)$ from $\mathbf{0}$ to u_q with $X(v_i) = 1$, $1 \le i \le n$, if also $X(u_q + e_j) = 1$, then we can extend the path with the vertex $u_q + e_j$ and $\kappa(\pi(u_q + e_j)) = 1$. This cannot be the case for any $u_q + e_j \in \mathscr{E}_q$, because $\pi(u_q + e_j) = b_{i_q}$ for $u_q + e_j \in \mathscr{E}_q$. Therefore $\kappa(b_{i_q}) = 0$ forces (3.14). In particular, on A the conditional probability in (3.13) is at most

$$P\{X(u_q + e_j) = 0 \text{ for } u_q + e_j \in \mathscr{E}_q,\ 1 \le q \le p/2\} = 2^{-\nu},$$

where

$$\nu = \text{cardinality of } \bigcup_1^{p/2} \mathscr{E}_q.$$

Statement (3.12) will follow once we show that on A,

$$\nu \ge \frac{p}{2}\left\lfloor \frac{d}{2} \right\rfloor - \frac{p}{4}. \tag{3.15}$$

To obtain (3.15) we observe first that by definition the cardinality of $\mathscr{E}_q$ is at least $\lfloor d/2 \rfloor$. The $\mathscr{E}_q$ are not necessarily disjoint. Clearly $\mathscr{E}_q$ and $\mathscr{E}_r$ are disjoint if $b_{i_q} \ne b_{i_r}$, because $\pi(v) = b_{i_q}$ for $v \in \mathscr{E}_q$, $\pi(v) = b_{i_r}$ for $v \in \mathscr{E}_r$. Our second observation is that for given q, $b_{i_q} = b_{i_r}$ can occur for at most one $r \ne q$. Indeed, if $b_{i_q} = b_{i_r} = b$, then either $a_{i_q} = b - (1,0)$, $a_{i_r} = b - (0,1)$ or $a_{i_q} = b - (0,1)$, $a_{i_r} = b - (1,0)$. There are therefore only two possibilities for a_{i_q} and a_{i_r} if $b_{i_q} = b_{i_r}$, and $b_{i_r} = b_{i_q}$ can indeed occur for only one $r \ne q$. The number of pairs (q,r) with $q \ne r$, $b_{i_q} = b_{i_r}$ is at most $p/4$, since q takes only $p/2$ values. Finally consider such a pair (q,r) with $b_{i_q} = b_{i_r}$. Then $\mathscr{E}_q \cap \mathscr{E}_r$ contains at most one vertex, since the vertices in $\mathscr{E}_q \cap \mathscr{E}_r$ have to be of the form $u_q + e_{j'}$ and of the form $u_r + e_{j''}$. This forces $e_{j'} - e_{j''} = u_r - u_q$ which determines j' and j'' uniquely (once u_q and u_r have been chosen). These observations show that

$$\begin{aligned}\text{cardinality of } \bigcup_{q=1}^{p/2} \mathscr{E}_q &\ge \frac{p}{2}\left\lfloor \frac{d}{2} \right\rfloor - (\text{number of pairs } (q,r) \text{ with } b_{i_q} = b_{i_r}) \\ &\ge \frac{p}{2}\left\lfloor \frac{d}{2} \right\rfloor - \frac{p}{4}.\end{aligned}$$

This proves (3.15), and hence (3.12).

Now that (3.12) is proven, the lemma follows easily. Note that any circuit $c^* = (w_1^*, \dots, w_p^*)$ which satisfies (3.6)–(3.8) must have the edge from $(-\frac{1}{2}, -\frac{1}{2})$

to $(\frac{1}{2}, -\frac{1}{2})$ as one of its edges f_i^*. This is so, because $\mathbf{0} \in \mathscr{C}$, but we can connect $\mathbf{0}$ to ∞ via the negative y–axis, and none of the edges from $(0, -k)$ to $(0, -k-1)$ for $k > 0$ goes from a point in $\mathscr{C}$ to a point outside $\mathscr{C}$, but the edge from $(0, 0)$ to $(0, -1)$ does. Hence this latter edge crosses one of the f_i^*. But then $(-\frac{1}{2}, -\frac{1}{2})$ must be one of the w_i^*, and without loss of generality we may number the w_i^* such that $w_1^* = (-\frac{1}{2}, -\frac{1}{2})$ and $w_2^* = (\frac{1}{2}, -\frac{1}{2})$. Moreover, all c^* which satisfy (3.6)–(3.8) must lie in $[-\frac{1}{2}, \infty) \times [-\frac{1}{2}, \infty)$, because $\mathscr{C}$ itself lies in $[0, \infty) \times [0, \infty)$. Therefore, for given p, the number of possible choices for p is at most equal to the number of self–avoiding closed polygons on $(\mathbb{Z})^2 + (\frac{1}{2}, \frac{1}{2})$ with p vertices, whose first two vertices are $(-\frac{1}{2}, -\frac{1}{2})$ and $(\frac{1}{2}, -\frac{1}{2})$ and whose bottom left vertex is $(-\frac{1}{2}, -\frac{1}{2})$. This number is at most μ^p, where μ is the connective constant for $\mathbb{Z}^2$ [see Hammersley (1961), equation (7), or Madras and Slade (1993), Theorem 3.2.3]. Thus, for given p the number of choices for c^* is at most μ^p. By (3.9), p must be at least $2m+4$ on the event $\{\tau = m\}$. Therefore, by virtue of (3.12),

$$P\{\tau = m\} \le \sum_{p=2m+4}^{\infty} \mu^p 2^{-p/2\lfloor d/2\rfloor + p/4} = \frac{z^{2m+4}}{1-z} \quad \text{for } z = \mu 2^{-1/2\lfloor d/2\rfloor + 1/4}. \tag{3.16}$$

PROOF OF (1.3). It is known that $\mu \le 2.7$ [see Madras and Slade (1993), page 12]. Therefore, for $d \ge 10$, the z of (3.16) satisfies

$$z = \mu 2^{-1/2\lfloor d/2\rfloor + 1/4} \le (2.7)2^{-9/4} \le 0.6$$

and the right-hand side of (3.16) is at most

$$\frac{5}{2}[(2.7)2^{-1/2\lfloor d/2\rfloor + 1/4}]^{2m+4} \le \frac{5}{2}[(2.7)2^{-9/4}]^{2m+4}. \tag{3.17}$$

Now, for any word ξ, let

$$\hat{\xi}^{(m)} = (\xi_1, \dots, \xi_{m+1}, 1, 1, \dots).$$

Then for any $k \le m$,

$$\{\tau(\xi) = k\} = \{\tau(\hat{\xi}^{(m)}) = k\}, \tag{3.18}$$

because both events occur if and only if there exists an oriented path $(\mathbf{0}, v_1, \dots, v_k)$ with $X(v_i) = \xi_i$, but no such path $(\mathbf{0}, v_1, \dots, v_{k+1})$ exists. Now the number of choices for $\hat{\xi}^{(m)}$ is at most 2^{m+1} and hence, by Lemma 1,

$$\sum_{m=0}^{\infty} P\{\tau(\hat{\xi}^{(m)}) = m \text{ for some } \hat{\xi}^{(m)}\} \le \sum_{m=0}^{\infty} 2^{m+1} C_2 [(2.7)2^{-9/4}]^{2m} < \infty \tag{3.19}$$

for $d \ge 10$ and a suitable constant C_2. Therefore, w.p.1, for only finitely many m does there exist a $\hat{\xi}^{(m)}$ with $\tau(\hat{\xi}^{(m)}) = m$. Then (3.18) shows that also, w.p.1,

$$\{\tau(\xi) = m \text{ for some } \xi\} \tag{3.20}$$

happens only finitely often, say only for $m \le N$ (with N random). But then take any path $(\mathbf{0}, v_1, \dots, v_{N+1})$ on $\mathbb{Z}_+^d$ and assume $X(v_i) = \eta_i \in \{0,1\}$ for $1 \le i \le N+1$. Then, for any choice of $\eta_{N+2}, \eta_{N+3}, \dots$ and with $\eta = (\eta_1, \eta_2, \dots)$, we have $\tau(\eta) = \infty$, since $\tau(\eta) \ge N+1$ by choice of $\eta_1, \dots, \eta_{N+1}$, while (3.20) with $m \ge N+1$ fails. Thus η is seen from 0 and a fortiori $(\eta_{N+2}, \eta_{N+3}, \dots)$ is seen from some v with $\sum_1^d v(i) = N+1$ for all choices of $(\eta_{N+2}, \eta_{N+3}, \dots)$. This means that $S_\infty = \Xi$. □

PROOF OF (1.4). Our first task is to make the estimate (3.19) somewhat more precise. In particular we want to choose M such that

$$\sum_{m=M}^{\infty} P\{\tau(\hat{\xi}^{(m)}) = m \text{ for some } \hat{\xi}^{(m)}\} \le \sum_{m=M}^{\infty} 2^{m+1}\frac{5}{2}[(2.7)2^{-9/4}]^{2m+4} < 1. \tag{3.21}$$

Some simple calculations show that $M = 2$ suffices for (3.21). Define

$$|v| = \sum_{i=1}^{d} v(i).$$

The argument following (3.19) now shows that for $d = 10$,

$$P\{\text{every } \xi \text{ is seen from some } v \text{ with } |v| = 2\} > 0. \tag{3.22}$$

Now take $D = 2^2 \times 10$ and define for $0 \le j \le 3$, the subsets of $\mathbb{Z}_+^D$:

$$B_j = \{v = (v(1), \dots, v(D)) \in \mathbb{Z}_+^D \colon v(i) = 0 \text{ for all } i \le 10j \text{ and all } i > 10(j+1)\}.$$

