Percolation of arbitrary words in one dimension

Geoffrey R. Grimmett, Thomas M. Liggett and Thomas Richthammer

University of Cambridge and University of California, Los Angeles

July 8, 2009

Abstract

We consider a type of long-range percolation problem on the positive integers, motivated by earlier work of others on the appearance of (in)finite words within a site percolation model. The main issue is whether a given infinite binary word appears within an iid Bernoulli sequence at locations that satisfy certain constraints. We settle the issue in some cases, and provide partial results in others.

1 Introduction

Let $W = (w_1, w_2, \dots) \in \{0, 1\}^{\mathbb{N}}$ be an infinite binary word, and $X = (X_1, X_2, \dots)$ and $Y = (Y_1, Y_2, \dots)$ be independent Bernoulli sequences (i.e., iid sequences of Bernoulli random variables) with parameters $p_X = P(X_i = 1), p_Y = P(Y_j = 1) \in (0, 1)$. Let M be a positive integer. An *admissible (M-)embedding* of W in Y is a sequence $(m_i : i \geq 1)$ of integers such that $Y_{m_i} = w_i$ and $1 \leq m_i - m_{i-1} \leq M$ for each $i \geq 1$. (By default, we take $m_0 = 0$.) We say that W *is M-seen in* Y if there exists an M-admissible embedding of W in Y. In this paper, we ask whether or not the events $\{W$ is M-seen in $Y\}$ and $\{X$ is M-seen in $Y\}$ can have strictly positive probability.

This question is motivated by analogous questions considered in [2, 6, 7] concerning percolation of words on graphs such as Z^d. These questions were partially answered for large d in [2], and on a modified version of Z^2 in [7]. A version of the above question has been answered in the affirmative in [5, 8] for large M and $d = 2$. Our problem may be set in the context of long-range percolation, through a consideration of the oriented

graph with vertices $\{1, 2, \dots\}$ in which there is an edge from i to j if $1 \leq j - i \leq M$. In this setting, our problem corresponds to ordinary site percolation when W is the constant word $w_i \equiv 1$, and to so-called AB site percolation when W is the alternating word with $w_{2i} \equiv 0, w_{2i-1} \equiv 1$ (or *vice versa*).

The following formulation of the problem appears to be similar in spirit to the famous problem of the clairvoyant demon posed by Peter Winkler. As above, let X and Y be independent Bernoulli sequences with parameters p_X, p_Y; for simplicity we assume $p_X = p_Y = \frac{1}{2}$. We color the point (i, j) of the first orthant of the square lattice Z^2 *red* if $X_i = Y_j$. Let the origin $(0, 0)$ be red also. For $M \geq 1$, we define an M-admissible path to be an infinite sequence $m = (m_0, m_1, m_2, \dots)$ satisfying $m_0 = 0$ and $1 \leq m_{i+1} - m_i \leq M$ for all i, such that every point (i, m_i), $i \geq 0$, is red. There exists an M-admissible path if and only if X is M-seen in Y. See [5] for an account of the clairvoyant demon problem, and [1, 9] for solutions to a related problem in which the admissible paths for that problem are undirected.

Since Y contains arbitrarily long sequences of 0's and arbitrarily long sequences of 1's, it is easy to see that $P(W \text{ is } M\text{-seen in } Y) = 0$ when W is periodic. However, the situation for general words is not so clear. In Section 5, we will show that the truth of the statements "for every M, $P(X \text{ is } M\text{-seen in } Y) = 0$" and "for every M, $P(W \text{ is } M\text{-seen in } Y) = 0$" (for an arbitrary infinite word W) is independent of the parameters p_X, p_Y of the Bernoulli sequences X, Y; see Theorem 13. Therefore, except in that section, we will assume that

$$p_X = p_Y = \tfrac{1}{2}.$$

In order to gain some insight into our problem, we consider the probability of M-seeing finite words W, and particularly how this probability depends on W. Let

$$\alpha = 1 - 2^{-M}, \quad \beta = 2^{-M}.$$

It is easy to check (as we will do in Section 3) that the probability of M-seeing a given word of length n is minimized by the constant word $W = (1, 1, \dots, 1)$ of length n, and that in this case, this probability equals α^n. In the other direction, we consider the alternating word $A_n = (1, 0, 1, 0, \dots)$ of length n; we could equally consider the alternating word beginning with 0. Since the infinite alternating word is periodic, the probability

$$v_n = P(A_n \text{ is } M\text{-seen in } Y) \tag{1}$$

tends to zero as $n \to \infty$. In Section 2, we will show how to compute v_n exactly, and hence determine the exponential rate at which this probability tends to zero. If $M = 2$, for example, $v_n \sim c(0.85 \dots)^n$. We will prove that the alternating word is most likely to be seen in two cases:

Theorem 1. (a) *Let $M = 2$ and $n \geq 1$. For any word W of length n,*

$$P(W \text{ is } M\text{-seen in } Y) \leq v_n.$$

2

(b) *Let $M \geq 2$, and let $W_{p,q}$ be the word $(1,1,1,\ldots,0,0,0)$ comprising p 1's followed by q 0's. Then*

$$P(W_{p,q} \text{ is } M\text{-seen in } Y) \leq v_{p+q}.$$

The first part of this theorem will be proved in Section 2, and the second in Section 4. As a consequence of Theorem 1(a), we have the following solution to our main problem in case $M = 2$:

Corollary 2. *If $M = 2$, $P(W \text{ is } M\text{-seen in } Y) = 0$ for every infinite word W, and thus $P(X \text{ is } M\text{-seen in } Y) = 0$.*

A useful tool in our analysis is the following sequence of 'spacing' random variables. Given a finite or infinite word W, define $T_0 = 0$ and, recursively,

$$T_{k+1} = \min\{i > T_k : Y_i = w_{k+1}\}, \quad \tau_{k+1} = T_{k+1} - T_k.$$

Note that, while the values of $\tau_1, \tau_2, \ldots$ depend on the choice of W, for any W they are iid random variables with the geometric distribution with parameter $\frac{1}{2}$. The values of $\tau_1, \tau_2, \ldots, \tau_n$ do not in general determine whether or not the word $(w_1, w_2, \ldots, w_n)$ is M-seen. (An example illustrating this is given in Section 3.) However, they do so for the constant and alternating words.

