The Annals of Probability
1998, Vol. 26, No. 3, 1212–1231

NEAREST-NEIGHBOR WALKS WITH LOW PREDICTABILITY PROFILE AND PERCOLATION IN $2+\varepsilon$ DIMENSIONS

By Olle Häggström[1] and Elchanan Mossel

Chalmers University of Technology and Hebrew University of Jerusalem

A few years ago, Grimmett, Kesten and Zhang proved that for supercritical bond percolation on $\mathbb{Z}^3$, simple random walk on the infinite cluster is a.s. transient. We generalize this result to a class of wedges in $\mathbb{Z}^3$ including, for any $\varepsilon \in (0,1)$, the wedge $\mathscr{W}_\varepsilon = \{(x, y, z) \in \mathbb{Z}^3\colon x \geq 0, |z| \leq x^\varepsilon\}$ which can be thought of as representing a $(2+\varepsilon)$-dimensional lattice. Our proof builds on recent work of Benjamini, Pemantle and Peres, and involves the construction of finite-energy flows using nearest-neighbor walks on $\mathbb{Z}$ with low predictability profile. Along the way, we obtain some new results on attainable decay rates for predictability profiles of nearest-neighbor walks.

1. Introduction. It is a classical theorem of Pólya [14] that simple random walk on the cubic lattice $\mathbb{Z}^d$ is recurrent for $d = 1, 2$ and transient for $d \geq 3$. By inspecting quantities such as the Green's function, one is immediately led to think that the critical dimension should be 2 rather than some other number between 2 and 3. A natural candidate for a $(2+\varepsilon)$-dimensional lattice, $\varepsilon \in (0,1)$, is the wedge

$$\mathscr{W}_\varepsilon = \left\{(x, y, z) \in \mathbb{Z}^3\colon x \geq 0, |z| \leq x^\varepsilon\right\},$$

since the number of points in $\mathscr{W}_\varepsilon$ within distance n from the origin grows like $n^{2+\varepsilon}$. Lyons [13] showed that, indeed, simple random walk on $\mathscr{W}_\varepsilon$ is transient for each $\varepsilon > 0$. (See [16] and [2] for a different interpretation of random walks in noninteger dimensions.)

Grimmett, Kesten and Zhang (GKZ) showed in [8] that the transience for $d = 3$ is highly robust, in the following sense. Suppose that every edge in the $\mathbb{Z}^3$ lattice is removed independently with probability $(1-p)$, thus being retained with probability p. Then, for each $p \in (p_c, 1]$, where p_c is the critical value for independent bond percolation on $\mathbb{Z}^3$ (see [7] for basic notation and results on percolation), there is a.s. an infinite cluster among the retained edges on which simple random walk is transient. Obviously, this result cannot be pushed further as far as the value of p is concerned, since by definition there is a.s. no infinite cluster of retained edges when $p < p_c$, and (as everyone believes, although a rigorous proof is still lacking) not at $p = p_c$ either.

Corollary 1.2 below provides an extension of the GKZ theorem to the wedges $\mathscr{W}_\varepsilon$ for all $\varepsilon > 0$. This answers a question of Benjamini, Pemantle and Peres [3].

Received June 1997; revised February 1998.

[1]Supported by a grant from the Swedish Natural Science Research Council.

AMS 1991 *subject classifications*. Primary 60K35, 60J45; secondary 60J15.

Key words and phrases. Percolation, random walk, transience, predictability profile, Ising model.

In fact, the result also covers somewhat thinner wedges, such as

$$\{(x, y, z) \in \mathbb{Z}^3 : x \geq 0, |z| \leq (\log(x+1))^{2+\varepsilon}\}$$

for any $\varepsilon > 0$.

Percolation on wedges has been studied previously, for example, by Chayes and Chayes [5]. Most of the wedges we study have the same critical value for bond percolation as $\mathbb{Z}^3$, as will be evident from the final part of the proof of Theorem 1.1. For this reason, p_c will always denote the critical probability for bond percolation on $\mathbb{Z}^3$. Our main result is the following.

THEOREM 1.1. *Suppose that $h_1(x)$ and $h_2(x)$ are increasing positive functions such that for $i = 1, 2$ we have*

$$\sum_{j=1}^{\infty} \frac{h_i(j)}{j^2} < \infty, \tag{1}$$

and consider independent bond percolation on the wedge

$$\mathscr{W}_{h_1, h_2} = \{(x, y, z) \in \mathbb{Z}^3 : x \geq 0, |y| \leq h_1(x), |z| \leq h_2(x)\}$$

with retention probability $p \in (p_c, 1]$. If

$$\sum_{j=1}^{\infty} \frac{1}{h_1(j) h_2(j)} < \infty, \tag{2}$$

then the set of retained edges will a.s. contain an infinite cluster on which simple random walk is transient. Conversely, if the sum in (2) *diverges, then the set of retained edges will a.s. not contain an infinite cluster on which simple random walk is transient.*

For instance, the set of retained edges will contain an infinite cluster on which simple random walk is transient if we take $h_1(x) = h_2(x) = \sqrt{x(\log(x+1))^{1+\varepsilon}}$ with $\varepsilon > 0$, but not if we take $h_1(x) = h_2(x) = \sqrt{x \log(x+1)}$. (The cutoff for having an infinite cluster at all is much lower; see [5].)

Once we have proved Theorem 1.1, the next result follows with very little extra effort.

COROLLARY 1.2. *Suppose that $h(x)$ is an increasing positive function for which*

$$\sum_{j=1}^{\infty} \frac{1}{j\sqrt{h(j)}} < \infty. \tag{3}$$

Independent bond percolation with retention probability $p \in (p_c, 1]$ on the wedge

$$\mathscr{W}_h = \{(x, y, z) \in \mathbb{Z}^3 : x \geq 0, \ |z| \leq h(x)\},$$

will then a.s. yield an infinite cluster on which simple random walk is transient.

An interesting aspect of Theorem 1.1 is that condition (2) coincides with Lyons' [13] criterion for transience of simple random walk on $\mathscr{W}_{h_1,h_2}$. In other words, we get transience of simple random walk on the infinite cluster if and only if simple random walk on $\mathscr{W}_{h_1,h_2}$ is transient. Corollary 1.2, however, does not share this feature and may leave room for improvement; see the discussion in Section 6.

The way we prove Theorem 1.1 is to sharpen the techniques of Benjamini, Pemantle and Peres [3], who gave a new proof (and some generalizations) of the GKZ theorem. Geoffrey Grimmett has kindly informed us that the original approach in [8] can be used to prove Theorem 1.1 and Corollary 1.2 under slightly stronger assumptions on h, h_1 and h_2.

The main part of the game is to prove transience for some $p < 1$; once this is done the modern renormalization technology developed by Grimmett and Marstrand [9], Antal and Pisztora [1] and others can be invoked to bring the result all the way down to the critical point p_c. The approach of Benjamini, Pemantle and Peres is to construct finite-energy flows on the set of retained edges using nearest-neighbor walks on Z whose so-called predictability profile is sufficiently small. By a nearest-neighbor walk, we mean a random process $\{S_n\}_{n=0}^{\infty}$ taking values in Z such that $|S_{n+1} - S_n| = 1$ for each n.

DEFINITION 1.3. For a random process $S = \{S_n\}_{n=0}^{\infty}$ taking values in the finite or countably infinite set V, the predictability profile $\{\mathrm{PRE}_S(k)\}_{k=1}^{\infty}$ of S is defined as

$$\mathrm{PRE}_S(k) = \sup \mathrm{P}\big[S_{n+k} = x \mid S_0, \dots, S_n\big],$$

where the supremum is taken over all $n \geq 0$, all $x \in V$ and all histories $S_0, \dots, S_n$.