Then each B_j is a copy of $\mathbb{Z}_+^{10}$ and by (3.22) the event

$$F_j = \{\text{all words } \xi \text{ are seen in } B_j \text{ from some vertex } v \in B_j \text{ with } |v| = 2\}$$

has strictly positive probability. Since $B_{j_1} \cap B_{j_2} = \{\mathbf{0}\}$ for $j_1 \ne j_2$, the events F_j are independent and

$$P\{F_j \text{ occurs for all } 0 \le j \le 3\} > 0. \tag{3.23}$$

Next let G_j be the event

$$G_j = \{\text{for all } v \text{ in } B_j \text{ with } |v| = k, \text{ one has } X(v) = \eta_j(k),\ k = 1, 2\},$$

where $\eta_j = (\eta_j(1), \eta_j(2))$ runs through the four elements of $\{0,1\}^2$ (in some order) as j varies from 0 to 3. Then also

$$P\{G_j \text{ occurs for all } 0 \le j \le 3\} > 0, \tag{3.24}$$

and since G_j and F_j depend only on vertices v with $|v| \le 2$ and $|v| > 2$, respectively, the events in (3.23) and (3.24) are independent. It follows that

$$P\{F_j \cap G_j \text{ occurs for all } 0 \le j \le 3\} > 0. \tag{3.25}$$

Now assume $F_j \cap G_j$ occurs for all j and let $\xi = (\xi_1, \xi_2, \dots)$ be some word. Then (ξ_1, ξ_2) equals some η_j, while $(\xi_3, \xi_4, \dots)$ is seen from some $\hat{v} \in B_j$ with $|\hat{v}| = 2$. Thus, there exists an oriented path $(\mathbf{0}, v_1, v_2 = \hat{v}, v_3, \dots)$ in B_j, from the origin and passing through $\hat{v}$, such that $X(v_k) = \eta_j(k) = \xi_k$, $k = 1, 2$, and $X(v_{2+\ell}) = \xi_{2+\ell}$, $\ell \geq 1$. Thus ξ is seen from the origin. This holds for all ξ, so that $S(\mathbf{0}) = \Xi$. Hence (3.25) implies $P\{S(\mathbf{0}) = \Xi\} > 0$ on $\mathbb{Z}_+^D$. As in Harris (1960), the ergodic theorem now shows that (1.4) holds. □

4. Percolation of all words on tubes in $\mathbb{Z}^d$. In this section we outline an alternative argument to that of the preceding section, which shows that in the unoriented case all words may appear on a small set, and that this already happens for $d \geq 132$.

Let $H \subset \mathbb{Z}^2$ be the set consisting of the seven vertices $(0,0)$, $(\pm 1, 0)$, $(0, \pm 1)$, $\pm(1,1)$ (see Figure 2). The basis of our argument is the following deterministic lemma, which gives a deterministic configuration on the "chimney" $H \times \mathbb{Z}_+ \subset \mathbb{Z}^3$ in which all words are seen from the origin.

LEMMA 2. *Assume $X^*(v)$ takes the following values on $H \times \mathbb{Z}_+$:*

$$X^*(v) = 0 \quad \text{if } (v(1), v(2)) = (0,0), \tag{4.1}$$

$$X^*(v) = 1 \quad \text{if } (v(1), v(2)) = \pm(0,1) \text{ or } \pm(1,1), \tag{4.2}$$

$$X^*(v) = \begin{cases} 1, & \text{if } (v(1), v(2)) = (1,0) \text{ and } v(3) \text{ is even} \\ & \text{or } (v(1), v(2)) = (-1,0) \text{ and } v(3) \text{ is odd,} \\ 0, & \text{if } (v(1), v(2)) = (1,0) \text{ and } v(3) \text{ is odd} \\ & \text{or } (v(1), v(2)) = (-1,0) \text{ and } v(3) \text{ is even.} \end{cases} \tag{4.3}$$

Then for each word ξ there is a path $(\mathbf{0}, v_1, v_2, \dots) \in H \times \mathbb{Z}_+$ such that $v_i(3) \geq v_{i-1}(3)$ and $X^*(v_i) = \xi_i$ for $i \geq 1$. In particular, each word is seen from the origin in such a configuration.

We do not give the proof of this lemma. A somewhat tedious induction on n can be used to show for every $(\xi_1, \dots, \xi_n)$ there is a path $(\mathbf{0}, v_1, \dots, v_n) \subset H \times \mathbb{Z}_+$ with $X^*(v_i) = \xi_i$ and $v_i(3) \geq v_{i-1}(3)$, $1 \leq i \leq n$.

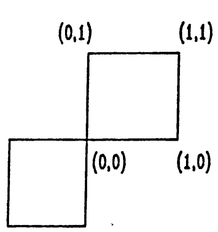

FIG. 2. *The set H in $\mathbb{Z}^2$.*

THEOREM 2. *If d is such that*

$$p_c(d-2,\ \textit{site, oriented}) < 2^{-7}, \tag{4.4}$$

then

$$P\left\{\exists \text{ a path } (v_0=\mathbf{0},v_1,v_2,\dots) \text{ in } H\times\mathbb{Z}^{d-2} \text{ such that } \sum_{\ell=3}^{d} v_i(\ell) \text{ is increasing in } i \text{ and such that every } \xi \text{ is seen in the tube } \{w\colon |w-v_i|\le 2 \text{ for some } i\}\right\} > 0. \tag{4.5}$$

Therefore,

$$P\{S(\mathbf{0})=\Xi\}>0 \tag{4.6}$$

and

$$P\{S_\infty=\Xi\}=1. \tag{4.7}$$

[*Here $p_c(d, site, oriented)$ is the critical probability for oriented site percolation on $\mathbb{Z}^d_+$.*] *In particular,* (4.5)–(4.7) *hold for $d\ge 132$.*

PROOF. We view $H\times\mathbb{Z}^{d-2}$ as a subset of $\mathbb{Z}^d$ and for any configuration ω we color the vertices of $\mathbb{Z}^{d-2}$ white or black. The vertex $w=(w(1),\dots,w(d-2))$ is colored white if all the conditions (4.8)–(4.10) below hold:

$$X((0,0,w(1),\dots,w(d-2)))=0, \tag{4.8}$$

$$X(\varepsilon_1,\varepsilon_2,w(1),\dots,w(d-2))=1 \text{ for } (\varepsilon_1,\varepsilon_2)=\pm(0,1) \text{ and for } (\varepsilon_1,\varepsilon_2)=\pm(1,1), \tag{4.9}$$

(4.10a) $X(\varepsilon_1,\varepsilon_2,w(1),\dots,w(d-2))=1$ if $(\varepsilon_1,\varepsilon_2)=(1,0)$ and $\sum_1^{d-2} w(i)$ is even, and if $(\varepsilon_1,\varepsilon_2)=(-1,0)$ and $\sum_1^{d-2} w(i)$ is odd,

(4.10b) $X(\varepsilon_1,\varepsilon_2,w(1),\dots,w(d-2))=0$ if $(\varepsilon_1,\varepsilon_2)=(1,0)$ and $\sum_1^{d-2} w(i)$ is odd, and if $(\varepsilon_1,\varepsilon_2)=(-1,0)$ and $\sum_1^{d-2} w(i)$ is even.

If any of (4.8)–(4.10) fails, then w is colored black. Conditions (4.8)–(4.10) are obvious analogues of (4.1)–(4.3). Basically only $v(3)$ in (4.3) has been replaced by $\sum_1^{d-2} w(i)$. It follows that if $(\mathbf{0},w_1,w_2,\dots)$ is a path on $\mathbb{Z}^{d-2}$ with all vertices colored white and with $\sum_{i=1}^{d-2} w_n(i)=n$ for all $n\ge 0$, then X^* defined on $H\times\mathbb{Z}_+$ by $X^*(\varepsilon_1,\varepsilon_2,n)=X(\varepsilon_1,\varepsilon_2,w_n)$ satisfies (4.1)–(4.3). Consequently, by Lemma 2, each word is seen from $\mathbf{0}$ along some path $(\mathbf{0},v_1,v_2,\dots)$ inside $H\times\{\mathbf{0},w_1,w_2,\dots\}$ with $\sum_{\ell=3}^d v_i(\ell)\ge\sum_{\ell=3}^d v_{i-1}(\ell)$.

It follows from the preceding that the probability in (4.5) is at least

$$P\{\exists \text{ oriented path } (\mathbf{0}, w_1, w_2, \dots) \text{ in } \mathbb{Z}_+^{d-2} \text{ with all vertices colored white}\}. \tag{4.11}$$

Since $H \times w'$ and $H \times w''$ are disjoint for $w' \neq w''$, the colors of the vertices in $\mathbb{Z}^{d-2}$ are independent. The probability of a particular vertex being white is 2^{-7}, because (4.8)–(4.10) prescribe the value of $X(\varepsilon_1, \varepsilon_2, w(1), \dots, w(d-2))$ for the seven values of $(\varepsilon_1, \varepsilon_2)$ in H. [Note that the values of $X(v)$ for $v \notin H \times \mathbb{Z}^{d-2}$ play no role at all in this argument.] If (4.4) holds, then the probability in (4.11) is strictly positive, so that (4.5) follows. Clearly (4.5) is stronger than (4.6). Finally (4.6) implies (4.7), again by the ergodic theorem as in Harris (1960).

The fact that (4.4) and hence (4.5)–(4.7) hold for $d \geq 132$ follows from the fact that [see Cox and Durrett (1983) with bond percolation replaced by site percolation]

$$\begin{aligned} &p_c(d-2, \text{ site, oriented}) \\ &\quad = \text{critical probability for oriented site percolation on } \mathbb{Z}_+^{d-2} \\ &\quad \leq P\{\text{two oriented simple random walks on} \\ &\qquad \mathbb{Z}_+^{d-2} \text{ (starting at } \mathbf{0}) \text{ intersect at some time} \geq 1\}. \end{aligned} \tag{4.12}$$

Imitating the proof of equation (2.2) in Cox and Durrett (1983), we find that the right-hand side of (4.12) is at most

$$\begin{aligned} &\frac{1}{d-2} + \frac{1}{(d-2)^2} - \frac{1}{(d-2)^3} + \sum_{k=3}^{d-2} k!(d-2)^{-k} \\ &\quad + \sum_{j=1}^{\infty} (d-2) \frac{(j(d-2))!}{(j!)^{d-2}} (d-2)^{-j(d-2)} \\ &\quad \leq \frac{1}{d-2} + \frac{1}{(d-2)^2} + \frac{5}{(d-2)^3} + 24 \frac{(d-5)}{(d-2)^4} \\ &\qquad + (2\pi)^{1/2} (d-2)^{3/2} \left(\frac{e^{1/13}}{\sqrt{2\pi}} \right)^{d-2} \left(1 + \frac{2}{d-5} \right). \end{aligned}$$

For $d \geq 132$ the last expression is less than 2^{-7} so that (4.4) holds. □

5. Words on branching process trees. In this section $\mathscr{G}$ will be a random tree, which is the family tree of a Bienaymé–Galton–Watson process. Readers unfamiliar with these trees can find a description in Harris [(1963), Section 6.2] or Jagers [(1975), Section 1.2]. These trees will be taken as rooted, ordered trees. That is, we think of the trees as imbedded in the plane, and with one vertex (which is called the *root*) singled out. The root will be denoted by $\langle 0 \rangle$. The set of vertices v which can be reached from $\langle 0 \rangle$ by a path $(\langle 0 \rangle, v_1, \dots, v_n = v)$ of $(n+1)$ vertices (including $\langle 0 \rangle$ and the endpoint v) are called the *vertices of the nth generation* (the zeroth generation consists of $\langle 0 \rangle$

only, by convention). The above path $(\langle 0\rangle, v_1, \dots, v_n = v)$ from $\langle 0\rangle$ to v is necessarily unique. The vertex v_{n-1} is called the *parent* of v_n. Also v_n is called a *child* of v_{n-1}, and a *descendant* of any of the vertices $\langle 0\rangle, v_1, \dots, v_{n-1}$. A typical vertex of the nth generation will be denoted as $\langle 0, i_1, \dots, i_n\rangle$ with $i_k \geq 1$. This vertex is sometimes referred to as the i_nth child of $\langle 0, i_1, \dots, i_{n-1}\rangle$.