Theorem 3. (a) *The constant word of length n is M-seen in Y if and only if $\tau_k \leq M$ for all $1 \leq k \leq n$.*
(b) *The alternating word A_n of length n is M-seen in Y if and only if*

$$T_k \leq kM \text{ for all } 1 \leq k \leq n \quad \text{and} \quad T_k - T_j < (k - j + 1)M \text{ for all } 0 \leq j < k \leq n.$$

This theorem will be proved in Section 3.

Let W be a word of length n, and let $N_n = N_n(W)$ be the number of M-admissible embeddings of W in Y. It is easy to see that

$$E(N_n) = (M/2)^n.$$

The second moment of N_n can be expressed in the following way. Let $J = (J_0. J_1, J_2, \ldots)$ and $K = (K_0, K_1, K_2, \ldots)$ be independent random walks on Z starting at $J_0 = K_0 = 0$ with, as step-size distribution, the uniform distribution on the finite set $\{1, 2, \ldots, M\}$.

Theorem 4. (a) *For any word W of length n, and any $M \geq 1$,*

$$E(N_n^2) = E(N_n)^2 E\left(\prod_{(r,s):J_r = K_s} 2 \cdot 1(w_r = w_s)\right), \tag{2}$$

where the product is over $r, s \in \{1, 2, \ldots, n\}$, and $1(A)$ denotes the indicator function of A.

3

(b) *Let X be the random word of length n, comprising random letters with the Bernoulli $\left(\frac{1}{2}\right)$ distribution. Then*

$$E(N_n(X)^2) = E(N_n(X))^2 E(2^{Z_n}),$$

where Z_n is the number of visits to zero between times 1 and n made by the random walk $J - K$.

This theorem will be proved in Section 6. We will also see there that $E(2^{Z_n})$ is asymptotic to a constant multiple of c_M^n for some $c_M > 1$. In the case $M = 2$, we have $c_2 = \frac{4}{3}$. By part (a) of this theorem, the constant word maximizes the variance of $N_n(W)$.

2 Recursions for v_n

Let $A_n = (1, 0, 1, 0, \dots)$ be the alternating word of length n starting with 1. (By symmetry, probabilities for alternating words starting with 0 are the same as for A_n.) In this section, we first compute the v_n given in (1), and then we prove Theorem 1(a). The computation of v_n is done recursively; the recursions satisfied by v_n will be used in Section 4 in the proof of Theorem 1(b).

If an admissible embedding of A_n in Y exists, we define the *standard embedding* to be that whose sequence of positions $(m_i : i \geq 1)$ is earliest in the usual lexicographic order. We will use the following notation:

- $v_{n,k} :=$ the probability that A_n possesses an admissible embedding, and its standard embedding starts at position k.

- $v_n := \sum_{k=1}^{M} v_{n,k} =$ the probability that A_n possesses an admissible embedding.

- $v_n' := v_{n,M}$.

Proposition 5. *The sequences v_n and v_n' satisfy the following recursions:*

$$v_n = \alpha v_{n-1} + (\alpha - M\beta)v_{n-1}' \quad and \quad v_n' = \beta v_{n-1} + (M-1)\beta v_{n-1}' \tag{3}$$

for $n \geq 1$, with initial conditions $v_0 = 1$ and $v_0' = 0$, and

$$v_{n+1} = (\alpha + (M-1)\beta)v_n - \beta(M - 2\alpha)v_{n-1} \tag{4}$$

for $n \geq 1$, with initial conditions $v_0 = 1$ and $v_1 = \alpha$.

Note that in (3), unlike (4), all the coefficients are nonnegative. This will enable us to compare solutions to recursive inequalities.

It is easy to solve the recursion (4) explicitly. The characteristic polynomial is

$$f(\lambda) = \lambda^2 - (\alpha + (M-1)\beta)\lambda + \beta(M - 2\alpha).$$

4

Note that, for $M \geq 2$,

$$f(0) = \beta(M - 2\alpha) > 0, \quad f(M\beta) = 2\beta(M\beta - \alpha) < 0,$$
$$f(\alpha) = \beta(M\beta - \alpha) < 0, \quad f(1) = 2\beta^2 > 0.$$

Therefore, the two roots of f lie in the disjoint intervals $(0, M\beta)$ and $(\alpha, 1)$, respectively. This implies that $v_n \to 0$ exponentially fast as $n \to \infty$, and is a quantification of the observation that Y contains no admissible embedding of A_n if there appears, sufficiently early in Y, a consecutive subsequence of $2M$ letters all of which are 0 (respectively 1). A word of caution: while v_n tends to zero exponentially rapidly, the convergence can still in a sense be quite slow. The larger root of f is of the order $1 - 2\beta^2$, and, for example, is $.9978\ldots$ when $M = 5$.

Proof. A Bernoulli sequence Y contributing to $v_{n,k}$ must satisfy $Y_k = 1$, and the preceding sequence $(Y_1, \ldots, Y_{k-1})$ cannot contain 1 and 0 in that order, since if it did, there would be an admissible embedding starting before position k. Therefore, $(Y_1, \ldots, Y_{k-1})$ must be of the form $(0, \ldots, 0, 1, \ldots, 1)$. We distinguish two cases for the starting sequence: all 0's (Case 1) and at least one 1 (Case 2). In Case 1, the only condition on $(Y_{k+1}, \ldots)$ is that it must contain an admissible embedding of the remainder of A_n. In Case 2, $(Y_{k+1}, \ldots)$ must contain an admissible embedding of the remainder whose standard embedding starts at (relative) position M. (If it started earlier, then the 1 at position $k - 1$ would initiate an earlier embedding.) This yields the recursion

$$v_{n,k} = \frac{1}{2^k} v_{n-1} + \frac{k-1}{2^k} v_{n-1,M} \quad \text{for } 1 \leq k \leq M, \ n \geq 1, \tag{5}$$

with initial condition $v_{1,k} = 2^{-k}$ for $1 \leq k \leq M$.

