The Annals of Probability
1998, Vol. 26, No. 3, 1198–1211

UNPREDICTABLE PATHS AND PERCOLATION

BY ITAI BENJAMINI, ROBIN PEMANTLE[1] AND YUVAL PERES[2]

Weizmann Institute, University of Wisconsin and Hebrew University

We construct a nearest-neighbor process $\{S_n\}$ on $\mathbf{Z}$ that is less predictable than simple random walk, in the sense that given the process until time n, the conditional probability that $S_{n+k} = x$ is uniformly bounded by $Ck^{-\alpha}$ for some $\alpha > 1/2$. From this process, we obtain a probability measure μ on oriented paths in $\mathbf{Z}^3$ such that the number of intersections of two paths, chosen independently according to μ, has an exponential tail. (For $d \geq 4$, the uniform measure on oriented paths from the origin in $\mathbf{Z}^d$ has this property.) We show that on *any* graph where such a measure on paths exists, oriented percolation clusters are transient if the retention parameter p is close enough to 1. This yields an extension of a theorem of Grimmett, Kesten and Zhang, who proved that supercritical percolation clusters in $\mathbf{Z}^d$ are transient for all $d \geq 3$.

1. Introduction. An oriented path from the origin in the lattice $\mathbf{Z}^d$ is determined by a sequence of vertices $\{y_n\}_{n\geq 0}$ where $y_0 = 0$ and for each $n \geq 1$, the increment $y_n - y_{n-1}$ is one of the d standard basis vectors. When these increments are chosen independently and uniformly among the d possibilities, we refer to the resulting random path as a *uniform random oriented path*. For $d \geq 4$, the number of intersections of two (independently chosen) uniform random oriented paths in $\mathbf{Z}^d$ has an exponentially decaying tail. Cox and Durrett (1983) used this fact to obtain upper bounds on the critical probability p_c^{or} for oriented percolation. (They attribute this idea to H. Kesten.)

In $\mathbf{Z}^3$, however, two independently chosen uniform random oriented paths have infinitely many intersections a.s. Perhaps surprisingly, there is a different measure on oriented paths in $\mathbf{Z}^3$ with exponential tail for the intersection number (see Theorem 1.3 below). The usefulness of such a measure goes beyond estimates for p_c^{or}, since on *any* graph, its existence implies that for p close enough to 1, a.s. some infinite cluster for oriented percolation is transient for simple random walk (see Proposition 1.2 below). In particular, for sufficiently large p, oriented clusters are transient in $\mathbf{Z}^d$ for all $d \geq 3$. This extends a theorem of Grimmett, Kesten and Zhang (1993), who established transience of simple random walk on the infinite cluster of ordinary percola-

Received February 1997; revised March 1998.
[1]Supported in part by NSF Grant DMS-93-00191, a Sloan Foundation Fellowship and a Presidential Faculty Fellowship.
[2]Supported in part by NSF Grant DMS-94-04391.
AMS 1991 *subject classifications.* Primary 60J45, 60J10; secondary 60J65, 60J15, 60K35.
Key words and phrases. Percolation, transience, electrical networks, multitype branching process.

tion in $\mathbf{Z}^d$, $d \geq 3$ (They obtain transience for *all* $p > p_c$ but in $\mathbf{Z}^d$ this can be reduced to the case of large p by renormalization; see Section 2).

We construct the required measure in three dimensions from certain nearest-neighbor stochastic processes on $\mathbf{Z}$ which are "less predictable than simple random walk."

DEFINITION. For a sequence of random variables $S = \{S_n\}_{n \geq 0}$ taking values in a countable set V, we define its *predictability profile* $\{\mathsf{PRE}_S(k)\}_{k \geq 1}$ by

$$\mathsf{PRE}_S(k) = \sup \mathbf{P}\left[S_{n+k} = x \mid S_0, \dots, S_n\right], \tag{1}$$

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

Thus $\mathsf{PRE}_S(k)$ is the maximal chance of guessing S correctly k steps into the future, given the past of S. Clearly, the predictability profile of simple random walk on $\mathbf{Z}$ is asymptotic to $Ck^{-1/2}$ for some $C > 0$.