We write τ for a generic tree and τ_n for the subtree consisting of its zeroth through nth generation. We shall also write $\tau(v)$ or $\tau(0, i, \dots, i_n)$ if $v = \langle 0, i_1, \dots, i_n\rangle$ for the subtree of τ consisting of v and its descendants. Finally, any string $\langle 0, i_1, \dots, i_n\rangle$ with $i_1 \geq 1, \dots, i_n \geq 1$ is a potential label of a vertex. However, for a given τ, there may not be a vertex in τ corresponding to this label. For instance if $\langle 0\rangle$ has only r children in τ, then only $\langle 0, i_1, \dots, i_n\rangle$ with $1 \leq i_1 \leq r$ can be the label of a vertex in τ. Similarly if $\langle 0, \bar{i}_1, \dots, \bar{i}_{k-1}\rangle$ is a vertex of τ with r children, then one must have $1 \leq i_k \leq r$ for all vertices $\langle 0, \bar{i}_1, \dots, \bar{i}_{k-1}, i_k, \dots, i_n\rangle$ which actually label a vertex of τ. If $\langle 0, i_1, \dots, i_n\rangle$ is the label of an actual vertex of τ, we also say that $\langle 0, i_1, \dots, i_n\rangle$ is *realized in* τ.

For the next section it is convenient to treat "periodic Bienaymé–Galton–Watson processes" rather than the standard branching processes only. We assume that there exist ν distributions on $\mathbb{Z}_+$, $p_k^{(i)}$, $k \geq 0$, $0 \leq i \leq \nu - 1$, such that

$$\sum_{k\geq 0} p_k^{(i)} = 1, \qquad 0 \leq i \leq \nu - 1, \tag{5.1}$$

and

(5.2) the numbers of children of all the vertices are independent; each vertex in generation $\ell\nu + i$, $\ell \geq 0$, $0 \leq i < \nu$, has k children with probability $p_k^{(i)}$.

We set

$$m^{(i)} = \sum_{k=0}^{\infty} k p_k^{(i)}. \tag{5.3}$$

This is the mean number of children of any vertex in generation $\ell\nu + i$, $\ell \geq 0$, $0 \leq i < \nu$. We allow $m^{(i)}$ to take the value $+\infty$. Here is a slightly more general version of Theorem 3 than stated in the Introduction.

THEOREM 3. *Let $\mathscr{G}$ be the random family tree of a Bienaymé–Galton–Watson process with period ν, as above. If*

$$M := m^{(0)} m^{(1)} \cdots m^{(\nu-1)} > 2^{\nu}, \tag{5.4}$$

then a.e. (*with respect to the branching process*) *on the event* $\{\mathscr{G}$ *is infinite*$\}$, $\mathscr{G}$ *contains finitely many vertices* $v_1, \dots, v_r$ *for which*

$$P\{S(v_1, \dots, v_r) = \Xi\} > 0 \tag{5.5}$$

and almost everywhere on $\{\mathscr{G}$ *is infinite*$\}$, $\mathscr{G}$ *has the property*

$$P\{S_\infty = \Xi\} = 1. \tag{5.6}$$

PROOF. Without loss of generality we may assume that the distributions $\{p_k^{(i)}\}$ have bounded support; that is, for some $K < \infty$

$$p_k^{(i)} = 0 \quad \text{for } k > K,\ 0 \le i \le \nu - 1. \tag{5.7}$$

Obviously we can always achieve this by a truncation at sufficiently high K, so that (5.4) still holds for the truncated process. Moreover $\mathscr{G}$ is always at least as large as the family tree of the truncated process, so that (5.5) and (5.6) for the truncated process imply (5.5) and (5.6) for the original process.

We also must take care of some simple measure–theoretic issues. Let $\mathscr{T}$ be the space of all family trees (note that these are simply all locally finite rooted, ordered trees). Since we shall be picking a random element of $\mathscr{T}$, its vertex set is also random and we must therefore adopt a slightly more complicated choice for our configuration space Ω than before. The simplest way is to pick i.i.d. random variables $X(v)$, one for each *potential* $v = \langle 0 \rangle$ or $\langle 0, i_1, \dots, i_n \rangle$, $i_k \ge 1$. We then simply use only those $X(v)$ for which v is realized in our tree. Therefore, we take in this section $\Omega = \{0,1\}^{\mathscr{W}}$, where $\mathscr{W} = \langle 0 \rangle \cup \bigcup_{n \ge 1} \{\langle 0, i_1, \dots, i_n \rangle \colon i_k \ge 1, 1 \le k \le n\}$. We then define

$$\Theta = \{(\tau, \omega) \in \mathscr{T} \times \Omega\colon \tau \text{ is infinite and every } \xi \text{ is seen somewhere on } \tau \text{ in the restriction of } \omega \text{ to the realized vertices on } \tau\}.$$

We also introduce the obvious σ–field $\mathscr{C}$ say, in $\mathscr{T}$, namely, the smallest σ–field which contains all sets of the form $\{\tau_n = \overline{\tau}\}$ for $\overline{\tau}$ any finite tree of height n (i.e., with exactly the zeroth through nth generation nonempty). As before $\mathscr{B}$ is the σ–field in Ω generated by $X(0)$ and all $X(0, i_1, \dots, i_n)$. One should then first show that

$$\Theta \in \mathscr{C} \times \mathscr{B}. \tag{5.8}$$

However, by Remark 1 in Section 2,

$$\Theta = \bigcup_{r, v_1, \dots, v_r} \{(\tau, \omega)\colon \tau \text{ is infinite}, v_1, \dots, v_r \text{ are realized in } \tau \text{ and } S(v_1, \dots, v_r; \omega) = \Xi \text{ on } \tau\}, \tag{5.9}$$

where the v_j run through all the potential vertices, that is, the elements of $\mathscr{W}$. Each of the sets in braces in (5.9) is $\mathscr{C} \times \mathscr{B}$ measurable by an argument very similar to the proof of Proposition 2, which we leave to the reader.

On $\mathscr{C}$ we have the measure induced by the periodic branching process specified by the $p_k^{(i)}$, $0 \le i \le \nu - 1$, $k \ge 0$, while on $\mathscr{B}$ we have the measure (still denoted by P) under which $X(0)$ and $X(0, i_1, \dots, i_n)$, $n \ge 1$, $i_k \ge 1$, are i.i.d. and take the values 0 and 1 each with probability $\frac{1}{2}$. For the remainder of this proof we use $\mathbb{P}$ for the product of these two measures on $\mathscr{C} \times \mathscr{B}$; $\mathbb{E}$ will denote expectation with respect to $\mathbb{P}$. By Fubini's theorem, (5.6) is then equivalent to

$$\mathbb{P}\{\{\tau \text{ is infinite}\} \setminus \Theta\} = 0. \tag{5.10}$$

Moreover, (5.5) is an immediate consequence of (5.6) and (5.9).

It therefore suffices to prove (5.10) and we shall do this now. Let $\varepsilon > 0$ be such that

$$M > 2^{\nu}(1+2\varepsilon)^{\nu}$$

[see (5.4) for M] and for each $n \geq 1$ and $(\eta_1, \dots, \eta_{n\nu}) \in \{0,1\}^{n\nu}$, define the event

$$H_n(\eta_1, \dots, \eta_{n\nu}) = \{\text{there are at least } (1+\varepsilon)^{n\nu} \text{ vertices } \langle 0, i_1, \dots, i_{n\nu}\rangle \text{ in the } n\nu\text{th generation of } \tau \text{ with } X(\langle 0, i_1, \dots, i_k\rangle) = \eta_k \text{ for all } k \leq n\nu\}. \tag{5.11}$$

Thus H_n is the event that $(\eta_1, \dots, \eta_{n\nu})$ is seen along at least $(1+\varepsilon)^{n\nu}$ self–avoiding paths on τ starting at the root; these paths do not have to be disjoint, but only have to have distinct final vertices. We shall show below that

$$\sum_{n=0}^{\infty} 2^{n\nu}\mathbb{P}\{H_n(1, \dots, 1) \setminus H_{n+1}(1, \dots, 1)\} < \infty. \tag{5.12}$$

Before doing this we show that (5.12) implies (5.10). It is easy to see that

$$\mathbb{P}\{H_n(\eta_1, \dots, \eta_{n\nu}) \setminus H_{n+1}(\eta_1, \dots, \eta_{(n+1)\nu})\}$$

has the same value for all $\eta_1, \dots, \eta_{(n+1)\nu}$ [compare the argument following (2.9)]. Therefore (5.12) shows that a.e. [$\mathbb{P}$] there exists an $N = N(\tau, \omega) < \infty$ such that for all $n \geq N$ and all $\eta_1, \dots, \eta_{(n+1)\nu}$, $H_n(\eta_1, \dots, \eta_{n\nu}) \setminus H_{n+1}(\eta_1, \dots, \eta_{(n+1)\nu})$ does not occur. However, it is well known [see Harris (1963), Remark 1.8.1.1] that, by virtue of (5.7), a.e. on $\{\tau \text{ is infinite}\}$, the number of vertices in the $n\nu$th generation of τ is at least $WM^n \geq W2^{n\nu}(1+2\varepsilon)^{n\nu}$ for all large n and some (random) $W > 0$. In particular, there will eventually be at least $2^{n\nu}(1+\varepsilon)^{n\nu}$ vertices in the $n\nu$th generation. Since there are $2^{n\nu}$ choices for $(\eta_1, \dots, \eta_{n\nu})$, $H_n(\overline{\eta}_1, \dots, \overline{\eta}_{n\nu})$ must occur for some $n \geq N$ and some choice of $\overline{\eta}_1, \dots, \overline{\eta}_{n\nu}$. By definition of N then also $H_m(\overline{\eta}_1, \dots, \overline{\eta}_{n\nu}, \eta_{n\nu+1}, \dots, \eta_{m\nu})$ occurs for any $m > n$ and all continuations $\eta_{n\nu+1}, \dots, \eta_{m\nu}$. Therefore every word is seen from some vertex in the $n\nu$th generation. Thus (5.12) indeed implies (5.10), and we now turn to the proof of (5.12).