The first relation in (3) is obtained by summing (5) over k, and the second by setting $k = M$. Finally, one can eliminate v'_n from (3) to obtain (4). $\qquad\square$

Proof of Theorem 1(a). Let $M = 2$ and let $W \in \{0,1\}^n$ be a word of length n. For $1 \leq m \leq n$, let W_m be the word comprising the last m digits of W. We will use notation similar to that at the beginning of this section.

- $w_{m,k} :=$ the probability that W_m possesses an admissible embedding, and its standard embedding starts at position $k \in \{1, 2\}$.

- $w_m := w_{m,1} + w_{m,2} =$ the probability that W_m has an admissible embedding.

- $w'_m := w_{m,2}$

We follow the same procedure as we did for the alternating word A_n. Unlike that case, we shall obtain only a recursive estimate from above.

Denote the first digit of W_m by $a \in \{0, 1\}$, and let $b = 1 - a$ be the complementary digit. A Bernoulli sequence Y contributing to $w_{m,1}$ must contain a at position 1 and the word W_{m-1} following. This gives

$$w_{m,1} = \tfrac{1}{2} w_{m-1}. \tag{6}$$

If Y contributes to $w_{m,2}$, there are two cases. If Y starts with a, then the second digit must also be a, and subsequently the Y must contain W_{m-1}, but not starting at the next digit. (Otherwise there would be a standard embedding starting at position 1.) If Y starts with b, then the second digit has to be a, and subsequently Y must contain W_{m-1}. This gives

$$w_{m,2} \leq \tfrac{1}{4}w_{m-1,2} + \tfrac{1}{4}w_{m-1}. \tag{7}$$

Note that equality need not hold in (7). Suppose that both Y and W_m begin with the letters aa, and that $(Y_3, Y_4, \dots)$ contains an admissible embedding of both W_{m-2} and W_{m-1}. In this case, Y does not contribute to the left side of (7) but it does to the first term on the right side.

From (6)–(7), we deduce the recursive inequalities

$$w_m \leq \tfrac{3}{4}w_{m-1} + \tfrac{1}{4}w'_{m-1} \quad \text{and} \quad w'_m \leq \tfrac{1}{4}w_{m-1} + \tfrac{1}{4}w'_{m-1}, \quad \text{where} \quad w_1 = \tfrac{3}{4}, w'_1 = \tfrac{1}{4}.$$

By comparison with the recursion formula (3) with $M = 2$,

$$v_m = \tfrac{3}{4}v_{m-1} + \tfrac{1}{4}v'_{m-1} \quad \text{and} \quad v'_m = \tfrac{1}{4}v_{m-1} + \tfrac{1}{4}v'_{m-1}, \quad \text{where} \quad v_1 = \tfrac{3}{4}, \; v'_1 = \tfrac{1}{4}, \tag{8}$$

we obtain by induction on m and the positivity of the coefficients in (8) that $w_m \leq v_m$ for $1 \leq m \leq n$. In particular, $w_n \leq v_n$ as claimed. $\qquad\square$

Remark. The above method does not work for $M \geq 3$, since in this case the coefficients of the recursive inequalities for w_n, w'_n do not match the coefficients of the recursion for v_n, v'_n.

3 Relations to the spacing random variables

In this section, we prove several results relating M-seen finite words to inequalities satisfied by the spacing variables τ_k and their partial sums. When the word is either constant or alternating, these are equivalences, and were stated as Theorem 3 in the Introduction. For general words, we only have one direction — if the word is seen, then the spacing variables satisfy certain inequalities. We will say that $W = (w_1, w_2, \dots)$ is seen at $(m_1, m_2, \dots)$ if $m_1 < m_2 < \cdots$ and $Y_{m_i} = w_i$ for each i.

First, we give an example to show that seeing a word W of length n is not in general determined by the values of $\tau_1, \dots, \tau_n$. Suppose $n = 4$, $M = 2$, $W = (1, 1, 0, 0)$, and the Bernoulli sequence starts with $110110\cdots$. Then $\tau_1 = 1$, $\tau_2 = 1$, $\tau_3 = 1$, $\tau_4 = 3$. If W is to be seen, then it must be seen at locations $m_1 = 2$, $m_2 = 4$, $m_3 = 6$, and $m_4 = 7$ or 8. Thus, it is seen if and only if one of the next two digits in the Bernoulli sequence is a 0, but this cannot be determined from the first four τ_k.

Proposition 6. *Let $W = (w_1, \dots, w_n)$. If W is seen at $(m_1, \dots, m_n)$, then $T_k \leq m_k$ for $1 \leq k \leq n$. In particular, if W is M-seen, then $T_k \leq kM$ for all $1 \leq k \leq n$.*

Proof. Let $m_0 := T_0 = 0$. We will prove $T_k \leq m_k$ by induction on k. For the induction step we assume $T_k \leq m_k$. Let $a = w_{k+1}$. By definition T_{k+1} is the first location of an a after location T_k, and m_{k+1} is some location of an a after location $m_k \geq T_k$, which immediately implies $T_{k+1} \leq m_{k+1}$. Finally, if W is M-seen, then

$$T_k \leq m_k = \sum_{i=1}^{k} (m_i - m_{i-1}) \leq kM.$$

$\square$

The next result implies that the probability of M-seeing a word of length n is minimized by the constant word, and in that case, this probability is α^n.