$\mathrm{PRE}_S(k)$ should be thought of as the maximal chance of guessing S correctly k steps into the future, given the process up to the present. In some sense, simple random walk is of course the least predictable of all nearest-neighbor walks on Z, but in the sense of asymptotics of the predictability profile as $k \to \infty$, it is not! Whereas by the local central limit theorem, simple random walk has predictability profile of the order $k^{-1/2}$, [3] constructed, for any $\alpha < 1$, nearest-neighbor walks with predictability profile $O(k^{-\alpha})$. The following theorem is an improvement of the result in [3].

THEOREM 1.4. *For any decreasing positive sequence* $\{f(k)\}_{k=1}^{\infty}$ *such that*

$$\sum_{j=1}^{\infty} \frac{f(j)}{j} < \infty, \tag{4}$$

there exists a constant $C < \infty$ *and a nearest-neighbor walk* $S = \{S_n\}_{n=0}^{\infty}$ *on* Z *such that*

$$\mathrm{PRE}_S(k) \leq \frac{C}{kf(k)} \tag{5}$$

for all $k \geq 1$.

For instance, taking $f(k) = 1/(\log(k+1))^{1+\varepsilon}$ for $\varepsilon > 0$ gives a nearest-neighbor walk whose predictability profile is $O((\log(k+1))^{1+\varepsilon}/k)$. Theorem 1.4 is sharp, as Hoffman [10] very recently has shown; if f is decreasing and the sum in (4) diverges, then the predictability profile in (5) is impossible to achieve.

Theorem 1.4 is a key ingredient in the proof of Theorem 1.1. Levin and Peres [11] have recently found another application of Theorem 1.4 in percolation theory.

We shall present two alternative constructions leading to a proof of Theorem 1.4. These are somewhat different, and were obtained independently by the two authors of this paper. The first construction is based on the Ising model on a tree, and the second is a kind of random walk in random environment. For both constructions, it will be convenient to note that the condition (4) is equivalent to having

$$\sum_{j=1}^{\infty} f(b^j) < \infty \tag{6}$$

for some (hence any) $b > 1$.

While proving Theorem 1.4 via the Ising construction, we derive in Corollary 2.3 a result of independent interest, concerning the spin-sum on the boundary of the Ising model with fixed interaction strength.

One more result about attainable decay rates of predictability profiles will be needed in our proof of Theorem 1.1.

PROPOSITION 1.5. *Suppose that* $\{f(j)\}_{j=0}^{\infty}$ *is an increasing positive sequence for which*

$$\sum_{j=1}^{\infty} \frac{f(j)}{j^2} < \infty. \tag{7}$$

Then there exists a nearest-neighbor walk $S^* = \{S_n^*\}_{n=0}^{\infty}$ *on* Z *such that* $|S_n^*| \le f(n) + 1$ *for all* n *and whose predictability profile satisfies*

$$\mathrm{PRE}_{S^*}(k) \le \frac{C}{f(k/8)}$$

for some $C < \infty$.

If $f(k)/f(k/2)$ is bounded, then the predictability profile is $O(1/f(k))$, which is of course optimal up to determination of the constant C. The use of $f(n)+1$ rather than $f(n)$ is only to make sure that the walk can get started, and is irrelevant for the asymptotics.

The rest of this paper is organized as follows. In Sections 2 and 3 we give the two alternative constructions which prove Theorem 1.4, and in Section 4 we prove Proposition 1.5. Section 5 contains proofs of Theorem 1.1 and Corollary 1.2, and the final section contains a short concluding discussion.

2. Unpredictable walks: first construction. In this section we give the first proof of Theorem 1.4. The main part of the proof is Lemma 2.1 concerning the distribution of the spin sum over the boundary for the Ising model on a regular tree. The tree-indexed Ising model has been studied before by many authors; see, for example, [12] and [4]. One difference between our set-up and previous ones is that here we allow the interaction strength to vary in the tree. Our analysis uses Fourier transforms and resembles a process constructed in [3].

Let $b \geq 2$ be an integer. We build a tree T_b, also known as the "hierarchical lattice" which has Z as the boundary set. The levels of the tree L_i will be defined inductively, starting at the boundary. The boundary L_0 is simply Z. To define L_{i+1}, denote the vertices of L_i from left to right by $\{v_i\}_{i=-\infty}^{\infty}$. Set L_{i+1} to be $\{w_i\}_{i=-\infty}^{\infty}$ (left to right) where w_i is the parent of $v_{bi} \cdots v_{bi+b-1}$. A vertex in L_i will be called a vertex of level i. Thus a vertex of level i has distance i from the boundary of T_b.

Let $\frac{1}{2} > \varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq 0$ be a sequence and consider the following labeling $\{\sigma(v)\}$ of the vertices of T_b by ± 1 valued random variables called *spins*. For each vertex v in level $i \geq 1$ and for each of its children w, assign $\sigma(w)$ to be $\sigma(v)$ with probability $1-\varepsilon_i$, and $-\sigma(w)$ with probability ε_i, independently for all children w. Using Kolmogorov's consistency theorem, we obtain a labeling of T_b, in which $\mathsf{P}[\sigma(v)=1] = \mathsf{P}[\sigma(v)=-1] = \frac{1}{2}$ for all $v \in T_b$. For each level N we denote by Y_N the spin sum on the boundary of the subtree which has a root w at level N, given $\sigma(w)=1$. It is clear that the distribution of Y_N does not depend on the choice of w.

LEMMA 2.1. *The process Y_N satisfies the inequality*

$$\mathsf{P}[Y_N = x] \leq \frac{C}{b^N \varepsilon_N \prod_{k=1}^{N-1}(1-2\varepsilon_k)}$$

for all $N \geq 1$, all x, where C depends only on b.

The following result is an immediate consequence.

COROLLARY 2.2. *If $\sum_{k=1}^{\infty} \varepsilon_k < \infty$, then Y_N satisfies the inequality*

$$\mathsf{P}[Y_N = x] \leq \frac{C}{b^N \varepsilon_N}$$

for all $N \geq 1$, all x, where C depends only on b and on $\prod_{k=1}^{\infty}(1-2\varepsilon_k)$.

The case of homogeneous interaction strength is of independent interest, so we state this as a separate corollary.

COROLLARY 2.3. *When the interaction-strength is constant, that is, $\varepsilon_k = \varepsilon$ for all k, we have*

$$\mathsf{P}\big[Y_N = x\big] \leq \frac{C}{\varepsilon(b(1-2\varepsilon))^N}$$

for all $N \geq 1$ and all x. Here C depends only on b.

PROOF OF LEMMA 2.1. By decomposing the sum in the definition of Y_{N+1} into b parts according to level N of the subtree having root at level $N+1$ (that is, by decomposing the sum according to the children of the root of that tree) we get

$$Y_{N+1} = \sum_{j=1}^{b} \sigma(v_j) Y_N^{(j)},$$

where $\{\sigma(v_j)\}_{j=1}^{b}$ are b i.i.d. spins, having distribution $(1-\varepsilon_{N+1}, \varepsilon_{N+1})$, and $\{Y_N^{(j)}\}_{j=1}^{b}$ are i.i.d. variables with the distribution of Y_N, that are independent of these spins. Consequently, the characteristic functions

$$\hat{Y}_N(\lambda) = \mathsf{E}\big(\exp(i\lambda Y_N)\big)$$

satisfy the recursion

$$\begin{aligned} \hat{Y}_{N+1}(\lambda) &= \big((1-\varepsilon_{N+1})\hat{Y}_N(\lambda) + \varepsilon_{N+1}\hat{Y}_N(-\lambda)\big)^b \\ &= \big(\Re\hat{Y}_N(\lambda) + i(1-2\varepsilon_{N+1})\Im\hat{Y}_N(\lambda)\big)^b, \end{aligned} \tag{8}$$

where $\Re$ denotes real part, and $\Im$ imaginary part. Define

$$\theta_N(\lambda) = \arg \hat{Y}_N(\lambda),$$

$$J_n = \left\{0 \le \lambda \le \frac{\pi}{2} \colon \theta_k(\lambda) < \frac{\pi}{2b},\ k = 0 \cdots n-1\right\}$$

and

$$I_n = \left\{0 \le \lambda \le \frac{\pi}{2} \colon \theta_k(\lambda) < \frac{\pi}{2b}, k = 0 \cdots n-1, \theta_n(\lambda) \ge \frac{\pi}{2b}\right\}.$$