Statement (5.12) follows almost immediately from standard large deviation estimates. Assume that $H_n = H_n(1, \dots, 1)$ occurred, so that there exist at least $(1+\varepsilon)^{n\nu}$ vertices $v_1, \dots, v_r$ in the $n\nu$th generation such that $X(v) = 1$ on each vertex other than the root on the paths in τ from $\langle 0\rangle$ to $\langle v_i\rangle$, $1 \leq i \leq r$. This event depends only on the first $n\nu$ generations of τ and the $X(v)$ for v in one of these first $n\nu$ generations. Therefore, conditionally on H_n and the $v_1, \dots, v_r$, the following random variables are independent of each other:

$$U_n(v_i) := \{\text{number of descendants } w \text{ of } v_i \text{ in the } (n+1)\nu\text{th generation of } \tau \text{ with the property that } X(u) = 1 \text{ for all } u \text{ on the path from } v_i \text{ to } w \text{ in } \tau\}.$$

Moreover,

$$U_n(v_i) \geq 0 \quad \text{and} \quad \mathbb{E}\{U_n(v_i) \mid H_n, v_1, \dots, v_r\} = 2^{-\nu}M \geq (1+2\varepsilon)^{\nu}. \tag{5.13}$$

Bernstein's inequality [see Rényi (1970), Section 7.4] now implies that there exists a constant $C_3 = C_3(\varepsilon) > 0$ such that

$$\mathbb{P}\left\{\sum_1^r U_n(v_i) \le r(1+\varepsilon)^\nu \mid H_n, v_1, \dots, v_r\right\} \le 2\exp(-C_3 r). \tag{5.14}$$

[Note that the $U_n(v_i)$ are bounded, because of (5.7).] However, if H_n occurs, and hence $r \ge (1+\varepsilon)^{n\nu}$, and

$$\sum_1^r U_n(v_i) > r(1+\varepsilon)^\nu \ge (1+\varepsilon)^{(n+1)\nu}, \tag{5.15}$$

then there are at least $(1+\varepsilon)^{(n+1)\nu}$ vertices w in the $(n+1)\nu$th generation of τ such that $X(u) = 1$ for all u other than $\langle 0\rangle$ on the path from $\langle 0\rangle$ to w [these are precisely the vertices counted by the $U_n(v_i)$]. Thus (5.15) implies H_{n+1}, and (5.14) shows that

$$\mathbb{P}\{H_{n+1} \text{ fails} \mid H_n\} \le 2\exp(-C_3(1+\varepsilon)^{n\nu}); \tag{5.16}$$

(5.12) is now immediate. □

6. Trees on which all words are seen from $(k+1)$, but not from k points. We use the notation of the preceding section. The trees in this section are assumed to be deterministic, rooted and ordered and oriented away from the root, with the following periodic form: for some $\nu \ge 1$, $\rho > 1$,

(6.1) each vertex in the ℓth generation has exactly ρ children if $\nu \mid \ell$ and exactly 1 child if $\nu \nmid \ell$.

These are infinite trees and when $\nu = 1$, we obtain the regular ρ–ary tree. The reader is advised to concentrate on the special case $\nu = 1$, $\rho = 3$, $k = 1$ (the regular 3–tree) when reading the proofs of Lemmas 3 and 4 below. Let $k \ge 0$ be given. We can then choose ν and ρ such that

$$\frac{k+1}{k} > \rho 2^{-\nu} \tag{6.2}$$

(with the left-hand side interpreted as $+\infty$ if $k = 0$), but also

$$\rho 2^{-\nu} \ge \frac{k+2}{k+1}. \tag{6.3}$$

LEMMA 3. *Let $k \ge 0$. If $\mathscr{G}$ is a rooted tree for which* (6.1) *and* (6.2) *hold, then for each k–tuple of vertices $v_1, \dots, v_k$,*

$$P\{\mu(S(v_1, \dots, v_k)) = 1\} = 0 \tag{6.4}$$

and a fortiori

$$P\{S(v_1, \dots, v_k) = \Xi\} = 0. \tag{6.5}$$

PROOF. The lemma is vacuous for $k = 0$, so that we assume $k \geq 1$. Let $w_1, \dots, w_k$ be k vertices of $\mathscr{G}$. Assume w_i belongs to the t_ith generation of $\mathscr{G}$. Then, by (6.1) and (6.2),

$$\sum_{i=1}^{k} (\text{number of descendants of } w_i \text{ in the } (t_i + \nu)\text{th generation})$$
$$= k\rho < (k+1)2^\nu.$$

Now consider the paths in $\mathscr{G}$ from the w_i to their descendants in the $(t_i+\nu)$th generation, $1 \leq i \leq k$. Let $(w_i, u_1, \dots, u_\nu)$ be a typical path of this form. Then $(X(u_1), \dots, X(u_\nu))$ is a word of length ν. We have $k\rho$ such paths, while there exist 2^ν words of length ν. Since $k\rho < (k+1)2^\nu$, at least one word of length ν occurs on at most k paths.

We now start with $w_{i,0} = v_i$, $1 \leq i \leq k$. Write $t_{i,0}$ for the generation number of $w_{i,0}$. By the above, there exists at least one string $(\eta_1, \dots, \eta_\nu) \in \{0,1\}^\nu$ which is seen at most k times on the paths of length ν starting at the $w_{i,0}$. If this $(\eta_1, \dots, \eta_\nu)$ is not seen on any such paths, then we do not see any word in $C_\nu(\eta_1, \dots, \eta_\nu)$ from $v_1, \dots, v_k$ [see (2.2) for C_ν], that is,

$$S(v_1, \dots, v_k) \cap C_\nu(\eta_1, \dots, \eta_\nu) = \varnothing. \tag{6.6}$$

In this case we make no further choices. If (6.6) fails, let $w_{j,1}$, $1 \leq j \leq r$, be all the descendants of $w_{1,0}, \dots, w_{k,0}$ in the generations $t_{i,0} + \nu$, $1 \leq i \leq k$, respectively, for which $\eta_1, \dots, \eta_\nu$ is seen in the last ν generations of the path from $\langle 0 \rangle$ to $w_{j,1}$ (such a path necessarily goes through one of the $w_{i,0}$). By our choice of $\eta_1, \dots, \eta_\nu$ we must have $r \leq k$. We now repeat our procedure starting from the $r \leq k$ vertices $w_{1,1}, \dots, w_{r,1}$. There then must exist $(\eta_{\nu+1}, \dots, \eta_{2\nu}) \in \{0,1\}^\nu$ such that $(\eta_{\nu+1}, \dots, \eta_{2\nu})$ is seen at most k times along the paths from the $w_{j,1}$ to their descendants in the $(t_{j,1} + \nu)$th generation (where $t_{j,1}$ is the generation number of $w_{j,1}$). Again, if $(\eta_{\nu+1}, \dots, \eta_{2\nu})$ is not seen at all on these paths, then

$$S(v_1, \dots, v_k) \cap C_{2\nu}(\eta_1, \dots, \eta_{2\nu}) = \varnothing. \tag{6.7}$$

We iterate this procedure. At each stage s, start from at most k vertices and find $(\eta_{s\nu+1}, \dots, \eta_{(s+1)\nu})$ which is seen at most k times from these vertices. At each stage there is a probability at least

$$2^{-k\rho}$$

that some string $(\eta_{s\nu+1}, \dots, \eta_{(s+1)\nu})$ is not seen at all and, consequently, a.e. $[P]$, there exists an $s \geq 0$ and $\eta_1, \dots, \eta_{(s+1)\nu}$ such that

$$S(v_1, \dots, v_k) \cap C_{(s+1)\nu}(\eta_1, \dots, \eta_{(s+1)\nu}) = \varnothing.$$

Thus (6.4) holds. □

LEMMA 4. *Let $k \geq 0$. If $\mathscr{G}$ is a rooted tree for which* (6.1) *and* (6.3) *hold, and $v_1, \dots, v_{k+1}$ are $k+1$ vertices of the $s\nu$th generation for some s, then*

$$P\{S(v_1, \dots, v_{k+1}) = \Xi\} > 0. \tag{6.8}$$

PROOF. We begin with a deterministic construction. If $w_1, \dots, w_\ell$ are ℓ vertices in the $t\nu$th generation with $\ell \ge k+1$ and $\pi_1, \dots, \pi_{\ell\rho}$ are the paths in $\mathscr{G}$ from $w_1, \dots, w_\ell$ to the $(t+1)\nu$th generation, then we can choose values $\overline{X}(u) \in \{0,1\}$ for u on these paths (excluding the points $w_1, \dots, w_\ell$) such that each $(\eta_1, \dots, \eta_\nu)$ is seen along at least $\lfloor \ell\rho 2^{-\nu} \rfloor \ge (\ell+1)$ of the paths π_q. To prove this, let $\pi_{j\rho+1}, \dots, \pi_{(j+1)\rho}$ be the paths which start at w_j. These paths are of the form $\pi_q = (w_j, u_{q,1}, \dots, u_{q,\nu})$ for $j\rho + 1 \le q \le (j+1)\rho$. Since w_j belongs to the $t\nu$th generation, it has ρ children; these are the ρ vertices $u_{q,1}$ and they are distinct. Thus $\pi_{q'}$ and $\pi_{q''}$ can at most have a vertex w_j in common, and all the values $\overline{X}(u_{q,r})$, $1 \le q \le \ell$, $1 \le r \le \nu$, can be taken arbitrarily. In particular, we can choose the $\overline{X}(u_{q,r})$ such that for each of the 2^ν possibilities for $(\eta_1, \dots, \eta_\nu)$, there are at least $\lfloor \ell\rho 2^{-\nu} \rfloor$ of the π_q with $\overline{X}(u_{q,r}) = \eta_r$, $1 \le r \le \nu$. This proves our claim, since [see (6.3)]

$$\ell\rho = \frac{\ell}{\ell+1}(\ell+1)\rho \ge \frac{k+1}{k+2}(\ell+1)\rho \ge (\ell+1)2^\nu. \tag{6.9}$$