Proposition 7. (a) *If W is a word of length n and $\tau_1 \leq M$, ..., $\tau_n \leq M$, then W is M-seen.*
(b) *If a constant word of length n is M-seen, then $\tau_1 \leq M$, ..., $\tau_n \leq M$.*

Proof. For part (a), note that if $\tau_i \leq M$ for each $i \leq n$, then W is M-seen at $(T_1, \ldots, T_n)$. For part (b), suppose that the constant word $(a, a, \ldots, a)$ of length n is seen at $(m_1, \ldots, m_n)$ where $1 \leq m_i - m_{i-1} \leq M$ for each i, i.e. up to location m_n there is no block of M consecutive non a's. As $T_n \leq m_n$ by Proposition 6, this implies $\tau_i \leq M$ for all $i \leq n$. $\square$

Proposition 8. *Let $W = (w_1, \ldots, w_n)$. Suppose that W is seen at $(m_1, \ldots, m_n)$ and that $0 < m_{i+1} - m_i \leq M$ for each i. If $w_k \neq w_{k+1} = \cdots = w_l$ for some $k+1 \leq l$ and if $\tau_l > M$, then $T_l \leq m_{k+1}$.*

Proof. Let $a = w_l$. In between locations T_{l-1} and T_l there is a block of at least M consecutive non a's. As there have to be a's at locations $m_{k+1} < \cdots < m_l$ and we have $m_{i+1} - m_i \leq M$, this block has to be before location m_{k+1} or after location m_l, i.e. $T_l \leq m_{k+1}$ or $m_l \leq T_{l-1}$. By Proposition 6 we have $T_l \leq m_l$, so the second alternative is not possible. $\square$

Proposition 9. *Let A_n be an alternating word of length n. Then A_n is M-seen if and only if*

$$T_k \leq kM \text{ for all } 1 \leq k \leq n \quad \text{and} \quad T_k - T_j < (k-j+1)M \text{ for all } 0 \leq j < k \leq n. \quad (9)$$

Proof. A special property of an alternating word is that

$$Y_i = w_k \text{ for } T_k \leq i < T_{k+1}. \quad (10)$$

Define

$$S_k = \min\{T_{k+1} - 1, S_{k-1} + M\}, \quad S_0 = 0, \quad \sigma_k = S_k - S_{k-1}.$$

Note that $S_k < T_{k+1}$ and $\sigma_k \leq M$. Since T_k is strictly increasing in k, we see inductively that S_k is strictly increasing in k also. Therefore $\sigma_k \geq 1$ for all $k \geq 1$. Consider the statement

$$T_k \leq S_k \text{ for all } 1 \leq k \leq n. \tag{11}$$

We will prove the following implications:

$$(11) \Rightarrow A_n \text{ is } M\text{-seen} \Rightarrow (9) \Rightarrow (11).$$

First suppose that (11) holds. Then $T_k \leq S_k < T_{k+1}$, so that $Y_{S_k} = w_k$ by (10). Since $\sigma_k \leq M$, it follows that A_n is M-seen, since it is seen at $(S_1, \ldots, S_n)$.

Next assume that A_n is seen at $(m_1, \ldots, m_n)$ where $m_i - m_{i-1} \leq M$ for all i. The first part of (9) follows from Proposition 6. To prove the second part, let $0 \leq j < k \leq n$. As $(Y_i : T_j \leq i < T_k)$ consists of $k-j$ constant blocks, the interval $[T_j, T_k)$ contains at most $k-j$ consecutive elements from $m_1, \ldots, m_n$, i.e. $m_l < T_j \leq m_{l+1} \leq m_{l+r} < T_k \leq m_{l+r+1}$ for some l and $r \leq k-j$. So $T_k - T_j < m_{l+r+1} - m_l = \sum_{i=l}^{l+r}(m_{i+1} - m_i) \leq (r+1)M$.

Finally, assume that (9) holds. To prove (11), we will prove the statement

$$T_k \leq S_i + (k-i)M \tag{12}$$

by induction on i (for fixed k). When $i = 0$, (12) becomes $T_k \leq kM$, which is part of assumption (9). When $i = k$, (12) is $T_k \leq S_k$, which is the desired conclusion in (11). For the induction step, suppose (12) holds for i with $0 \leq i < k$. To prove it for $i+1$, we need to check that

$$T_k \leq \min\{T_{i+2} - 1, S_i + M\} + (k-i-1)M. \tag{13}$$

The fact that $T_k \leq T_{i+2} - 1 + (k-i-1)M$ follows from (9) whose second assertion holds trivially for $j = k, k+1$, while $T_k \leq S_i + (k-i)M$ is just the induction hypothesis. This proves (13). $\qquad\square$

4 Two-block words

We prove Theorem 1(b) in this section. For $p, q, j \geq 0$, define

$$\sigma_{p,j} = P(\tau_1 \leq M, \ldots, \tau_p \leq M, T_{p+j} > pM),$$
$$\sigma'_{p,j} = P(\tau_1 \leq M, \ldots, \tau_p \leq M, T_{p+j} \leq pM)$$

and

$$u_{p,q} = \alpha^{p+q} + \beta \sum_{j=1}^{q} \alpha^{q-j} \sigma'_{p,j} = \alpha^p - \beta \sum_{j=1}^{q} \alpha^{q-j} \sigma_{p,j}.$$

Here we have used $\sigma_{p,j} + \sigma'_{p,j} = \alpha^p$. Note that $u_{p,0} = \alpha^p$ and $u_{0,q} = \alpha^q$. The next result will allow us to compute $\sigma_{p,j}$ fairly explicitly.