We will evaluate the Ł^1 norm of $\hat{Y}_N$ by looking at the decomposition $[0, \pi/2] = \bigcup_{k=0}^{N-1} I_k \cup J_N$, and bounding Y_N on each of these intervals, and the length of those intervals. The symbol $|I|$ will denote the length of the interval I. We can rewrite (8) as

$$\hat{Y}_{n+1}(\lambda) = \big|\hat{Y}_n(\lambda)\big|^b \big[\cos(\theta_N(\lambda)) + i(1-2\varepsilon_{n+1})\sin(\theta_N(\lambda))\big]^b, \tag{9}$$

and get, for $0 \le \theta_N(\lambda) \le \pi/2b$, that

$$\theta_{N+1}(\lambda) = b \arctan\big((1-2\varepsilon_{N+1})\tan(\theta_N(\lambda))\big).$$

Since arctan is decreasing and concave in $[0, \infty)$, we obtain for $0 \le \theta_N(\lambda) \le \pi/2b$,

$$\frac{\pi}{2} \ge b\theta_N(\lambda) \ge \theta_{N+1}(\lambda) \ge b(1-2\varepsilon_{N+1})\theta_N(\lambda). \tag{10}$$

By (10) for $\lambda \in I_n$, $\pi/2 \geq \theta_n(\lambda) \geq \pi/2b$, and using (9) we get

$$\begin{aligned} \left|\hat{Y}_{n+1}(\lambda)\right| &\leq \left(\cos^2\left(\frac{\pi}{2b}\right) + (1-2\varepsilon_{n+1})^2 \sin^2\left(\frac{\pi}{2b}\right)\right)^{b/2} \\ &\leq \left(1 - 2\varepsilon_{n+1}\sin^2\left(\frac{\pi}{2b}\right)\right)^{b/2} \leq \exp(-\rho\varepsilon_{n+1}b), \end{aligned}$$

where ρ denotes $\sin^2(\pi/2b)$. Inductive usage of (9) for $\lambda \in I_n$ and $N > n$ now gives

$$\left|\hat{Y}_N(\lambda)\right| \leq \exp(-\rho\varepsilon_{n+1}b^{N-n}). \tag{11}$$

By (10) we have

$$\left|I_n\right| \leq \left|J_n\right| \leq \frac{\pi}{2b^n \prod_{k=1}^{n-1}(1-2\varepsilon_k)}, \tag{12}$$

and by (11),

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\hat{Y}_N(\lambda)\right| d\lambda &= \frac{2}{\pi}\int_0^{\pi/2}\left|\hat{Y}_N(\lambda)\right| d\lambda \\ &\leq \frac{2}{\pi}\left(\sum_{k=0}^{N-1}|I_k|\exp(-\rho\varepsilon_{k+1}b^{N-k}) + |J_N|\right). \end{aligned}$$

Inserting (12) yields

$$\begin{aligned} &\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\hat{Y}_N(\lambda)\right| d\lambda \\ &\quad \leq \frac{1}{\prod_{k=1}^{N-1}(1-2\varepsilon_k)}\left(\sum_{k=0}^{N-1} b^{-k}\exp(-\rho\varepsilon_{k+1}b^{N-k}) + b^{-N}\right). \end{aligned} \tag{13}$$

In order to evaluate the sum in the right-hand side of (13), we take the last n such that $\rho\varepsilon_{n+1}b^{N-n} > 1$, and get

$$\sum_{k=n}^{N-1} b^{-k}\exp(-\rho\varepsilon_{k+1}b^{N-k}) + b^{-N} \leq \sum_{k=n}^{N} b^{-k} \leq b^{-n}\sum_{k=0}^{\infty} b^{-k}. \tag{14}$$

Since $\{\varepsilon_k\}$ is decreasing, we get

$$\sum_{k=0}^{n-1} b^{-k}\exp(-\rho\varepsilon_{k+1}b^{N-k}) \leq \sum_{k=0}^{n-1} b^{-k}\exp(-b^{n-k}) \leq b^{-n}\sum_{k=0}^{\infty} b^k\exp(-b^k). \tag{15}$$

Furthermore, since $\rho\varepsilon_{n+2}b^{N-n-1} \leq 1$, we have that

$$b^{-n} \leq \frac{1}{\rho b^{N-1}\varepsilon_{n+2}} \leq \frac{1}{\rho b^{N-1}\varepsilon_N}. \tag{16}$$

Now combining (13), (14), (15) and (16) we see that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\hat{Y}_N(\lambda)|\, d\lambda \leq \frac{C}{b^N\varepsilon_N\prod_{k=1}^{N-1}(1-2\varepsilon_k)},$$

where

$$C = \frac{b\left(\sum_{k=0}^{\infty} b^{-k} + \sum_{k=0}^{\infty} b^k \exp(-b^k)\right)}{\sin^2(\pi/2b)}.$$

Using the inversion formula we finally achieve the bound

$$\begin{aligned} \mathrm{P}[Y_N = x] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{Y}_N(\lambda)\exp(-i\lambda x)\,d\lambda \\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} |\hat{Y}_N(\lambda)|\,d\lambda \le \frac{C}{b^N \varepsilon_N \prod_{k=1}^{N-1}(1-2\varepsilon_k)}. \end{aligned} \qquad \square$$

PROOF OF THEOREM 1.4. We can assume that for all k, $0 < f(k) < \frac{1}{2}$. Take any $b \ge 2$ and set $\varepsilon_k = f(b^k)$. By (6) we have that $\sum_{k=1}^{\infty} \varepsilon_k < \infty$. We look at the tree T_b with the spin assignment induced by $\{\varepsilon_k\}$. Fix $v_0 \in L_0$, let $v_1, v_2 \cdots$ be the elements of the boundary of T_b to the right of v_0 and set

$$S_n = \sum_{k=1}^{n} \sigma(v_k).$$

We claim that S_n has the desired predictability profile. To see this, fix $n \ge 0$, $k > 0$ and note that $S_{n+k} = S_n + \sum_{j=n+1}^{n+k} \sigma(v_j)$. If we now take the unique h satisfying $2b^h \le k < 2b^{h+1}$, there will exist a vertex w at level h (i.e., at distance h from the boundary) for which all of the descendants at level 0 are in the set $\{v_{n+1}, \dots, v_{n+k}\}$. From this it follows (by conditioning on the spins of all v_i which are not descendants of w, and on the spin of w) that

$$\sup_{x \in \mathbb{Z}} \mathrm{P}[S_{n+k} = x | S_0, \dots, S_n] \le \sup_{x \in \mathbb{Z}} \mathrm{P}[Y_h = x]. \tag{17}$$

Now using Corollary 2.2 and (17) we get

$$\mathrm{PRE}_S(k) \le \frac{C}{b^h \varepsilon_h} \le \frac{2bC}{kf(b^h)} \le \frac{2bC}{kf(k)}, \tag{18}$$

and the proof is complete. $\square$

REMARK. The above construction can be modified in such a way that S gets stationary increments. Indeed, by considering the entire bi-infinite boundary of T_b we get a bi-infinite process. If in the construction of all levels j of T_b we take s to be uniform shift in $\{0 \cdots b-1\}$ and set each w_i in L_j to be the parent of $v_{bi+s} \cdots v_{bi+b-1+s}$, then we get a process with stationary increments and the desired predictability profile.