If the $\overline{X}(u)$'s for u in the $(t\nu+1)$th to $(t+1)\nu$th generation are chosen as above, then for each $\eta_1, \dots, \eta_\nu$ we can find $\ell_1 := \lfloor \ell\rho 2^{-\nu} \rfloor$ vertices $w_{1,1}, \dots, w_{\ell_1,1}$ in the $(t+1)\nu$th generation such that $\eta_1, \dots, \eta_\nu$ is seen on the self–avoiding path from one of the w_i to $w_{j,1}$ for $1 \le j \le \ell_1$. Repeating the above procedure we can find for each $\eta_1, \dots, \eta_{2\nu}$, $\ell_2 := \lfloor \ell_1 \rho 2^{-\nu} \rfloor$ vertices $w_{1,2}, \dots, w_{\ell_2,2}$ in the $(t+2)\nu$th generation such that $(\eta_1, \dots, \eta_{2\nu})$ is seen on the self–avoiding path from one of the w_i (and through one of the $w_{j,1}$) to $w_{k,2}$, for $1 \le k \le \ell_2$. After j iterations, we have for each $\eta_1, \dots, \eta_{j\nu}$ at least ℓ_j vertices $w_{k,j}$ such that $\eta_1, \dots, \eta_{j\nu}$ is seen on a self–avoiding path from some w_i to $w_{k,j}$, $1 \le k \le \ell_j$. Here ℓ_j is defined recursively by

$$\ell_0 = k+1, \qquad \ell_j = \lfloor \ell_{j-1}\rho 2^{-\nu} \rfloor. \tag{6.10}$$

We apply the above procedure J times for some large J to be determined below and with $\ell = k+1$ and with $w_1, \dots, w_{k+1}$ equal to $(k+1)$ given vertices $(v_1, \dots, v_{k+1})$ in the $s\nu$th generation. Then there is at least one choice of the $\overline{X}(u)$ for u in the first $(s+J)\nu$ generations of $\mathscr{G}$, for which each string $(\eta_1, \dots, \eta_{J\nu})$ is seen at least ℓ_J times from one of $v_1, \dots, v_{k+1}$. Then also

$$P\{\text{each } (\eta_1, \dots, \eta_{J\nu}) \text{ is seen along at least } \ell_J \text{ paths starting at some } v_i, 1 \le i \le k+1\} > 0. \tag{6.11}$$

Now choose $\varepsilon > 0$ such that

$$\rho \ge 2^\nu (1+2\varepsilon)^\nu \tag{6.12}$$

and replace the H_n of (5.11) by

$$\tilde{H}_n(\eta_1, \dots, \eta_{(J+n)\nu}) := \{\text{there are at least } \ell_J(1+\varepsilon)^{n\nu} \text{ vertices } \langle 0, i_1, \dots, i_{(s+J+n)\nu}\rangle \text{ in the } (s+J+n)\nu\text{th generation of } \mathscr{G} \text{ with } \langle 0, i_1, \dots, i_{s\nu}\rangle \in \{v_1, \dots, v_{k+1}\} \text{ and } X(0, i_1, \dots, i_{s\nu+r}) = \eta_r \text{ for } 1 \le r \le (J+n)\nu\}.$$

Then (6.11) says that

$$P\{\tilde{H}_0(\eta_1,\dots,\eta_{J\nu}) \text{ for all } (\eta_1,\dots,\eta_{J\nu})\} > 0. \tag{6.13}$$

Furthermore, the same argument as used for (5.14) shows that on the event $\tilde{H}_n(\eta_1,\dots,\eta_{(J+n)\nu})$,

$$P\{\tilde{H}_{n+1}(\eta_1,\dots,\eta_{(J+n+1)\nu}) \text{ fails}|\ X(u) \text{ for all } u \text{ in the first } (s+J+n)\nu \text{ generations}\} \le 2e^{-C_3\ell_J(1+\varepsilon)^{n\nu}}.$$

Therefore,

$$\begin{aligned} &P\{\tilde{H}_n(\eta_1,\dots,\eta_{(J+n)\nu}) \text{ fails for some } n \ge 1 \text{ and } (\eta_1,\dots,\eta_{(J+n)\nu}) \mid X(u) \text{ for all } u \text{ in the first } (s+J)\nu \text{ generations}\} \\ &\quad\le \sum_{n=1}^{\infty} 2^{(J+n)\nu} 2e^{-C_3\ell_J(1+\varepsilon)^{(n-1)\nu}} \end{aligned} \tag{6.14}$$

on the event

$$\{\tilde{H}_0(\eta_1,\dots,\eta_{J\nu}) \text{ for all } (\eta_1,\dots,\eta_{J\nu})\}. \tag{6.15}$$

Finally we fix J such that the right-hand side of (6.14) is less than $\frac{1}{2}$. To see that this is possible, note first that by (6.10) and (6.9) we have $\ell_j \ge k + j + 1$ (by induction on j). Once ℓ_j is so large that

$$\ell_j \rho 2^{-\nu} \ge \ell_j(1+\varepsilon)^{\nu} + 1$$

[see (6.12)], it follows that

$$\ell_{j+1} \ge \ell_j(1+\varepsilon)^{\nu}.$$

Thus, there exists a j_0 such that for $j \ge j_0$,

$$\ell_j \ge (k + j_0 + 1)(1+\varepsilon)^{\nu(j-j_0)},$$

and one easily sees from this that the right-hand side of (6.14) will be less than 1/2 for large J. For such a J it follows from (6.13) and (6.14) that

$$\begin{aligned} &P\{\tilde{H}_n(\eta_1,\dots,\eta_{(J+n)\nu}) \text{ occurs for all } n \ge 0 \text{ and all } (\eta_1,\dots,\eta_{(J+n)\nu})\} \\ &\quad\ge \tfrac{1}{2} P\{\tilde{H}_0(\eta_1,\dots,\eta_{J\nu}) \text{ occurs for all } \eta_1,\dots,\eta_{J\nu}\} > 0. \end{aligned}$$

Since each $(\eta_1,\eta_2,\dots)$ is seen from one of $v_1,\dots,v_{k+1}$ if $\tilde{H}_n(\eta_1,\dots,\eta_{(J+n)\nu})$ occurs for all n and all $\eta_1,\eta_2,\dots,\eta_{(J+n)\nu}$, this proves (6.8). □

Lemmas 3 and 4 together yield the following theorem.

THEOREM 4. *Let $k \ge 0$ be given. Then if $\mathscr{G}$ is a rooted tree for which* (6.1)–(6.3) *hold, it satisfies* (6.4), (6.5) *and* (6.8), *so that not all* (*not even almost all*) *words can be seen from k vertices, but they are seen with probability* 1 *from* $(k+1)$ *vertices; that is,*

$$P\{S(v_1,\dots,v_{k+1}) = \Xi \text{ for some } (k+1)\text{-tuple of vertices}\} = 1. \tag{6.16}$$

PROOF. Only (6.16) needs further proof. However, this is immediate from the following ergodic considerations. Let $v_j = \langle 0, i_{1,j}, \dots, i_{s\nu,j}\rangle$, $j = 1, \dots, k+1$ be $k+1$ vertices in the $s\nu$th generation. Then (6.8) holds for these $v_1, \dots, v_{k+1}$. Consider the random sets

$$S^{(n)}(v_1, \dots, v_{k+1}) := S(v_1^{(n)}, \dots, v_{k+1}^{(n)}),$$

where

$$v_j^{(n)} = \langle 0, 1, \dots, 1, i_{1,j}, \dots, i_{s\nu,j}\rangle$$

with $n\nu$ ones between the 0 and $i_{1,j}$. Then $v_j^{(n)}$, $1 \le j \le k+1$, have the same relation to $\langle 0, 1, \dots, 1\rangle$ (with $n\nu$ ones) as v_j, $1 \le j \le n$, have to the root $\langle 0\rangle$. Therefore, $\{S^{(n)}(v_1, \dots, v_{k+1})\colon n \ge 0\}$ is a stationary ergodic sequence [compare Harris (1960)] and each $S^{(n)}(v_1, \dots, v_{k+1})$ has the same distribution as $S^{(0)}(v_1, \dots, v_{k+1}) = S(v_1, \dots, v_{k+1})$. It therefore follows from (6.8) and the ergodic theorem that w.p.1, $S^{(n)}(v_1, \dots, v_{k+1}) = \Xi$ for infinitely many n (in fact for a sequence of n's of strictly positive density). □

COROLLARY 2. *On the oriented regular* 3*–tree* (*each vertex other than the root has degree* 4) *one does not see almost all words from one vertex, but one does see all words from two vertices. On the oriented regular b–tree* [*each vertex other than the root has degree* $(b+1)$] *with* $b \ge 4$, *one sees all words from one vertex.*

COROLLARY 3. *On the positively oriented* $\mathbb{Z}_+^3$ *one does not see almost all words from one vertex, so that* (1.7) *holds.*

Corollary 2 follows by taking $\nu = 1$, $\rho = 3$ or $\rho \ge 4$ in this section. Corollary 3 does not follow directly, but the proof of Lemma 3 with $\nu = 1$, $\rho = 3$, $k = 1$ goes through without changes for $\mathscr{G} = \mathbb{Z}_+^3$.