Lemma 10. *For $p, j \geq 0$ and arbitrary l,*

$$P(\tau_1 \leq M, \ldots, \tau_p \leq M, T_{p+j} > lM) = \sum_{i=0}^{p} \binom{p}{i}(-\beta)^i P(T_{p+j} > (l-i)M).$$

Proof. Use the fact that for any geometric random variable τ, the conditional distribution of $\tau - M$ given $\tau > M$ is the same as the distribution of τ, to write

$$P(\tau_1 \leq M, \ldots, \tau_m \leq M, T_k > lM) - P(\tau_1 \leq M, \ldots, \tau_{m+1} \leq M, T_k > lM)$$
$$= P(\tau_1 \leq M, \ldots, \tau_m \leq M, \tau_{m+1} > M, T_k > lM)$$
$$= \beta P(\tau_1 \leq M, \ldots, \tau_m \leq M, T_k > (l-1)M)$$

for any $k > m$. Now use induction on m, together with the relation $\binom{m}{i} + \binom{m}{i-1} = \binom{m+1}{i}$. $\square$

Lemma 11. *For $p, q \geq 0$, $P(W_{p,q}$ is M-seen$) \leq u_{p,q}$.*

Proof. If $W_{p,q}$ is seen at $(m_1, \ldots, m_{p+q})$ where $m_{i+1} - m_i \leq M$ for all i, then $\tau_1 \leq M, \ldots, \tau_p \leq M$ by Proposition 7(b). Furthermore, if $1 \leq j \leq q$ and $\tau_{p+j} > M$, then $T_{p+j} \leq m_{p+1} \leq (p+1)M$ by Proposition 8. Considering the largest $j \geq 1$ (if any) for which $\tau_{p+j} > M$, we see that

$$P(W_{p,q} \text{ is seen}) \leq P(\tau_1 \leq M, \ldots, \tau_{p+q} \leq M) + \sum_{j=1}^{q} \tilde{\sigma}_{p,j},$$

where

$$\tilde{\sigma}_{p,j} = P(\tau_1 \leq M, \ldots, \tau_p \leq M, \tau_{p+j} > M, \tau_{p+j+1} \leq M, \ldots, \tau_{p+q} \leq M, T_{p+j} \leq (p+1)M).$$

Using the same trick as in the proof of the previous lemma we obtain

$$\tilde{\sigma}_{p,j} = \alpha^{q-j} P(\tau_1 \leq M, \ldots, \tau_p \leq M, \tau_{p+j} > M, T_{p+j} \leq (p+1)M) = \alpha^{q-j} \beta \sigma'_{p,j},$$

so the result follows by definition of $u_{p,q}$. $\square$

For any function $f_{p,q}$, $p, q \geq 0$, define a generalized mixed second derivative by

$$\Delta f_{p,q} = f_{p+1,q+1} - M\beta f_{p,q+1} - (\alpha - \beta)f_{p+1,q} + \beta(M - 2\alpha)f_{p,q}.$$

This particular choice of coefficients is designed to correspond to the coefficients in (4). In order to do so, the middle coefficients would have to sum to $-(\alpha + (M - 1)\beta)$. This particular decomposition was chosen by computing numerically $\Delta u_{p,q}$ for various values of the parameters, and checking to see which one made this expression ≤ 0.

Lemma 12. $\Delta u_{p,q} \leq 0$ *for $p, q \geq 0$.*

Proof. If $f_{p,q} = \alpha^p$, then

$$\Delta f_{p,q} = \alpha^{p+1} - M\beta\alpha^p - (\alpha - \beta)\alpha^{p+1} + \beta(M - 2\alpha)\alpha^p = 0.$$

Therefore, if we let

$$w_{p,q} = \sum_{j=1}^{q} \alpha^{q-j}\sigma_{p,j},$$

we have $\Delta u_{p,q} = -\beta\Delta w_{p,q}$, so we need to show that $\Delta w_{p,q} \geq 0$. We will do so by computing the generating function of this expression as a function of q.

By Lemma 10,

$$\sigma_{p,j} = \sum_{i=0}^{p} \binom{p}{i}(-\beta)^{p-i}P(T_{p+j} > iM).$$

Since T_m is a sum of m independent geometric random variables and can thus be interpreted as the waiting time for the m-th success in a sequence of Bernoulli experiments, and since S_k, the number of successes in the first k of these experiments is binomially distributed, we have

$$P(T_j > k) = P(S_k < j) = \frac{1}{2^k}\sum_{l=0}^{j-1}\binom{k}{l}.$$

Therefore, since $2^{-M} = \beta$,

$$\sigma_{p,j} = \sum_{i=0}^{p}\binom{p}{i}(-\beta)^{p-i}\beta^i\sum_{l=0}^{p+j-1}\binom{iM}{l} = \beta^p\sum_{i=0}^{p}\sum_{l=0}^{p+j-1}(-1)^{p-i}\binom{p}{i}\binom{iM}{l}.$$

It follows that for $0 < x < 1$,

$$(1-x)\sum_{j=1}^{\infty}\sigma_{p,j}x^{j-1} = \beta^p\sum_{i=0}^{p}\sum_{l=0}^{\infty}\binom{p}{i}\binom{iM}{l}(-1)^{p-i}x^{(l-p)^+}. \tag{14}$$

Note that for every polynomial function g of degree $\deg g < p$ we have

$$\sum_{i=0}^{p}\binom{p}{i}g(i)(-1)^{p-i} = 0. \tag{15}$$

It suffices to check this for all polynomial functions of the form $g(i) = \binom{i}{l}$, $0 \leq l < p$:

$$\sum_{i=0}^{p}\binom{p}{i}\binom{i}{l}(-1)^{p-i} = \binom{p}{l}\sum_{i=l}^{p}\binom{p-l}{p-i}(-1)^{p-i} = \binom{p}{l}(1-1)^{p-l} = 0.$$

Applying (15) to the functions $g(i) = \binom{iM}{l}$, $0 \le l < p$, it follows that the positive part at the end of (14) is not needed:

$$(1-x)\sum_{j=1}^{\infty} \sigma_{p,j} x^{j-1} = \beta^p \sum_{i=0}^{p} \sum_{l=0}^{\infty} \binom{p}{i}\binom{iM}{l}(-1)^{p-i} x^{l-p}$$

$$= \beta^p x^{-p} \sum_{i=0}^{p} \binom{p}{i}(-1)^{p-i}(1+x)^{iM}$$

$$= \beta^p x^{-p}[(1+x)^M - 1]^p.$$

Therefore,

$$(1-x)\sum_{q=1}^{\infty} w_{p,q} x^{q-1} = \frac{1-x}{1-\alpha x}\sum_{j=1}^{\infty}\sigma_{p,j}x^{j-1} = \frac{1}{1-\alpha x}\beta^p x^{-p}[(1+x)^M - 1]^p.$$