3. Unpredictable walks: second construction. The processes we construct in this section are a kind of random walk in random environment. For the usual random walk in random environment [17], [15], the environment is fixed in time and varies in space. In contrast, the environment in our set-up varies in time but not in space, so that our processes are closer related to

the birth and death chains in random environment studied by Torrez [18]. Whereas transform methods were needed to obtain the desired predictability profile in the previous section, the methods in the present section are purely probabilistic.

Let S_n be the sum $\sum_{i=1}^{n} \sigma_i$ of $\{-1,1\}$-valued random variables $\{\sigma_i\}_{i=1}^{\infty}$ which are independent conditioned on the random environment $\{p_i\}_{i=1}^{\infty}$. At each time i, σ_i takes value $+1$ with probability p_i and -1 with probability $1-p_i$. Fix an integer $b>1$, and furthermore let $\{a_j\}_{j=1}^{\infty}$ be a positive sequence such that

$$\sum_{j=1}^{\infty} a_j < \tfrac{1}{2}.$$

The random environment is obtained as

$$p_i = \tfrac{1}{2} + p_i^{(1)} + p_i^{(2)} + \cdots,$$

where $\{p_i^{(1)}\}_{i=1}^{\infty}, \{p_i^{(2)}\}_{i=1}^{\infty}, \dots$ are independent processes defined by the following.

1. For each i and j, the distribution of $p_i^{(j)}$ is uniform on $[-a_j, a_j]$.
2. The value $p_i^{(j)}$ is constant in i for $i = 1, \dots, b^j$. At time b^j+1 it switches to a new independent value uniform on $[-a_j, a_j]$ which is kept until time $2b^j$ and so on.

This defines the distribution of $\{\sigma_i\}_{i=1}^{\infty}$. Concretely, the process can be realized by letting $\{U_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables, independent also of the $\{p_i^{(j)}\}_{i=1}^{\infty}$ processes, uniformly distributed on $[0,1]$, and letting

$$\sigma_i = \begin{cases} +1, & \text{if } U_i < p_i, \\ -1, & \text{otherwise.} \end{cases}$$

Theorem 1.4 is an immediate consequence of the following result and the observation (6). We write $\lfloor x \rfloor$ for the integer part of x.

PROPOSITION 3.1. *Given b and $\{a_j\}_{j=1}^{\infty}$, there exists a $C<\infty$ such that*

$$\mathrm{PRE}_S(k) \le \frac{C}{k a_{\lfloor \log_b(k/2) \rfloor}}.$$

for all k.

PROOF. We may safely assume that $k \ge 2b$. Let

$$m_k = \lfloor \log_b(k/2) \rfloor.$$

Suppose that we know $S_0, \dots, S_n$ and want to predict S_{n+k}. The time interval $I_{n+1}^{n+k} = \{n+1, \dots, n+k\}$ will certainly contain some subinterval $I = \{l, \dots, l+b^{m_k}-1\}$ on which the $p_i^{(m_k)}$ process stays constant; fix such an I and write $p_I^{(m_k)}$ for the common value of $p_i^{(m_k)}$, $i \in I$. Clearly, $p_I^{(m_k)}$ is independent of the

$p_i^{(m_k)}$ process outside of this interval, and thus also of $\{\sigma_i\}_{i\notin I}$. The cardinality b^{m_k} of I satisfies

$$b^{m_k} = b^{\lfloor \log_b(k/2)\rfloor} \geq b^{\log_b(k/2)-1} = \frac{k}{2b}.$$

Write $\tilde{p}_i^{(m_k)}$ for $p_i - p_i^{(m_k)}$, and note that $\{p_i^{(m_k)}\}_{i=1}^{\infty}$ is independent of $\{\tilde{p}_i^{(m_k)}\}_{i=1}^{\infty}$. For $i \in I$, define the random variables

$$X_i = U_i - \tilde{p}_i^{(m_k)}$$

and

$$Y_i = \begin{cases} 1, & \text{if } X_i < -a_{m_k}, \\ 0, & \text{if } X_i \in [-a_{m_k}, a_{m_k}], \\ -1, & \text{if } X_i > a_{m_k}. \end{cases}$$

The Y_i's are not independent, but each of them independently takes value 0 with probability $2a_{m_k}$. Therefore, $\#\{i \in I\colon Y_i = 0\}$ has a binomial $(b^{m_k}, 2a_{m_k})$ distribution.

Next, suppose that in addition to $S_0, \dots, S_n$, we are also informed of the values of $\{\sigma_i\}_{i \in I_{n+1}^{n+k}\setminus I}$ and $\{Y_i\}_{i\in I}$. Given this extra information, we know that S_{n+k} has to be in the interval

$$L = \Bigg\{S_n + \sum_{i\in I_{n+1}^{n+k}\setminus I} \sigma_i + \sum_{i\in I} Y_i - \sum_{i\in I} 1_{\{Y_i=0\}}, \dots, S_n + \sum_{i\in I_{n+1}^{n+k}\setminus I} \sigma_i + \sum_{i\in I} Y_i + \sum_{i\in I} 1_{\{Y_i=0\}}\Bigg\}.$$

Now comes the key step of the proof, which is to note that the conditional distribution (given the extra information) of S_{n+k} is in fact *uniform* on the set

$$L_{\text{even}} = \Bigg\{l \in L\colon l - \Bigg(S_n + \sum_{i\in I_{n+1}^{n+k}\setminus I} \sigma_i + \sum_{i\in I} Y_i - \sum_{i\in I} 1_{\{Y_i=0\}}\Bigg) \text{ is even}\Bigg\}. \tag{19}$$

This follows from the fact that

$$S_{n+k} - \Bigg(S_n + \sum_{i\in I_{n+1}^{n+k}\setminus I} \sigma_i + \sum_{i\in I} Y_i - \sum_{i\in I} 1_{\{Y_i=0\}}\Bigg) = 2\,\#\big\{i \in I\colon Y_i = 0, X_i < p_I^{(m_k)}\big\}$$

and the observation that the random variables $p_I^{(m_k)}$ and $\{X_i\}_{i\in I\colon Y_i=0}$ are conditionally i.i.d. and uniformly distributed on $[-a_{m_k}, a_{m_k}]$. Since L_{even} has cardinality $\#\{i \in I\colon Y_i = 0\} + 1$, we get for any x that

$$\mathsf{P}\big[S_{n+k} = x \mid S_0, \dots, S_n\big] \le \sum_{j=0}^{b^{m_k}} \frac{\mathsf{P}[\#\{i \in I\colon Y_i = 0\} = j]}{j+1}$$

$$\le \mathsf{P}\big[\#\{i \in I\colon Y_i = 0\} < b^{m_k} a_{m_k}\big] + \frac{1}{b^{m_k} a_{m_k}}.$$

The first term in the last expression tends to 0 exponentially fast in $b^{m_k} a_{m_k}$ and is thus asymptotically negligible (as $b^{m_k} a_{m_k} \to \infty$) compared to the second term. Hence, we can find a C' such that

$$\mathsf{P}\big[S_{n+k} = x \mid S_0, \dots, S_n\big] \le \frac{C'}{b^{m_k} a_{m_k}} = \frac{2bC'}{k a_{m_k}}.$$

Setting $C = 2bC'$ gives the desired result. $\square$

REMARK. Just as in Section 2, we can modify the above construction in order to obtain a process which has further desirable properties without losing anything essential in the upper bound for the predictability profile. If the $\{p_i^{(j)}\}_{i=1}^{\infty}$ processes are extended to negative times in the obvious way, and each of the processes independently is shifted by a random time lag equidistributed on $\{0, \dots, b^j - 1\}$, then the $\{\sigma_i\}_{i=1}^{\infty}$ process becomes stationary, so that $\{S_i\}_{i=1}^{\infty}$ gets stationary increments. If, furthermore, for each j the time intervals that $p_i^{(j)}$ stays constant are turned into an aperiodic renewal process (for instance by letting each time interval that $p_i^{(j)}$ stays fixed independently have length b^j or $b^j + 1$ with probability 1/2 each), then the $\{\sigma_i\}_{i=1}^{\infty}$ process becomes ergodic and Bernoulli, that is, isomorphic (in the sense of ergodic theory) to an i.i.d. process.