7. A tree on which one sees almost all but not all words. To obtain an example with the properties in the title of this section, we take $\mathscr{G}$ of the following form: $\mathscr{G}$ is a rooted oriented tree and for sequences $\{n_i\}_{i\ge 1}$ and $\{r_i\}_{i \ge 1}$ to be determined below it holds that

$$\text{number of children of any given vertex } v = \begin{cases} n_i, & \text{if } v \text{ belongs to the } r_i\text{th generation,} \\ 1, & \text{if the generation number of } v \text{ is not one of the } r_i \end{cases} \tag{7.1}$$

(see Figure 3). We denote this tree by $\mathscr{G}(n_i, r_i)$.

THEOREM 5. *Let* $\mathscr{G} = \mathscr{G}(n_i, r_i)$, $\ell(i) = r_{i+1} - r_i$ *and*

$$\nu(i) = \prod_{k=1}^{i} n_k = \textit{number of vertices in the } (r_i + 1)\textit{th generation of } \mathscr{G}. \tag{7.2}$$

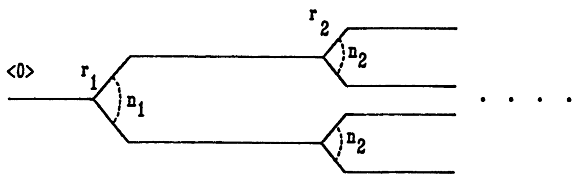

FIG. 3. *Illustration of* $\mathscr{G}(n_i, r_i)$.

It is possible to choose $\{n_i\}$ *and* $\{r_i\}$ *such that*

$$\sum_{i=1}^{\infty}\{1-[1-(1-2^{-\ell(i)})^{n_i}]^{\nu(i-1)}\} < \infty \tag{7.3}$$

and

$$\sum_{i=1}^{\infty}\{1-[1-(1-2^{-\ell(i)})^{\nu(i)}]^{2^{\ell(i)}}\} = \infty. \tag{7.4}$$

For such a choice of $\{n_i\}$ *and* $\{r_i\}$, *it holds on* $\mathscr{G}$ *that*

$$P\{S_\infty = \Xi\} = 0, \tag{7.5}$$

but

$$P\{\mu(S_\infty) = 1\} = 1. \tag{7.6}$$

Also

(7.7) $P\{\mu(S(v_1,\dots,v_k)) = 1$ *for any finite number of vertices* $v_1,\dots,v_k\} = 0$.

PROOF. Denote by $v_{1,i},\dots,v_{\nu(i),i}$ the vertices of the (r_i+1)th generation. Between the (r_i+1)th generation and the r_{i+1}th generation the only paths on $\mathscr{G}(n_i, r_i)$ are the "straight line segments" starting from the $v_{j,i}$, $1 \le j \le \nu(i)$, and consisting of $\ell(i) := r_{i+1} - r_i$ vertices. We also give a name to the parents of the $v_{j,i}$, $1 \le j \le \nu(i)$, that is, to the vertices in the r_ith generation. If we denote these vertices by $u_{1,i},\dots,u_{\nu(i-1),i}$, then we can number the vertices $v_{j,i}$ in such a way that the children of $u_{s,i}$ are precisely the vertices $\{v_{j,i}\colon (s-1)n_i < j \le sn_i\}$ for $1 \le s \le \nu(i-1)$. Let $\pi_{j,i}$, $1 \le j \le \nu(i)$, be the paths in $\mathscr{G}$ from the $u_{k,i}$ in the r_ith generation to the r_{i+1}th generation. A typical such path is $\pi_{j,i} = (u_{s,i}, w_{1,j,i},\dots,w_{\ell(i),j,i})$ with $w_{1,j,i} = v_{j,i}$ and $u_{s,i}$ equal to its parent [i.e., $(s-1)n_i < j \le sn_i$] and with $w_{k+1,j,i}$ equal to the unique child in $\mathscr{G}$ of $w_{k,j,i}$ if $1 \le k < \ell(i)$. Now define the following events:

$$\begin{aligned} K_i(\overline{\eta}) &= K_i(\eta_1,\dots,\eta_{\ell(i)}) \\ &= \{\text{for some } 1 \le s \le \nu(i-1), \text{ the word } \overline{\eta} = (\eta_1,\dots,\eta_{\ell(i)}) \text{ is not seen} \\ &\qquad \text{along any of the paths } \pi_{j,i} \text{ with } (s-1)n_i < j \le sn_i\}, \end{aligned}$$

$$\begin{aligned} L_i = \{&\text{there exists a word } \overline{\eta} = (\eta_1,\dots,\eta_{\ell(i)}) \in \{0,1\}^{\ell(i)} \text{ which is} \\ &\text{not seen along any of the paths } \pi_{j,i},\ 1 \le j \le \nu(i)\}. \end{aligned}$$

Then K_i occurs if and only if $\overline{\eta}$ is not seen on some group of paths, descending from one vertex in the r_ith generation. Note that the $\pi_{j,i}$ for different j have at most their initial points in common, and recall that for a word to be seen along a path, the path's initial vertex is ignored. Therefore, the words seen along the $\pi_{j,i}$ for different j are independent. It follows that

$$P\{K_i(\overline{\eta})\} = 1 - [1 - (1 - 2^{-\ell(i)})^{n_i}]^{\nu(i-1)}. \tag{7.8}$$

We note in passing that we could complete the proof below by only estimating the probability that $\overline{\eta}$ is not seen along any path $\pi_{j,i}$ descending from *one* $u_{s,i}$, say along the paths $\pi_{j,i}$ with $1 \le j \le n_i$. [Thus we would replace "for $1 \le s \le \nu(i-1)$" in the definition of K_i by "for $s = 1$."] This would allow us to drop the exponent $\nu(i-1)$ in (7.8) and (7.3). This would give a slightly better theorem, but would complicate the proof, so we shall not pursue this.

Our aim is to make sure that K_i happens for only finitely many i (for some given $\overline{\eta}$'s), but that L_i happens for infinitely many i. This will be possible, because L_i is a *union* over $\overline{\eta}$'s and there will be many choices for $\overline{\eta}$. To estimate the probability of L_i, note that along each path $\pi_{j,i}$ some word of length $\ell(i)$ is seen, and the words seen along distinct paths are independent, each word of length $\ell(i)$ having probability $2^{-\ell(i)}$ of being seen along a given $\pi_{j,i}$. Thus, if we interpret each of the $2^{\ell(i)}$ possible words as an urn and each of the $\nu(i)$ paths as a ball which is to be put into the urns, each urn having probability $2^{-\ell(i)}$ of receiving the ball, then $P\{L_i\}$ is simply the probability of finding at least one urn empty. Thus

$$\begin{aligned} P\{L_i\} &= 1 - P\{\text{each urn contains at least one ball}\} \\ &\ge 1 - [P\{\text{a given urn contains at least one ball}\}]^{2^{\ell(i)}} \\ &= 1 - [1 - (1 - 2^{-\ell(i)})^{\nu(i)}]^{2^{\ell(i)}}. \end{aligned} \tag{7.9}$$

The inequality here can be found in Mallows (1968).

We first show how to choose n_i and r_i in such a way that (7.3) and (7.4) hold. This is easy to do in a recursive way. Take $r_1 = 0$. Assume that $n_1, \dots, n_{i-1}$ and $r_1, \dots, r_i$ have already been chosen. Then $\ell(1), \dots, \ell(i-1)$ and $\nu(1), \dots, \nu(i-1)$ are also determined. Recall also that $\nu(i) = n_i \nu(i-1)$ and express (7.3) and (7.4) as

$$\sum_{i=1}^{\infty} \{1 - [1 - \rho_i]^{\nu(i-1)}\} < \infty \tag{7.10}$$

and

$$\sum_{i=1}^{\infty} \{1 - [1 - \rho_i^{\nu(i-1)}]^{2^{\ell(i)}}\} = \infty, \tag{7.11}$$

respectively, with

$$\rho_i = (1 - 2^{-\ell(i)})^{n_i}.$$

We now relate $\ell(i)$ (and hence r_{i+1}) and n_i to each other such that

$$[1-\rho_i]^{\nu(i-1)} > 1 - \frac{1}{i^2}.$$

This will guarantee (7.10) and (7.3) and can be achieved by making ρ_i sufficiently small or $n_i 2^{-\ell(i)}$ sufficiently large. It still allows us to take $2^{\ell(i)}$ as large as we wish, if we make n_i correspondingly larger. In particular, we can still take $\ell(i)$ so large that

$$\left[1-\rho_i^{\nu(i-1)}\right]^{2^{\ell(i)}} < \frac{1}{2}.$$

This will make the ith summand in (7.11) or (7.4) greater than $\frac{1}{2}$, and doing this for all i, we obtain (7.4).

Now that we have shown that (7.3) and (7.4) can be satisfied, we show that they imply (7.6) and (7.7). Equation (7.5) is implied by (7.7), by Remark 1 in Section 2 and the fact that $\mu(S(v_1,\dots,v_k)) < 1$ implies $S(v_1,\dots,v_k) \neq \Xi$. We first prove (7.7). Since there are only countably many k–tuples $v_1,\dots,v_k$, it suffices to show for any fixed $v_1,\dots,v_k$ that

$$\mu(S(v_1,\dots,v_k)) < 1 \quad \text{a.e. } [P]. \tag{7.12}$$

However, a.e. $[P]$, L_i occurs for infinitely many i, by virtue of (7.9) and (7.4) and independence of the L_i for $i \geq 1$. Thus, w.p.1 there exist $i(1) < i(2) < \cdots$ and corresponding words

$$\overline{\eta}_p = (\eta_{1,p},\dots,\eta_{\ell(i(p)),p}) \in \{0,1\}^{\ell(i(p))}$$

so that $\overline{\eta}_p$ is not seen on any path starting in generation $r_{i(p)}+1$ and ending in generation $r_{i(p)+1}$. In particular, if v_q belongs to the $\gamma(q)$th generation, then one does not see from v_q any infinite word $\xi = (\xi_1, \xi_2, \dots)$ with

$$\xi_{r_{i(p)}+j-\gamma(q)} = \eta_{j,p}, \qquad 1 \leq j \leq \ell(i(p)), \tag{7.13}$$

for any $r_{i(p)} \geq \gamma(q)$. Therefore, if $p(1),\dots,p(k)$ are such that

$$r_{i(p(q))} \geq \gamma(q), \qquad 1 \leq q \leq k,$$

and

$$r_{i(p(q+1))-\gamma(q+1)} > r_{i(p(q))+1}, \qquad 1 \leq q < k$$

(so that the intervals $[r_{i(p(q))-\gamma(q)+1}, r_{i(p(q)+1)}]$ for the various q are disjoint), then one does not see from any $v_1,\dots,v_k$ any word ξ in the cylinder of words which satisfy (7.13) for $p = p(1)$ and $p = p(2)\cdots$ and $p = p(k)$. Thus, this whole cylinder in Ξ is missing from $S(v_1,\dots,v_k)$ and (7.12) follows.