Using this expression, we can write

$$(1-x)\sum_{q=0}^{\infty}\Delta w_{p,q} x^{q-1} = \frac{\beta^{p+1} x^{-p}}{x^2(1-\alpha x)}[(1+x)^M - 1]^p P(x),$$

where

$$P(x) = (1+x)^M[1-(\alpha-\beta)x] - 1 - (M+\beta-\alpha)x + x^2(M-2\alpha).$$

Note that $P(1) = 2^M(1-\alpha+\beta) - 2 = 0$, so we may define a polynomial Q by $P(x) = (1-x)Q(x)$. If Q has nonnegative coefficients, it will follow that $\Delta w_{p,q} \ge 0$ for all $p, q \ge 0$ as required. The coefficients of Q are the partial sums of the coefficients of P. The constant and linear terms in P vanish. The coefficient of x^2 is $\binom{M}{2} - 2(\alpha - M\beta)$, while for $k \ge 3$, the coefficient of x^k is $\binom{M}{k} - (\alpha-\beta)\binom{M}{k-1}$. Therefore, we need to check that the following expression is nonnegative for $2 \le l \le M$:

$$\binom{M}{2} - 2(\alpha - M\beta) + \sum_{k=3}^{l}\left[\binom{M}{k} - (\alpha-\beta)\binom{M}{k-1}\right]$$

$$= \binom{M}{l} + 2\beta\sum_{k=2}^{l-1}\binom{M}{k} - 2(\alpha - M\beta) = \binom{M}{l} - 2\beta\sum_{k=l}^{M}\binom{M}{k}.$$

This is nonnegative for $2 \le l < M$ since $\beta\sum_{k=0}^{M}\binom{M}{k} = 1$, and for $l = M$ since then the right side above is $1 - 2\beta$. It equals 0 when $l > M$. $\square$

Proof of Theorem 1(b). Let

$$\delta_{p,q} = v_{p+q} - u_{p,q}, \quad p, q \ge 0.$$

By (3) we have $v_{n+1} \geq \alpha v_n$ and thus $v_n \geq \alpha^n$, so

$$\delta_{p,0} = v_p - \alpha^p \geq 0 \quad \text{and} \quad \delta_{0,q} = v_q - \alpha^q \geq 0 \quad \text{for all } p, q \geq 0. \tag{16}$$

By (4) and Lemma 12 we have

$$\delta_{p+1,q+1} - M\beta\delta_{p,q+1} - (\alpha - \beta)\delta_{p+1,q} + \beta(M - 2\alpha)\delta_{p,q}$$
$$= v_{p+q+2} - (\alpha + (M - 1)\beta)v_{p+q+1} + \beta(M - 2\alpha)v_{p+q} - \Delta u_{p,q} \geq 0,$$

which can be rewritten as

$$\delta_{p+1,q+1} - M\beta\delta_{p,q+1} \geq (\alpha - \beta)\delta_{p+1,q} - \beta(M - 2\alpha)\delta_{p,q}. \tag{17}$$

We will now prove by induction on q the statement that $\delta_{p+1,q} \geq M\beta\delta_{p,q}$ for all $p \geq 0$. By (16), for the basis step we have to show that

$$v_{n+1} - \alpha^{n+1} \geq M\beta(v_n - \alpha^n).$$

This follows from

$$v_{n+1} - M\beta v_n = (\alpha - M\beta)(v_n + v_n') \geq (\alpha - M\beta)\alpha^n,$$

where we have used (3), $\alpha - M\beta > 0$, $v_n' \geq 0$, and $v_n \geq \alpha^n$. For the induction step, assume that the statement is true for a given $q \geq 0$. Using (16) it follows that $\delta_{p,q} \geq 0$ for that q and all p. Therefore, since $0 \leq M - 2\alpha \leq M(\alpha - \beta)$ (which is equivalent to $M\beta \leq \alpha$), (17) can be written as

$$\delta_{p+1,q+1} - M\beta\delta_{p,q+1} \geq \frac{M - 2\alpha}{M}[\delta_{p+1,q} - M\beta\delta_{p,q}] \geq 0,$$

where the final inequality follows from the induction hypothesis. This proves that $\delta_{p+1,q} \geq M\beta\delta_{p,q}$ for all $p, q \geq 0$, and hence that $\delta_{p,q} \geq 0$ for all $p, q \geq 0$ and therefore $u_{p,q} \leq v_{p+q}$. Now apply Lemma 11 to complete the proof. $\square$

5 Independence of parameter choice

Let $X = (X_n : n \geq 1)$ and $Y = (Y_n : n \geq 1)$ be two independent Bernoulli sequences with parameters p_X and p_Y respectively, and let $W \in \{0,1\}^{\mathbb{N}}$ be an arbitrary infinite word.

Theorem 13. (a) *The validity of the assertion "$\forall M \geq 2 : P(X$ is M-seen in $Y) = 0$" does not depend on the values of $p_X, p_Y \in (0,1)$.*
(b) *The validity of the assertion "$\forall M \geq 2 : P(W$ is M-seen in $Y) = 0$" does not depend on the value of $p_Y \in (0,1)$.*

The main idea of the proof of Theorem 13 is to use a coupling of two Bernoulli sequences so that one can be M-seen in the other for sufficiently large M.

Lemma 14. *Let X, X' be Bernoulli sequences with parameters $p, p' \in (0, 1)$ respectively.*
(a) If $p' \in [p^2, 1 - (1-p)^2]$, there is a coupling such that X' can be 3-seen in X.
(b) There is an $M \geq 2$ and a coupling such that X' can be M-seen in X.