REMARK. There is some similarity between the two processes we have constructed, in that the dependence structure is hierarchical in both processes. An obvious question is whether in fact the two processes coincide. The answer is no: they are different. One way to see this is as follows. In the Ising model construction, the absolute value of

$$Y_{bi+b-1} - Y_{bi-1} = \sigma(v_{bi}) + \sigma(v_{bi+1}) + \cdots + \sigma(v_{bi+b-1})$$

is independent of all other increments (because it is independent of the spin of the parent of $v_{bi}, \dots, v_{bi+b-1}$). The random environment construction is easily seen not to have such an independence property.

REMARK. When an unpredictable walk has a predictability profile which is asymptotically of the order $k^{-\alpha}$, it is natural to ask whether it converges to a fractional Brownian motion with index α under appropriate space-time scaling. We suspect that neither of our processes exhibit such behavior. It would be interesting to see a construction of a process which combines the desired predictability and scaling properties.

4. Proof of Proposition 1.5. The purpose of this section is to give examples which prove Proposition 1.5. Fix a sequence $\{f(n)\}_{n=0}^{\infty}$ satisfying the assumptions of the proposition, and define another sequence $\{B(n)\}_{n=0}^{\infty}$ (B as in boundary) by letting

$$B(n) = 2^j \quad \text{where } j = \begin{cases} 0, & \text{if } f(n) \le 1, \\ \max\{i \in \{1, 2, \dots\} \colon 2^i \le f(n)\}, & \text{otherwise.} \end{cases}$$

Since $\{f(n)\}_{n=0}^{\infty}$ is increasing, the $\{B(n)\}_{n=0}^{\infty}$ sequence is also increasing. The process $S^* = \{S_n^*\}_{n=0}^{\infty}$ which we will use to prove Proposition 1.5 will be obtained by taking another nearest-neighbor walk $S = \{S_n\}_{n=0}^{\infty}$ and reflecting the path of S each time it attempts to cross the boundaries $\pm B(n)$. In order for S^* to have the required predictability profile we need that the distance of S_n from the lattices $B(k)\mathbb{Z}$ is smooth for $k \le n$. A precise version of this statement is given in the following key lemma.

LEMMA 4.1. *With $\{f(n)\}_{n=0}^{\infty}$ and $\{B(n)\}_{n=0}^{\infty}$ as above, there exists a nearest-neighbor walk $S = \{S_n\}_{n=0}^{\infty}$ such that for all $n \ge k$,*

$$\max_{x,(S_0,\dots,S_n)} \mathrm{P}\big[S_{n+k} \bmod B(k) = x \mid S_0, \dots, S_n\big] \le \frac{C}{f(k/4)}. \tag{20}$$

Once we have Lemma 4.1, the proof of Proposition 1.5 is as follows.

PROOF OF PROPOSITION 1.5. We first describe how S^* is obtained from S, where S is chosen as in Lemma 4.1. For this we shall use an auxiliary $\{-1, 1\}$-valued random sequence $\{Z_n\}_{n=0}^{\infty}$ which indicates whether S^* currently is moving in the same or in the opposite direction as S. Initially we have $S_0 = 0$ and set $S_0^* = 0$, $Z_0 = 1$. We then obtain S^* and Z inductively as follows. Suppose that S_{n-1}^* and Z_{n-1} have been determined. We then set

$$S_n^* = S_{n-1}^* + Z_{n-1}(S_n - S_{n-1}),$$
$$Z_n = Z_{n-1},$$

unless $S_{n-1}^* + Z_{n-1}(S_n - S_{n-1})$ happens to fall outside of the range

$$\{-B(n), \dots, B(n)\},$$

in which case we instead let

$$S_n^* = S_{n-1}^* - Z_{n-1}(S_n - S_{n-1}),$$
$$Z_n = -Z_{n-1}.$$

This guarantees that $\{S_n^*\}_{n=0}^{\infty}$ is a nearest-neighbor walk which for each n sits in the interval $\{-B(n), \dots, B(n)\}$.

Now, we show that S^* has the right predictability profile. Fix n and k, and assume for the moment that

$$n \ge k. \tag{21}$$

We then have for each $i \geq n$ that

$$B(i) \geq B(k) \geq \frac{f(k)}{2} \geq 2^{\alpha_k}, \tag{22}$$

where α_k denotes $\lfloor \log_2(f(k)/2) \rfloor$. We claim that there exists a constant d such that for all $i \geq n$,

$$\begin{aligned} (S_i - S_i^*) \bmod B(k) &= d \quad \text{if } Z_i = 1, \\ (S_i + S_i^*) \bmod B(k) &= d \quad \text{if } Z_i = -1. \end{aligned} \tag{23}$$

This follows by induction. Indeed, if we take

$$d = (S_n - Z_n S_n^*) \bmod B(k),$$

then the statement surely holds for $i = n$. If $Z_i = Z_{i+1}$ the induction step follows directly from the definition of S^*. Otherwise, by (22) $S_i^* \bmod B(k) = 0$, and again the induction step follows from the definition of S^*. By (23) we have that for $i \geq n$ for every given location of S_i^* there are only two possible locations of $S_i \bmod B(k)$. This, in combination with (20) and the fact that $(S_0^*, \dots, S_n^*)$ is reconstructible from $(S_0, \dots, S_n)$, implies that

$$\begin{aligned} &\max_{x, (S_0^*, \dots, S_n^*)} \mathrm{P}\big[S_{n+k}^* = x \mid S_0^*, \dots, S_n^*\big] \\ &\quad \leq \max_{x, (S_0, \dots, S_n)} \mathrm{P}\big[S_{n+k}^* = x \mid S_0, \dots, S_n\big] \\ &\quad \leq 2 \max_{x, (S_0, \dots, S_n)} \mathrm{P}\big[S_{n+k} \bmod B(k) = x \mid S_0, \dots, S_n\big] \\ &\quad \leq \frac{2C}{f(k/4)}. \end{aligned} \tag{24}$$

Now Proposition 1.5 is almost proved; we only need to remove the assumption (21). To do this, we just note that

$$\begin{aligned} &\max_{x, (S_0^*, \dots, S_n^*)} \mathrm{P}[S_{n+k}^* = x \mid S_0^*, \dots, S_n^*] \\ &\quad \leq \max_{x, \left(S_0^*, \dots, S_{\lceil n+k/2 \rceil}^*\right)} \mathrm{P}\big[S_{n+k}^* = x \mid S_0^*, \dots, S_{\lceil n+k/2 \rceil}^*\big] \end{aligned}$$

and that the right-hand side is less than $2C/f(k/8)$ by (24). □

We now go on to prove Lemma 4.1. Note first that if we restrict to the case $f(k) = o(\sqrt{k})$, then by the local central limit theorem with well-known error estimates, it suffices to let S be simple random walk. For the more general case considered in the lemma, a more complicated construction is clearly necessary.