THEOREM 1.1. (a) *For any $\alpha < 1$ there exists an integer-valued stochastic process $\{S_n\}_{n \geq 0}$ such that $|S_n - S_{n-1}| = 1$ a.s. for all $n \geq 1$ and*

$$\mathsf{PRE}_S(k) \leq C_\alpha k^{-\alpha} \quad \textit{for some } C_\alpha < \infty, \textit{ for all } k \geq 1. \tag{2}$$

(b) *Furthermore, there exists such a process where the ± 1 valued increments $\{S_n - S_{n-1}\}$ form a stationary ergodic process.*

Part (b) is not needed for the applications in this paper and is included because such processes may have independent interest. The approach that naturally suggests itself to obtain processes with a low predictability profile is to use a discretization of fractional Brownian motion. However, we could not turn this idea into a rigorous construction. Instead, we construct the processes described in Theorem 1.1 from a variant of the Ising model on a regular tree, by summing the spins along the boundary of the tree (see Section 4). This may be a case of the principle "when you have a hammer, everything looks like a nail," and we would be interested to see alternative constructions.

DEFINITIONS. (i) Let $G = (V_G, E_G)$ be an infinite directed graph with all vertices of finite degree and let $v_0 \in V_G$. Denote by $\Upsilon = \Upsilon(G, v_0)$ the collection of infinite directed paths in G which emanate from v_0 and tend to infinity (i.e., the paths in Υ visit any vertex at most infinitely many times). The set $\Upsilon(G, v_0)$, viewed as a subset of $E_G^{\mathbf{N}}$, is a Borel set in the product topology.

(ii) Let $0 < \theta < 1$. A Borel probability measure μ on $\Upsilon(g, v_0)$ has *exponential intersection tails* with parameter θ [in short, EIT(θ)] if there exists C such that

$$\mu \times \mu\{(\varphi, \psi) : |\varphi \cap \psi| \geq n\} \leq C\theta^n \tag{3}$$

for all n, where $|\varphi \cap \psi|$ is the number of edges in the intersection of φ and ψ.

(iii) If such a measure μ exists for some basepoint v_0 and some $\theta < 1$, then we say that G *admits random paths with* $\mathrm{EIT}(\theta)$. Analogous definitions apply to undirected graphs.

(iv). *Oriented percolation* with parameter $p \in (0,1)$ on the directed graph G is the process where each edge of G is independently declared *open* with probability p and *closed* with probability $1 - p$. The union of all directed open paths emanating from v will be called the *oriented open cluster* of v and denoted $C(v)$.

(v) A subgraph Λ of G is called *transient* if, when the orientations on the edges are ignored, Λ is connected and simple random walk on it is a transient Markov chain. As explained in Doyle and Snell (1984), the latter property is equivalent to finiteness of the effective resistance from a vertex of Λ to infinity, when each edge of Λ is endowed with a unit resistor.

PROPOSITION 1.2. *Suppose a directed graph G admits random paths with* $\mathrm{EIT}(\theta)$. *Consider oriented percolation on G with parameter p. If $p > \theta$ then with probability* 1 *there is a vertex v in G such that the directed open cluster $C(v)$ is transient.*

REMARK. The proof of the proposition, given in Section 2, also applies to site percolation. If the graph G is *undirected* and admits random (undirected) paths with exponential intersection tails, then the same proof shows that for p close enough to 1, a.s. some infinite cluster of ordinary percolation on G is transient.

Recall that a path $\{\Gamma_n\}$ in $\mathbf{Z}^d$ is called *oriented* if each increment $\Gamma_{n+1} - \Gamma_n$ is one of the d standard basis vectors. The difference of two independent, uniformly chosen, oriented paths in $\mathbf{Z}^d$ is a random walk with increments generating the $d-1$ dimensional hyperplane $\{\sum_{i=1}^d x_i = 0\}$. As noted by Cox and Durrett (1983), it follows that the uniform measure on oriented paths in $\mathbf{Z}^d$ has $\mathrm{EIT}(\theta_d)$. Clearly, a different approach is needed for $d = 3$.