Proof. (a) Let X be a Bernoulli sequence with parameter p. For a given sequence $x \in \{0, 1\}^{\mathbb{N}}$, we define a random subsequence $F(x)$ by partitioning x into disjoint blocks each comprising 2 consecutive letters, and replacing every block 00 by 0, 11 by 1, and 01 and 10 by 1 with probability p_1 and by 0 with probability $p_0 = 1 - p_1$. This is done independently for each block and independently of the choice of the Bernoulli sequence X. The sequence $X' := F(X)$ is a Bernoulli sequence with parameter

$$p' = p^2 + 2p(1-p)p_1$$

such that X' can be 3-seen in X (since for every k, $X'_k = X_{2k}$ or X_{2k-1}). Since p_1 can be chosen arbitrarily in $[0, 1]$, p' has a possible range of $[p^2, 1 - (1-p)^2]$.
(b) Let $f_1(p) = p^2$ and $f_2(p) = 1 - (1-p)^2$. We note that $f_1(p) \leq p \leq f_2(p)$ and $f_1^k(p) \to 0$ and $f_2^k(p) \to 1$ for $k \to \infty$, where $f^k(p) = f \circ \cdots \circ f(p)$. Thus for arbitrary given $p, p' \in (0, 1)$, there exist $p_0, \ldots, p_k \in (0, 1)$ such that $p_0 = p$, $p_k = p'$ and $p_{i+1} \in [f_1(p_i), f_2(p_i)]$ for all i. Thus by (a) there exist Bernoulli sequences $X^{(0)}, \ldots, X^{(k)}$ such that $X^{(i)}$ has parameter p_i and $X^{(i+1)}$ can be 3-seen in $X^{(i)}$ for all i. Thus $X' := X^{(k)}$ can be 3^k-seen in $X := X^{(0)}$. $\square$

Proof of Theorem 13. (a) Let $p_X, p_Y, p'_X, p'_Y \in (0, 1)$. By Lemma 14 there are Bernoulli sequences X, Y, X', Y' with these parameters such that X' can be M_X-seen in X, Y can be M_Y-seen in Y', and (X, X') is independent of (Y, Y'). In particular we have

$$\{X \text{ is } M\text{-seen in } Y\} \subset \{X' \text{ is } M_X M_Y M\text{-seen in } Y'\} \quad \text{for all } M.$$

Thus the validity of "$\forall M' \geq 2 : P(X' \text{ is } M'\text{-seen in } Y') = 0$" implies the validity of "$\forall M \geq 2 : P(X \text{ is } M\text{-seen in } Y) = 0$".
(b) Let $p_Y, p'_Y \in (0, 1)$. By Lemma 14 there are Bernoulli sequences Y, Y' with these parameters such that Y can be M_Y-seen in Y'. In particular we have

$$\{W \text{ is } M\text{-seen in } Y\} \subset \{W \text{ is } M_Y M\text{-seen in } Y'\} \quad \text{for all } M.$$

Thus the validity of "$\forall M' \geq 2 : P(W \text{ is } M'\text{-seen in } Y') = 0$" implies the validity of "$\forall M \geq 2 : P(W \text{ is } M\text{-seen in } Y) = 0$". $\square$

6 Variance of the number of embeddings

We prove Theorem 4. Let $W = (w_1, w_2, \ldots, w_n)$ be a word of length n, and let $j = (j_0, j_1, \ldots, j_n)$ and $k = (k_0, k_1, \ldots, k_n)$ be strictly increasing sequences of integers with

$j_0 = k_0 = 0$ and gaps not exceeding M. Then

$$E(N_n^2) = \sum_{j,k} P(j \text{ and } k \text{ are } M\text{-admissible embeddings})$$

$$= \sum_{j,k} (\tfrac{1}{2})^{|j \cup k|} I_{j,k} = \sum_{j,k} (\tfrac{1}{2})^{2n-|j \cap k|} I_{j,k},$$

where

$$I_{j,k} = I_{j,k}(W) = \prod_{(r,s) \neq 0 : \, j_r = k_s} 1(w_r = w_s),$$

$j \cup k = \{j_1, j_2 \ldots, j_n\} \cup \{k_1, k_2, \ldots, k_n\}$ viewed as a set, and $|j \cap k| = |\{(r,s) \neq (0,0) : j_r = k_s\}|$. Therefore,

$$E(N_n^2) = (M/2)^{2n} E(2^{|J \cap K|} I_{J,K}),$$

as claimed in part (a).

Part (b) follows from the fact that, for the random word X,

$$E(I_{j,k}(X)) = (\tfrac{1}{2})^{|j \cap k| - Z(j,k)},$$

where $Z(j,k) = |\{l \neq 0 : j_l = k_l\}|$.

To determine the asymptotics of $E(2^{Z_n})$, we proceed as follows. Let $\tau, \tau_1, \tau_2, \ldots$ be iid with the distribution of the hitting time of 0 for the random walk $J - K$ starting at 0. Then $P(Z_n \geq k) = P(\tau_1 + \cdots + \tau_k \leq n)$, so

$$E(2^{Z_n}) = 1 + \sum_{k=1}^{\infty} 2^{k-1} P(\tau_1 + \cdots + \tau_k \leq n).$$

Therefore, with $f(z) = Ez^\tau$,

$$(1-z) \sum_{n=0}^{\infty} z^n E(2^{Z_n}) = \frac{1 - f(z)}{1 - 2f(z)}, \tag{18}$$

whenever $|2f(z)| < 1$. Two methods are presented for extracting the asymptotics of $E(2^{Z_n})$ from (18).

The function f is smooth on the real interval $[0, 1)$, so if we define $c = c_M > 1$ by $Ec^{-\tau} = \tfrac{1}{2}$, the right side of (18) is asymptotic to a constant multiple of $(1 - cx)^{-1}$ as $x \uparrow c^{-1}$. By a Tauberian theorem (e.g., Theorem 5 in [4, Sect. XIII.5]), it follows that $E(2^{Z_n})$ is asymptotic to a constant multiple of c^n. When $M = 2$, $Ex^\tau = 1 - \sqrt{1-x}$ (see, for example, [3, Sect. XIV.4]), so $c = c_2 = \tfrac{4}{3}$ in this case.