Proof of Lemma 4.1. Take S to be the unpredictable nearest-neighbor walk of Section 3 with parameters b and $\{a_j\}_{j=1}^{\infty}$ chosen as follows. First let $b = 2$. Then write $g(x)$ for $f(x)/x$, and note that

$$g(x) \leq 2g(y) \quad \text{for } y \in \{x, \dots, 2x\}$$

so that

$$\sum_{j=1}^{\infty} \frac{f(2^j)}{2^j} = \sum_{j=1}^{\infty} g(2^j) = \sum_{j=1}^{\infty} \sum_{k=2^j}^{2^{j+1}-1} \frac{g(2^j)}{2^j} \leq \sum_{j=1}^{\infty} \sum_{k=2^j}^{2^{j+1}-1} \frac{2g(k)}{2^j}$$
$$\leq \sum_{j=1}^{\infty} \sum_{k=2^j}^{2^{j+1}-1} \frac{4g(k)}{k} = 4 \sum_{k=1}^{\infty} \frac{g(k)}{k} = 4 \sum_{k=1}^{\infty} \frac{f(k)}{k^2} < \infty.$$

We can therefore find a constant $C_1 > 0$ small enough so that

$$C_1 \sum_{j=1}^{\infty} \frac{f(2^j)}{2^j} < \frac{1}{2},$$

and let

$$a_j = \frac{C_1 f(2^j)}{2^j}$$

for each j. This defines S. We now go on to estimate the left-hand side of (20) for $n \geq k$. Let the notation of Section 3 be in force, and suppose that we condition S on the same extra information $\mathscr{I}$ as in the proof of Proposition 3.1. Write $|L_{\text{even}}|$ for the cardinality of the set L_{even} defined in (19). We have from the proof of Proposition 3.1 that $|L_{\text{even}}|$ is a binomial random variable with mean

$$2^{m_k} 2a_{m_k} = 2C_1 f\big(2^{\lfloor \log_2(k/2) \rfloor}\big) \in \big[2C_1 f(k/4), 2C_1 f(k/2)\big].$$

Letting E be the event that $|L_{\text{even}}| \in [C_1 f(k/4), 4C_1 f(k/2)]$, we have by standard large deviations theory that

$$\mathsf{P}[\neg E \mid S_0, \dots, S_n] \leq C_2 \exp(-C_3 f(k/4))$$

for some constants $C_2, C_3 \in (0, \infty)$ not depending on k. Furthermore, on the event E, we have for all k, that the map mod $B(k)$ maps at most

$$\frac{2 \cdot 4C_1 f(k/2)}{B(k)} \leq \frac{2 \cdot 4C_1 f(k/2)}{\frac{1}{2} f(k/2)} = 16C_1$$

different elements of L_{even} on the same $x \in \mathbb{Z}$, so that (still on the event E),

$$\begin{aligned}
&\max_x \mathsf{P}\big[S_{n+k} \bmod B(k) = x \mid (S_0, \dots, S_n), \mathscr{I}\big] \\
&\qquad \leq 16C_1 \max_x \mathsf{P}\big[S_{n+k} = x \mid (S_0, \dots, S_n), \mathscr{I}\big] \\
&\qquad \leq \frac{16}{f(k/4)}.
\end{aligned}$$

Hence, we have for large k that

$$\begin{aligned}
&\max_{x, (S_0, \dots, S_n)} \mathsf{P}\big[S_{n+k} \bmod B(k) = x \mid S_0, \dots, S_n\big] \\
&\qquad \leq \frac{16}{f(k/4)} + C_2 \exp(-c_3 f(k/4)) \leq \frac{C}{f(k/4)}
\end{aligned}$$

for some $C < \infty$, and Lemma 4.1 is proved. $\square$

We have also tried to find an alternative proof of Lemma 4.1 using the Ising model construction in Section 2, but we are only able to do this under stronger conditions on f. Since the tree-indexed Ising model is of independent interest, we end this section by showing how that is done.

We will assume that $g(j) = f(j)/j$ is a decreasing function of j which is bounded above by 1/4, and that f satisfies

$$\sum_{k=1}^{\infty} \sqrt{f(2^k)/2^k} < \infty. \tag{25}$$

These conditions hold for every function of the type $f(n) = n^{\alpha}$, for $\alpha < 1$, but are stronger than (7). Actually, condition (25) can be replaced by the weaker condition

$$\sum_{k=1}^{\infty} \bigl(f(2^k)/2^k\bigr)^{\alpha} < \infty \quad \text{for some } \alpha < 1,$$

but this requires more delicate estimates than those given below, so for brevity and simplicity we restrict to the case where (25) holds.

Take S to be the unpredictable process of Section 2, with $b = 2$. Choose $\varepsilon_n = \sqrt{f(2^n)/2^n}$. If we do the same conditioning as in the first proof of Theorem 1.4, It is easy to see that in order to prove (20), it is enough to show that for $2^K \le f(2^N)$, and for all $x \in \{0, \dots, 2^K - 1\}$,

$$\mathrm{P}[Y_n \bmod 2^K = x] \le C2^{-K}.$$

Therefore, it suffices to show that for all $x \in \{0, \dots, 2^K - 1\}$,

$$\bigl|\mathrm{P}[Y_n \bmod 2^K = x] - \mathrm{P}[Y_n \bmod 2^K = 0]\bigr| \le C2^{-K}. \tag{26}$$

By the inversion formula, the left-hand side in the last inequality equals

$$\begin{aligned} &\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{Y}_N(\lambda)(1 - \exp(-i\lambda x)) \sum_{l=-2^{N-K}}^{2^{N-K}} \exp(-i\lambda l 2^K)\, d\lambda\right| \\ &\qquad \le 2^{N+2}\int_0^{\pi/2} |\hat{Y}_N(\lambda)|\, |\lambda|\, d\lambda. \end{aligned} \tag{27}$$

In order to estimate the last integral, we use the same interval decomposition that we used in Section 2. Note that since $\sum_{j=1}^{\infty} \varepsilon_j < \infty$ we get $|I_l| \le C_1 2^{-l}$ using (12). Furthermore, by (11), we have for $\lambda \in I_n$ that $|\hat{Y}_N(\lambda)| \le \exp(-\rho\varepsilon_{n+1}2^{N-n})$. It follows that

$$\int_0^{\pi/2} \bigl|\hat{Y}_N(\lambda)\bigr||\lambda|\, d\lambda \le C_1^2\left(\sum_{l=0}^{N-1} 4^{-l} \exp -\rho\varepsilon_{l+1}2^{N-l}) + 4^{-N}\right). \tag{28}$$

For $l \le (N + K)/2$, we get

$$2^{N-l}\varepsilon_{l+1} \ge 2^{N-l}\varepsilon_N = 2^{N-l}\sqrt{f(2^N)/2^N} \ge 2^{N-l}\sqrt{2^K/2^N} = 2^{(N+K)/2-l},$$

so the sum in (28) is bounded by

$$C_1^2\left(\sum_{l=0}^{(N+K)/2} 4^{-l}\exp(-\rho 2^{(N+K-2l)/2}) + \sum_{l=(N+K)/2+1}^{N} 4^{-l}\right) \le C_2 2^{-N-K}. \tag{29}$$

Inserting (29) into (27), we get (26) and thus the desired predictability profile.

5. Transience in wedges.

In this section, we will prove Theorem 1.1 and Corollary 1.3. For the proof of Theorem 1.1, we shall make use of Proposition 1.5 and also of the following two results from [3].