This proves (7.7). As we already observed in (1.13), by Fubini's theorem and Proposition 1, it suffices for (7.6) to show that for every $\xi \in \Xi$,

$$\rho(\xi) = P\{\xi \in S_\infty\} = P\{\xi \text{ is seen on } \mathscr{G}\} = 1. \tag{7.14}$$

To prove (7.14), let $\varepsilon > 0$ be given and fix i_0 so that

$$\sum_{i=i_0}^{\infty}\{1 - [1-(1-2^{-\ell(i)})^{n_i}]^{\nu(i-1)}\} \le \varepsilon. \tag{7.15}$$

Then, for any fixed $\xi = (\xi_1, \xi_2, \dots)$ take $R = r_{i_0}$ and

$$\overline{\eta}_i = (\xi_{r_i+1-R}, \dots, \xi_{r_{i+1}-R}), \qquad i \ge i_0.$$

Then by (7.8) and (7.15),

$$P\{\text{no } K_i(\overline{\eta}_i) \text{ with } i \ge i_0 \text{ occurs}\} \ge 1 - \varepsilon. \tag{7.16}$$

However, when none of the $K_i(\overline{\eta}_i)$ with $i \ge i_0$ occur, then for each u_{s,i_0} there is a path π_{j,i_0} descending from this u_{s,i_0} along which the word $\overline{\eta}_{i_0} = (\xi_1, \dots, \xi_{\ell(i_0)})$ is seen. The last vertex of this path is some vertex in the r_{i_0+1}th generation, say, u_{s_1,i_0+1}. Since $K_{i_0+1}(\overline{\eta}_{i_0+1})$ does not occur, there is in turn a path π_{j,i_0+1} descending from this vertex along which $\overline{\eta}_{i_0+1} = (\xi_{\ell(i_0)+1}, \dots, \xi_{\ell(i_0)+\ell(i_0+1)})$ is seen, and so forth. Therefore, ξ is seen from each u_{s,i_0} in the r_{i_0}th generation. In particular, by (7.16),

$$P\{\xi \in S_\infty\} \ge 1 - \varepsilon.$$

This holds for each $\varepsilon > 0$, so that (7.14), and hence (7.6), follows. □

8. An example in which the random word percolates, but $\mathbb{1}$ does not. To conclude, we construct a graph $\mathscr{G}$ with the property of the title of this section, that is, a graph for which (1.19) holds. In fact, our graph is such that it rules out percolation of any word $\bar{\xi}$ which eventually omits a fixed finite string of letters. More specifically, let

$$\begin{aligned}\Xi_0 = \{\bar{\xi}\colon \exists\, m < \infty \text{ and } \eta_0, \dots, \eta_m \in \{0,1\}, \text{ such} \\ \text{that there exist only finitely many } \ell \text{ for} \\ \text{which } \bar{\xi}_{\ell+j} = \eta_j,\ 0 \le j \le m\}.\end{aligned} \tag{8.1}$$

Then our $\mathscr{G}$ will satisfy

$$P\{\mu(S_\infty) = 1, \text{ but } S_\infty \cap \Xi_0 = \varnothing\} = 1. \tag{8.2}$$

The graph $\mathscr{G}$ will be a sequence of complete graphs connected successively by strings of single edges. More precisely, let $\mathscr{V} = \{v(1), v(2), \dots\}$ be the vertex set of $\mathscr{G}$ [in this section $v(i)$ will no longer denote the ith component of v]. For two sequences of integers $\{N_m\}, \{M_m\}$ with

$$1 \le N_1 < M_1 < N_2 < M_2 \cdots, \tag{8.3}$$

$\mathscr{G}$ will have an edge between any pair $\{v(i'), v(i'')\}$ for which either

$$i' \ne i'' \quad \text{and} \quad N_m \le i', i'' \le M_m \quad \text{for some } m, \tag{8.4a}$$

or

$$i'' = i' \pm 1 \quad \text{and} \quad M_m \le i', i'' \le N_{m+1} \quad \text{for some } m. \tag{8.4b}$$

Thus, the part of $\mathscr{G}$ with vertices $v(i), N_m \le i \le M_m$, is a copy of the complete graph on

$$n_m := M_m - N_m + 1$$

vertices. The only connection between $v(M_m)$ and $v(N_{m+1})$ is the path $(v(M_m), v(M_m+1), \dots, v(N_{m+1}-1), v(N_{m+1}))$ (see Figure 4).

THEOREM 6. *If*

$$N_{m+1} - M_m = m \tag{8.5}$$

and

$$2^{-3m} n_m \to 1 \quad \text{as } m \to \infty, \tag{8.6}$$

then (8.2) *holds.*

PROOF. First we show that w.p.1, no word of Ξ_0 is seen on $\mathscr{G}$. This is easy because any infinite path $(v_1, v_2, \dots)$ on $\mathscr{G}$ must contain all the strings $v(M_m), v(M_{m+1}), \dots, v(N_{m+1})$ from some m on. In addition, each fixed word $\bar\eta = (\eta_0, \dots, \eta_k)$ of finite length occurs w.p.1 infinitely often on these strings, that is,

$$P\{X(v(M_m + j)) = \eta_j,\ 0 \le j \le k, \text{ occurs for infinitely many } m\} = 1. \tag{8.7}$$

If the event in braces in (8.7) occurs, then no word which contains $\bar\eta = (\eta_0, \dots, \eta_k)$ only finitely often is seen on $\mathscr{G}$. Since there are only countably many choices for $\bar\eta$, this shows that

$$P\{S_\infty \cap \Xi_0 = \varnothing\} = 1.$$

This proves part of (8.2). To show that almost all words are seen on $\mathscr{G}$, we consider the set of words, Ξ_1, in which the time of first occurrence of any finite word is not "too large." The precise definition of Ξ_1 is as follows. For any $\xi = (\xi_1, \xi_2, \dots) \in \Xi \setminus \Xi_0$, $\ell \ge 0$, and any finite word $\bar\eta = (\eta_0, \dots, \eta_m)$, let

$$\nu(\xi, \bar\eta, \ell) = \inf\{t > \ell\colon\ \xi_{t+j} = \eta_j,\ 0 \le j \le m\} - \ell. \tag{8.8}$$

Thus, ν is the number of steps one has to move forward from ℓ until $\bar\eta$ occurs in ξ. Now we take

$$\begin{aligned}\Xi_1 = \{\xi\colon\ &\exists\ m_1 < \infty \text{ such that for all } m \ge m_1 \text{ and}\\ &\text{all } \bar\eta \text{ of length } (m+1) \text{ and all } 0 \le \ell \le M_m \text{ it}\\ &\text{holds that } \nu(\xi, \bar\eta, \ell) \le (m+1)2^{2m}\}.\end{aligned} \tag{8.9}$$

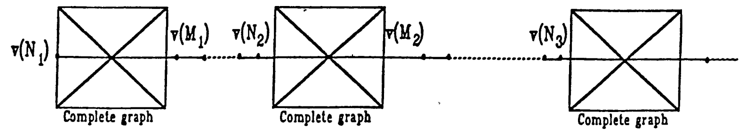

FIG. 4. *Illustration of* $\mathscr{G}$.

Our first task is to show that

$$\mu(\Xi_1) = 1. \tag{8.10}$$

To see this, note that for any fixed $\bar{\eta}$ of length $(m+1)$ and any ℓ,

$$\begin{aligned}&\mu\{\xi\colon \nu(\xi,\bar{\eta},\ell) > (m+1)2^{2m}\}\\ &\quad\le \mu\{\xi\colon (\xi_{\ell+1+r(m+1)+j})_{0\le j\le m} \ne \bar{\eta} \text{ for } 0 \le r < 2^{2m}\}\\ &\quad= (1-2^{-m-1})^{2^{2m}} \le \exp(-2^{2m}\cdot 2^{-m-1}) = \exp(-2^{m-1}).\end{aligned}$$

There are 2^{m+1} possible choices for $\bar{\eta}$ of length $(m+1)$ and (M_m+1) choices for $\ell \in [0, M_m]$. Thus

$$\begin{aligned}&\sum_m\sum_{\ell=0}^{M_m}\sum_{\substack{\bar{\eta}\text{ of length}\\(m+1)}} \mu\{\xi\colon \nu(\xi,\bar{\eta},\ell) > (m+1)2^{2m}\}\\ &\qquad\le \sum_m (M_m+1)2^{m+1}\exp(-2^{m-1}).\end{aligned} \tag{8.11}$$

However, by (8.5) and (8.6),

$$\begin{aligned}M_m &= M_1 + \sum_{k=2}^{m}(M_k - N_k) + \sum_{k=2}^{m}(N_k - M_{k-1})\\ &\le M_1 + \sum_{k=2}^{m} n_k + \sum_{k=2}^{m}(k-1) \sim \tfrac{8}{7}2^{3m},\end{aligned}$$

so that the right-hand side of (8.11) is finite. This establishes (8.10).

Next we define a subcollection of configurations which has full measure and in which all words of length at most $(m+1)2^{2m}$ are represented among the $X(v(i))$ with $N_m < i < M_m$ for some m. The precise definition is the following.

(8.12) $\Omega_1 = \{\omega\colon \exists\, m_2$ such that for all $m \ge m_2$ and for all choices of $\kappa(1);\dots,\kappa((m+1)2^{2m}) \in \{0,1\}$ there exist distinct $v(i_j)$ with $N_m < i_j < M_m$ and $X(v(i_j)) = \kappa(j)$, $1 \le j \le (m+1)2^{2m}\}$.