In order to get full benefit from the Tauberian theorem, we need to know that the sequence $E(2^{Z_n})c^{-n}$ is monotone; otherwise we would only have convergence in the Cesàro sense. We check this next. Let $u_n = P(J_n - K_n = 0)$, $v_0 = 1$,

$$v_n = \sum_{k=1}^{n} u_k v_{n-k}, \quad n \geq 1, \tag{19}$$

14

and $V_n = \sum_{k=0}^{n} v_k$. We will check that

$$u_n \downarrow, \tag{20}$$

$$V_n^2 \geq V_{n+1}V_{n-1}, \quad n \geq 1 \tag{21}$$

and

$$E(2^{Z_n}) = V_n. \tag{22}$$

Once these are checked, it will follow from (21) that V_{n+1}/V_n is decreasing, and by (22) and the Cesàro convergence noted above, $V_{n+1}/V_n \downarrow c$. Therefore $V_{n+1} \geq cV_n$, so V_n/c^n is increasing as required.

To check (20), note first that by the Schwarz inequality,

$$P(J_n - K_n = m) = \sum_l P(K_n = l)P(J_n = l + m)$$

$$\leq \sqrt{\sum_l P^2(K_n = l) \sum_l P^2(J_n = l + m)}$$

$$= \sum_l P^2(K_n = l) = u_n.$$

Then use this to write

$$u_{n+1} = \sum_m P(J_n - K_n = m)P(J_1 - K_1 = -m) \leq u_n.$$

We check (21) by induction on n. Note that it is equivalent to

$$\frac{v_n}{V_{n-1}} \geq \frac{v_{n+1}}{V_n}. \tag{23}$$

Dividing (19) by V_{n-1} expresses the left side of (23) as an average of $u_1, \ldots, u_n$, and of course the right side is an average of $u_1, \ldots, u_{n+1}$. By (20), it suffices to check that the two averaging measures are stochastically ordered. But this is equivalent to $V_{n-1}V_j \geq V_nV_{j-1}$ for $1 \leq j \leq n$, which is a consequence of the induction hypothesis.

Finally, we check (22). Write

$$E(2^{Z_n}) = E \prod_{i=1}^{n} \left(1 + 1(J_i = K_i) \right) = \sum_{A \subseteq \{1,\ldots,n\}} P(J_i = K_i, \forall\, i \in A).$$

So, it is enough to show that

$$v_n = \sum_{n \in A \subseteq \{1,\ldots,n\}} P(J_i = K_i, \forall\, i \in A),$$

or equivalently, that the right side above satisfies the recursion (19). But this is easily checked by breaking up the sum according to the value of the smallest element k of A.

Finally, we give a second analysis of (18). This approach is simpler, but gives somewhat less than the first, in that it does not imply the log-concavity statement in (21). With c given as after (18), f is analytic on the disk $|z| < 1$, and $f'(c^{-1}) > 0$. Therefore,

$$h(z) := \frac{1 - f(z)}{(1 - z)(1 - 2f(z))}$$

has a simple pole at $z = c^{-1}$.

On the other hand, $|f(z)| \leq f(|z|)$, since the power-series expansion of f has real nonnegative coefficients. Therefore, $|f(z)| < \frac{1}{2}$ for $|z| \leq c^{-1}$, $z \neq c^{-1}$, since τ is not concentrated on a non-trivial arithmetic progression. It follows that $z = c^{-1}$ is the only root of $f(z) = \frac{1}{2}$ in some disk $|z| < c^{-1} + \epsilon$ with $\epsilon > 0$. Hence, $z = c^{-1}$ is the only singularity of h in that disk, whence

$$h(z) = \frac{a}{1 - cz} + g(z)$$

for some $a > 0$ and with g analytic in the disk. The coefficients of the power series for g grow as $\mathrm{O}(d^n)$ for some $0 < d < c$, so it follows from (18) that $E(2^{Z_n}) = ac^n + \mathrm{O}(d^n)$.

Acknowledgments

References

[1] P. N. Balister, B. Bollobás and A. M. Stacey, *Dependent percolation in two dimensions*, Probab. Th. Rel. Fields **117** (2000), 495–513. MR 2001h:60173

[2] I. Benjamini and H. Kesten, *Percolation of arbitrary words in $\{0, 1\}^{\mathbb{N}}$*, Ann. Probab. **23** (1995) 1024–1060. MR 97a:60140

[3] W. Feller, *An Introduction to Probability Theory and its Applications, Volume I, 3rd edition*, Wiley (1968). MR 37 #3604

[4] W. Feller, *An Introduction to Probability Theory and its Applications, Volume II*, Wiley (1966). MR 35 #1048

[5] G. R. Grimmett, *Three problems for the clairvoyant demon*, (2009) arxiv:0903.4749.

[6] H. Kesten, V. Sidoravicius and Y. Zhang, *Almost all words are seen in critical site percolation on the triangular lattice*, Elect. J. Probab. **3** (1998) Paper #10, 1–75. MR 99j:60155

[7] H. Kesten, V. Sidoravicius, and Y. Zhang, *Percolation of arbitrary words on the close-packed graph of Z^2*, Elect. J. Probab. **6** (2001) Paper #4, 1–27. MR 2002i:60181

[8] B. N. B. de Lima, *A note about the truncation question in percolation of words*, Bull. Braz. Math. Soc. New Ser. **39** (2008), 183–189. MR 2419900

[9] P. Winkler, *Dependent percolation and colliding random walks*, Rand. Struct. Alg. **16** (2000), 58–84. MR 2001c:60160

Statistical Laboratory
University of Cambridge
Wilberforce Road
Cambridge CB3 0WB, UK

email: g.r.grimmett@statslab.cam.ac.uk

Department of Mathematics
University of California, Los Angeles
405 Hilgard Ave.
Los Angeles CA 90095, USA

email: tml@math.ucla.edu
email: richthammer@math.ucla.edu