LEMMA 5.1 (Benjamini, Pemantle and Peres [3]). *Let $S = \{S_n\}_{n=0}^{\infty}$ be a random process taking values in the countable set V. If the predictability profile for S satisfies $\sum_{k=1}^{\infty}\mathrm{PRE}_S(k) < \infty$, then there exist $C < \infty$ and $\theta \in (0,1)$ such that for any sequence $\{v_n\}_{n=0}^{\infty}$ and all $l \ge 1$ we have*

$$\mathrm{P}\big[\#\{n \ge 0\colon S_n = v_n\} \ge l\big] \le C\theta^l.$$

For an infinite graph G with a distinguished vertex v_0, we write $\Upsilon = \Upsilon(G, v_0)$ for the set of paths in G which start from v_0 and visit any vertex at most finitely many times. A probability measure μ on $\Upsilon(G, v_0)$ is said to have exponential intersection tails with parameter $\theta \in (0,1)$ [abbreviated as EIT(θ)] if there exists a C such that

$$\mu \times \mu\big\{(\varphi, \psi)\colon |\varphi \cap \psi| \ge n\big\} \le C\theta^n \tag{30}$$

for all n, where $|\varphi \cap \psi|$ is the number of edges contained both in φ and in ψ. It is shown in [3] how such a measure can be used to construct, with positive probability, a finite-energy flow from v_0 to "infinity" in the graph obtained from G by independent edge-thinning with retention probability $p > \theta$. This implies the following lemma.

LEMMA 5.2 (Benjamini, Pemantle and Peres [3]). *Consider independent bond percolation with retention probability p on an infinite graph G, and suppose that there exists a vertex v_0 in G and a measure μ on $\Upsilon(G, v_0)$ with the EIT(θ) property. If $p > \theta$, then the set of retained edges a.s. contains an infinite cluster on which simple random walk is transient.*

PROOF OF THEOREM 1.1. We begin with the second half of the theorem [divergence of the sum in (2) implies nonexistence of transience of infinite clusters], which is easy. Lyons [13] showed that if the sum in (2) is infinite, then simple random walk on $\mathscr{W}_{h_1,h_2}$ is recurrent. The second half of the theorem is thus an immediate consequence of Rayleigh's monotonicity principle, which states that by removing edges from a graph one can only make simple random walk on the graph "more recurrent" (see [6] for a more precise formulation and a proof).

For the first half of the theorem, suppose that the sum in (2) is finite. In order to show that the wedge $\mathscr{W}_{h_1,h_2}$ contains transient clusters for p close to 1, we will construct a measure μ on $\Upsilon(\mathscr{W}_h, v_0)$ which has the EIT(θ) property for some $\theta < 1$. Here we take $v_0 = (0,0,0)$. Let us assume that $h_1(0)$ and $h_2(0)$ are both at least 1. This is no loss of generality because if $\min(h_1(0), h_2(0)) < 1$ we can instead consider the wedge

$$\big\{(x, y, z) \in \mathrm{Z}^3\colon x \geq 0, |y| \leq h_1(x+k)|z| \leq h_2(x+k)\big\},$$

where k is chosen large enough so that $\min(h_1(0), h_2(0)) \geq 1$. This new wedge satisfies the assumptions of the theorem, including (2). Furthermore, the new wedge is a subset of a translate of $\mathscr{W}_{h_1, h_2}$, so Rayleigh's monotonicity principle then completes the result for $\mathscr{W}_{h_1, h_2}$.

Let $S^1 = \{S_n^1\}_{n=0}^\infty$ and $S^2 = \{S_n^2\}_{n=0}^\infty$ be two independent nearest-neighbor walks on Z starting at 0 chosen in such a way that for $i = 1, 2$, we have:

(i) $S_n^i \leq h_i(n)$ for all n, and
(ii) $\mathrm{PRE}_{S^i}(k) \leq C_i / h_i(k/8)$ for some $C_i < \infty$.

Such processes exist by Proposition 1.5 [this is where the assumption (1) is needed]. Now construct a $\mathscr{W}_{h_1, h_2}$-valued process $S = \{S_n\}_{n=0}^\infty$ as follows. Let

$$S_n = \begin{cases} \left(\left\lfloor \dfrac{n}{3} \right\rfloor, S^1_{\lfloor n/3 \rfloor}, S^2_{\lfloor n/3 \rfloor}\right), & \text{for } n = 0, 3, 6, \dots, \\ \left(\left\lfloor \dfrac{n}{3} \right\rfloor + 1, S^1_{\lfloor n/3 \rfloor}, S^2_{\lfloor n/3 \rfloor}\right), & \text{for } n = 1, 4, 7, \dots, \\ \left(\left\lfloor \dfrac{n}{3} \right\rfloor + 1, S^1_{\lfloor n/3 \rfloor + 1}, S^2_{\lfloor n/3 \rfloor}\right), & \text{for } n = 2, 5, 8, \dots. \end{cases}$$

and note that at each time step S_n changes exactly one of its three coordinates, and does this by ± 1. Hence, S_n can be viewed as a random path in $\mathscr{W}_{h_1, h_2}$ or, more precisely, a random element of $\Upsilon(\mathscr{W}_{h_1, h_2}, v_0)$. Let μ be the induced probability measure on $\Upsilon(\mathscr{W}_{h_1, h_2}, v_0)$.

Since S^1 and S^2 are independent, we get for $k = 3, 6, 9, \dots$ that

$$\mathrm{PRE}_S(k) \leq \mathrm{PRE}_{S^1}(k/3)\mathrm{PRE}_{S^2}(k/3) \leq \frac{C_1 C_2}{h_1(k/24) h_2(k/24)}$$

and similarly for $k = 1, 4, 7, \dots$ and $k = 2, 5, 8, \dots$. By the assumed convergence (2), we get that

$$\sum_{k=1}^\infty \mathrm{PRE}_S(k) < \infty.$$

By Lemma 5.1, we thus have for some $\tilde{C} < \infty$ and $\tilde{\theta} \in (0, 1)$ that

$$\mathrm{P}\big[\#\{n \geq 0\colon S_n = v_n\} \geq l\big] \leq \tilde{C} \tilde{\theta}^l \tag{31}$$

for any sequence $\{v_n\}_{n=0}^\infty$ taking values in $\mathscr{W}_{h_1, h_2}$.

Now pick two paths S and S' in $\mathscr{W}_{h_1, h_2}$ according to $\mu \times \mu$. By considering the x-coordinate, we see that we can have $S_i = S'_j$ only if $|i - j| \leq 2$. By conditioning on S' and applying (31) five times, with $\{v_n\}_{n=0}^\infty$ being the S' sequence delayed by $0, \pm 1, \pm 2$ time units, we get that the probability of having at least n vertices in the intersection of S and S' is less than $C\theta^n$, where we can take $C = 5\tilde{C}$ and $\theta = \tilde{\theta}^{1/5}$. Hence, (30) holds, so that μ has the EIT(θ) property. By Lemma 5.2, we can thus find a $p < 1$ such if we do bond percolation on $\mathscr{W}_{h_1, h_2}$ with retention probability p, then the set of retained edges contains a transient infinite cluster.

It remains to extend the result to all $p \in (p_c, 1)$. First note that what we have done so far easily extends from bond percolation to site percolation (see the remark after Proposition 1.2 in [3]).

Next, we claim that we can find a $\tilde{p} < 1$ with the property that site percolation with retention probability $\tilde{p}$ yields transient infinite clusters on the shrunk wedge $\mathscr{W}_{h_1/N, h_2/N}$ for any $N < \infty$. To see this, pick two functions $\tilde{h}_1(x)$ and $\tilde{h}_2(x)$ such that

$$\sum_{j=1}^{\infty} \frac{1}{\tilde{h}_1(j)\tilde{h}_2(j)} < \infty$$

and having the additional properties that $\lim_{x\to\infty} \tilde{h}_i(x)/h_i(x) = 0$ for $i = 1, 2$ (such functions are easily constructed) and pick $\tilde{p} < 1$ such that $\mathscr{W}_{\tilde{h}_1, \tilde{h}_2}$ gets transient clusters. The claim now follows from Rayleigh's monotonicity principle and the observation that for any N there exists a translate of $\mathscr{W}_{\tilde{h}_1, \tilde{h}_2}$ which is contained in $\mathscr{W}_{h_1/N, h_2/N}$.