For a given m, there are

$$2^{(m+1)2^{2m}}$$

choices for $\{\kappa(j)\colon 1 \le j \le (m+1)2^{2m}\}$. The probability that for at least one of these choices one cannot find $v(i_j)$ with $N_m < i_j < M_m$ and $X(v(i_j)) = \kappa(j), 1 \le j \le (m+1)2^{2m}$, is at most

$$\begin{aligned}&2^{(m+1)2^{2m}}P\{\text{there are fewer than } (m+1)2^{2m} \text{ zeroes}\\ &\qquad\text{or fewer than } (m+1)2^{2m} \text{ ones among the}\\ &\qquad X(v(i)),\ N_m < i < M_m\}\\ &\quad\le 2^{(m+1)2^{2m}+1}P\{B(n_m-2,\tfrac{1}{2}) < (m+1)2^{2m}\},\end{aligned}$$

where $B(n, \frac{1}{2})$ is a binomial random variable corresponding to n trials with success probability $\frac{1}{2}$. Under (8.6), standard exponential bounds [see, for instance, Grimmett and Stirzaker (1992), Theorem 2.2.1] show that

$$\sum_m 2^{(m+1)2^{2m}+1} P\{B(n_m - 2, \tfrac{1}{2}) < (m+1)2^{2m}\} < \infty,$$

so that indeed

$$P\{\Omega_1\} = 1. \tag{8.13}$$

By virtue of (8.10) and (8.13) it now suffices for (8.2) to prove that in any configuration $\bar{\omega}$ from Ω_1, any word $\bar{\xi}$ from Ξ_1 is seen in $\bar{\omega}$. To this end, fix $\bar{\omega} \in \Omega_1$ and $\bar{\xi} \in \Xi_1$, and let $m_0 < \infty$ be such that for all $m \geq m_0$,

$$\nu(\bar{\xi}, \bar{\eta}, \ell) \leq (m+1)2^{2m} \quad \text{for any } \bar{\eta} = (\eta_0, \dots, \eta_m) \text{ and } 0 \leq \ell \leq M_m, \tag{8.14}$$

and

(8.15) there exists distinct $v(i_j)$ with $N_m < i_j < M_m$ and $X(v(i_j)) = \kappa(j), 1 \leq j \leq (m+1)2^{2m}$, for any choice of $(\kappa(1), \dots, \kappa((m+1)2^{2m})$.

Such an $m_0 < \infty$ exists by definition of Ξ_1 and Ω_1. To conclude, take $v = v(N_{m_0})$ and define

$$\eta^{(m)} = (X(v(M_m)), X(v(M_m + 1)), \dots, X(v(N_{m+1}))). \tag{8.16}$$

Note that this word depends only on $\bar{\omega}$ and that it has length $N_{m+1} - M_m + 1 = m + 1$ [by (8.5)]. We now choose the path $(v, v_1, v_2, \dots)$ such that $\bar{\xi}$ is seen from v along this path. By (8.14),

$$\rho_0 := \nu(\bar{\xi}, \eta^{(m_0)}, 0) \leq (m_0 + 1)2^{2m_0}$$

[because $\eta^{(m_0)}$ has length $(m_0 + 1)$]. By (8.15) we can therefore find vertices $v(i_j)$ with $N_{m_0} < i_j < M_{m_0}$ such that

$$X(v(i_j)) = \bar{\xi}(j), \qquad 1 \leq j \leq \rho_0 - 1. \tag{8.17}$$

All these vertices are adjacent to each other and to v in $\mathscr{G}$. We can therefore take $v_j = v(i_j)$, $1 \leq j < \rho_0$, and then take $v, v_1, \dots, v_{\rho_0 - 1}$ as the first ρ_0 vertices of our path. The next $(m_0 + 1)$ vertices will be $v(M_{m_0}), v(M_{m_0} + 1), \dots, v(N_{m_0+1})$. Note that $v(M_{m_0})$ is adjacent to $v_{\rho_0 - 1}$. Moreover, the X–values on these vertices have the required value, since

$$X(v(M_{m_0+j})) = \eta_j^{(m_0)} \text{ [see (8.16)]} = \bar{\xi}_{\rho_0 + j} \quad \text{[by our choice of } \rho_0 = \nu(\bar{\xi}, \eta^{(m_0)}, 0) \text{ and (8.8)]}. \tag{8.18}$$

Furthermore, the vertices $v, v_1, \dots, v_{\rho_0 - 1}, v_{\rho_0} = v(M_{m_0})$, $v_{\rho_0 + 1} = v(M_{m_0} + 1), \dots, v_{\rho_0 + m_0} = v(N_{m_0+1})$ are all distinct.

To see how this procedure can be iterated, assume that we already have found the path $(v, v_1, \dots, v_\ell)$ with $X(v_j) = \bar{\xi}_j$, $1 \le j \le \ell$, $v_\ell = v(N_{m+1})$ for some $m \ge m_0$ and

$$\{v, v_1, \dots, v_\ell\} \subset \{v(1), \dots, v(N_{m+1})\}, \tag{8.19}$$

so that, in particular,

$$\ell \le N_{m+1} + 1 \le M_{m+1}. \tag{8.20}$$

Then, because $\eta^{(m+1)}$ has length $(m+2)$, we have, by (8.14) with m replaced by $(m+1)$,

$$\nu(\bar{\xi}, \eta^{(m+1)}, \ell) \le (m+2)2^{2(m+1)}.$$

Then, by (8.15), we can choose $v(i_j)$ with

$$N_{m+1} < i_j < M_{m+1} \quad \text{and} \quad X(v(i_j)) = \bar{\xi}_{\ell+j} \qquad \text{for } 1 \le j < \rho_{m+1-m_0} := \nu(\bar{\xi}, \eta^{(m+1)}, \ell). \tag{8.21}$$

We then extend our path $(v, v_1, \dots, v_\ell)$ by the $v(i_j)$, $1 \le j < \rho_{m+1-m_0}$, followed by the $(m+2)$ vertices $v(M_{m+1}), v(M_{m+1}+1), \dots, v(N_{m+2})$. We then are again in the situation (8.19), (8.20) with m replaced by $(m+1)$ and ℓ by $\ell' = \ell + \rho_{m+1-m_0} + m + 1$. Moreover, we now have $X(v_j) = \bar{\xi}_j$ for $j \le \ell'$ for the same reasons as in (8.17), (8.18). Indeed, by (8.21),

$$X(v_{\ell+j}) = X(v(i_j)) = \bar{\xi}_{\ell+j} \quad \text{for } 1 \le j < \rho_{m+1-m_0},$$

and by the definition (8.8) of $\rho_{m+1-m_0} = \nu(\bar{\xi}, \eta^{(m+1)}, \ell)$,

$$X(v_{\ell+\rho_{m+1-m_0}+j}) = X(v(M_{m+1}+j)) = \eta_j^{(m+1)} = \bar{\xi}_{\nu(\bar{\xi},\eta^{(m+1)},\ell)+\ell+j}, \quad 0 \le j \le m+1.$$

Thus we can repeat the procedure and obtain a path $(v, v_1, v_2, \dots)$ with $X(v_j) = \bar{\xi}_j$ for all $j \ge 1$. □

Acknowledgments. The authors are grateful for helpful conversations on Section 2 with S. Kalikow and for substantial help on Section 3 from R. Durrett and G. Grimmett.

REFERENCES

BREIMAN, L. (1968). *Probability*. Addison–Wesley, Reading, MA.

CAMPANINO, M. and RUSSO, L. (1985). An upper bound for the critical probability for the three-dimensional cubic lattice. *Ann. Probab.* **13** 478–491.

COX, J. T. and DURRETT, R. (1983). Oriented percolation in dimension $d \ge 4$: bounds and asymptotic formulas. *Math. Proc. Cambridge Philos. Soc.* **93** 151–162.

DEKKING, F. M. (1989). On the probability of occurrence of labelled subtrees of a randomly labelled tree. *Theoret. Comput. Sci.* **65** 149–152.

DEKKING, F. M. (1991). Branching processes that grow faster than binary splitting. *Amer. Math. Monthly* **98** 728–731.

DEKKING, F. M. and PAKES, A. G. (1991). On family trees and subtrees of simple branching processes. *J. Theoret. Probab.* **4** 353–369.

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth & Brooks/Cole, Belmont, CA.

EVANS, S. N. (1992). Polar and nonpolar sets for a tree indexed process. *Ann. Probab.* **20** 579–590.

GRIMMETT, G. R. and STIRZAKER, D. R. (1992). *Probability and Random Processes*, 2nd ed. Oxford Univ. Press.

HAMMERSLEY, J. M. (1961). The number of polygons on a lattice. *Proc. Cambridge Philos. Soc.* **57** 516–523.

HARRIS, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56** 13–20.

HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, Berlin.

JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley, New York.

KESTEN, H. (1982). *Percolation Theory for Mathematicians*. Birkhäuser, Boston.

LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, Berlin.

LYONS, R. (1992). Random walks, capacity and percolation on trees. *Ann. Probab.* **20** 2043–2088.

MADRAS, N. and SLADE, G. (1993). *The Self-Avoiding Walk*. Birkhäuser, Boston.

MAI, T. and HALLEY, J. W. (1980). AB percolation on a triangular lattice. In *Ordering in Two Dimensions* (S. K. Sinha, ed.) 369–371. North–Holland, Amsterdam.

MALLOWS, C. L. (1968). An inequality involving multinomial probabilities. *Biometrika* **55** 422–424.

MENSHIKOV, M. V. and ZUYEV, S. A. (1992). Models of ρ-percolation. In *Petrozavodsk Conference on Probabilistic Methods in Discrete Mathematics*.

RÉNYI, A. (1970). *Probability Theory*. North–Holland, Amsterdam.

RUDIN, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw–Hill, New York.

WIERMAN, J. C. (1989). AB percolation: a brief survey. In *Combinatorics and Graph Theory* **25** 241–251. Banach Center Publications.

WIERMAN, J. C. and APPEL, M. S. (1987). Infinite AB percolation clusters exist on the triangular lattice. *J. Phys. A* **20** 2533–2537.

MATHEMATICAL SCIENCES INSTITUTE
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853