For N a multiple of 8, let

$$\mathscr{W}^N_{h_1, h_2} = \left\{ v \in N\mathbb{Z}^3 \colon v + w \in \mathscr{W}_{h_1, h_2} \text{ for all } w \in \left\{ -\frac{5N}{8}, \dots, \frac{5N}{8} \right\}^3 \right\},$$

that is, $\mathscr{W}^N_{h_1, h_2}$ consists of those points of the stretched lattice $N\mathbb{Z}^3$ which sit at the center of a cube of side-length $5N/4$ contained entirely in $\mathscr{W}_{h_1, h_2}$. For $v \in \mathscr{W}^N_{h_1, h_2}$, write $Q_N(v)$ for the cube of side-length $5N/4$ centered at v. For bond percolation on $\mathscr{W}_{h_1, h_2}$ with retention probability p, let $A_p(N)$ be the random set of vertices $v \in \mathscr{W}^N_{h_1, h_2}$ with the property that the set of retained edges in $Q_N(v)$ contains a connected component which connects all six faces of $Q_N(v)$ but contains no other connected component of diameter greater than $N/10$. It follows from Proposition 2.1 in [1] that for any $p > p_c$, the set $A_p(N)$ stochastically dominates site percolation with parameter $p^*(N)$ on $\mathscr{W}^N_{h_1, h_2}$, with $\lim_{N\to\infty} p^*(N) = 1$. Picking N so large that $p^*(N) \geq \tilde{p}$, the proof can now be finished as the proof of Corollary 2.1 in [3]. □

PROOF OF COROLLARY 1.2. We may without loss of generality assume that $h(x)$ does not grow too rapidly: let us for concreteness assume that

$$\lim_{n\to\infty} \frac{h(x)}{\sqrt{x}} = 0. \tag{32}$$

Indeed, if $h(x)$ fails (32), then we can instead prove the result for the wedge $\mathscr{W}_g$ with

$$g(x) = \min\{h(x), x^{1/4}\}$$

and use Rayleigh's monotonicity principle to carry over to the case of $\mathscr{W}_h$; it is easy to see that $g(x)$ satisfies (32) as well as the assumptions of the corollary.

If

$$\frac{h(x)}{x^2} \quad \text{is decreasing in } x, \tag{33}$$

then we can set $h_1(x) = x/\sqrt{h(x)}$ and $h_2(x) = h(x)$, and apply Theorem 1.1 to get that the wedge $\mathscr{W}_{h_1,h_2}$ contains a transient infinite cluster for all $p > p_c$. Since $\mathscr{W}_{h_1,h_2} \subset \mathscr{W}_h$, we can use Rayleigh's monotonicity principle to obtain the same conclusion for $\mathscr{W}_h$.

If (33) happens to fail, then the result does not follow directly from Theorem 1.1, but rather from its proof. The only modification of the proof which is needed is in the choice of nearest-neighbor walks S^1 and S^2. If S^1 is chosen in such a way that $\mathrm{PRE}_{S^1}(k) \le C_1\sqrt{h(x)}/x$ (this is possible by Theorem 1.4) and S^2 is chosen as before with $h_2(x) = h(x)$, then the rest of the proof of Theorem 1.1 goes through for proving the corollary. □

6. Discussion. Lyons [13] has an exact condition on h for simple random walk on $\mathscr{W}_h$ to be transient, namely that

$$\sum_{j=1}^{\infty} \frac{1}{jh(j)} < \infty.$$

An obvious question given the results of the present paper is whether transience survives under independent thinning of the edge set for all such wedges. Corollary 1.2 does not quite cover all such cases. For instance, taking $h(x) = (\log(1+x))^{1+\varepsilon}$ for $\varepsilon \in (0,1]$ yields a wedge for which Lyons' condition yields transience but for which Corollary 1.2 does not apply to give transience after thinning. The result of Hoffman mentioned in the introduction indicates that the approach in Section 5 may be difficult to adapt in order to strengthen Corollary 1.2 in this direction.

It is worth noting that the GKZ theorem cannot be extended to general transient graphs. For instance, there exists a graph G such that (i) simple random walk on G is transient, (ii) the critical value for percolation on G is strictly less than 1, whereas (iii) for any $p < 1$, percolation on G with parameter p yields a.s. no infinite cluster on which simple random walk is transient. A somewhat artificial example (which we attribute to mathematical folklore) of such a graph is to take the square lattice Z^2 and attach to the origin a tree T whose growth rate is chosen in such a way that the critical value for percolation on T is 1 while at the same time simple random walk on T is transient.

Some open problems have already been mentioned. A more extensive list of open problems in this area can be found in [3].

Acknowledgments. Both authors are grateful to Yuval Peres for suggesting the problems, for encouragement and for many helpful discussions.

They have also benefited from discussions with Geoffrey Grimmett and Harry Kesten.

REFERENCES

[1] ANTAL, P. and PISZTORA, A. (1996). On the chemical distance in supercritical Bernoulli percolation. *Ann. Probab.* 24 1036–1048.

[2] BENJAMINI, I., PEMANTLE, R. and PERES, Y. (1996). Random walks in varying dimensions, *J. Theoret. Probab.* 9 221–244.

[3] BENJAMINI, I., PEMANTLE, R. and PERES, Y. (1998). Unpredictable paths and percolation, *Ann. Probab.* 26 1198–1211.

[4] BLEHER, P. M., RUIZ, J. and ZAGREBNOV, V. A. (1995). On the purity of the limiting Gibbs state for the Ising model on the Bethe lattice. *J. Statist. Phys.* 79 473–482.

[5] CHAYES, J. T. and CHAYES, L. (1986). Critical points and intermediate phases on wedges of $\mathbb{Z}^d$. *J. Phys. A* 19 3033–3048.

[6] DOYLE, P. and SNELL, J. L. (1984). *Random Walks and Electric Networks*. Mathematical Monograph 22. Math. Assoc. Amer., Washington, D.C.

[7] GRIMMETT, G. (1989). *Percolation*. Springer, New York.

[8] GRIMMETT, G., KESTEN, H. and ZHANG, Y. (1993). Random walk on the infinite cluster of the percolation model. *Probab. Theory Related Fields* 96 33–44.

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

[10] HOFFMAN, C. (1997). Unpredictable nearest neighbor processes. Preprint.

[11] LEVIN, D. and PERES, Y. (1998). Energy and cutsets in infinite percolation clusters. In *Proceedings of the Cortona Workshop on Random Walks and Discrete Potential Theory* (M. Piccardello and W. Woess, eds.). To appear.

[12] LYONS, R. (1989). The Ising model and percolation on trees and tree-like graphs. *Comm. Math. Phys.* 125 337–353.

[13] LYONS, T. (1983). A simple criterion for transience of a reversible Markov chain. *Ann. Probab.* 11 393–402.

[14] POLYA, G. (1921). Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irfahrt im Strassennetz. *Math. Ann.* 84 149–160.

[15] RÉVÉSZ, P. (1990). *Random Walk in Random and Non-random Environments*. World Scientific, Singapore.

[16] SCOTT, D. (1990). A non-integral-dimensional random walk. *J. Theoret. Probab.* 3 1–7.

[17] SOLOMON, F. (1975). Random walks in a random environment. *Ann. Probab.* 3 1–31.

[18] TORREZ, W. C. (1978). The birth and death chain in a random environment: instability and extinction theorems. *Ann. Probab.* 6 1026–1043.

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
412 96 GÖTEBORG
SWEDEN
E-MAIL: olleh@math.chalmers.se

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
GIVAT RAM
JERUSALEM 91904
ISRAEL
E-MAIL: mossel@math.huji.ac.il