THEOREM 1.3. *There exists a measure on oriented paths from the origin in* $\mathbf{Z}^3$ *that has exponential intersection tails.*

The rest of this paper is organized as follows. In Section 2 we prove Proposition 1.2 by constructing a flow of finite energy on the percolation cluster. For $d \geq 4$, this yields a "soft" proof that if the parameter p is close enough to 1, then oriented percolation clusters in $\mathbf{Z}^d$ are transient with positive probability; for $d = 3$, Theorem 1.3 is needed to obtain the same conclusion. We also explain there how transience of ordinary percolation clusters in $\mathbf{Z}^d$ for $d \geq 3$ and *all* $p > p_c$ can be reduced to transience of these clusters for p close to 1. In Section 3 we relate the predictability profile of a random process to the tail of its collision number with a fixed sequence, and establish Theorem 1.3. Theorem 1.1 is proved in Section 4. The main ingredient in the proof is an estimate on the distribution of the population vector in a certain two-type branching process, when this vector is projected in a nonprincipal eigendirection. Section 5 contains auxiliary remarks and problems. After this paper was submitted, the methods introduced here were refined and extended by several different authors, and some of the questions we raised were solved. The paper concludes with a brief survey of these recent developments.

2. Exponential intersection tails imply transience of clusters. To show that an infinite connected graph Λ is transient, it suffices to construct a nonzero flow f on Λ, with a single source at v_0, such that the *energy* $\sum_{e \in E_\Lambda} f(e)^2$ of f is finite [see Doyle and Snell (1984)].

PROOF OF PROPOSITION 1.2. The hypothesis means that there is some vertex v_0 and a probability measure μ on $\Upsilon = \Upsilon(g, v_0)$ that has EIT(θ). We first assume that

(4) *The paths in the closed support of μ are self-avoiding and tend to infinity uniformly.*

A path is *self-avoiding* if it never revisits a vertex; the second part of the assumption means that there is a sequence $r_N \to \infty$ such that for all $N \geq 1$ and all paths φ in the support of μ, the endpoint of φ_N is not in $B(v_0, r_N)$, where $B(v_0, r)$ denotes the ball of radius r centered at v_0 in the usual graphical distance. The assumption (4) certainly holds in our main application, where μ is supported on oriented paths in $G = \mathbf{Z}^d$; at the end of the proof we show how to remove the assumption (4).

For $N \geq 1$ and any infinite path $\varphi \in \Upsilon(G, v_0)$, denote by φ_N the finite path consisting of the first N edges of φ. Consider the random variable

$$Z_N = \int_\Upsilon p^{-N} \mathbf{1}_{\{\varphi_N \text{ is open}\}} \, d\mu(\varphi). \tag{5}$$

Except for the normalization factor p^{-N}, this is the μ-measure of the paths that stay in the open cluster of v_0 for N steps.

Since each edge is open with probability p (independently of other edges), $\mathbf{E}(Z_N) = 1$. The second moment of Z_N is

$$\begin{aligned} \mathbf{E}(Z_N^2) &= \mathbf{E} \int_\Upsilon \int_\Upsilon p^{-2N} \mathbf{1}_{\{\varphi_N \text{ and } \psi_N \text{ are open}\}} \, d\mu(\varphi) \, d\mu(\psi) \\ &\leq \int_\Upsilon \int_\Upsilon p^{-|\varphi \cap \psi|} \, d\mu(\varphi) \, d\mu(\psi). \end{aligned} \tag{6}$$

By (3), the last integral is at most $\sum_{n=0}^\infty C\theta^n p^{-n} = C/(1 - \theta p^{-1})$.

By Cauchy–Schwarz,

$$\mathbf{P}\big[|C(v_0)| \geq N\big] \geq \mathbf{P}[Z_N > 0] \geq \frac{\mathbf{E}(Z_N)^2}{\mathbf{E}(Z_N^2)} \geq \frac{1 - \theta p^{-1}}{C}. \tag{7}$$

This shows that the cluster $C(v_0)$ is infinite with positive probability.

The next step is to construct a flow f of finite mean energy on $C(v_0)$. For each $N \geq 1$ and every directed edge e in E_G, we define

$$f_N(e) = \int_{\Upsilon} p^{-N} \mathbf{1}_{\{\varphi_N \text{ is open}\}} \mathbf{1}_{\{e \in \varphi_N\}} \, d\mu(\varphi). \tag{8}$$

Then f_N is a flow on $C(v_0) \cap B(v_0, r_N + 1)$ from v_0 to the complement of $B(v_0, r_N)$, that is, for any vertex $v \in B(v_0, r_N)$ except v_0, the incoming flow to v equals the outgoing flow from v. The *strength* of f_N (the total outflow from v_0) is exactly Z_N.

Next, we estimate the expected energy of f_N:

$$\begin{aligned} \mathbf{E} \sum_{e \in E_G} f_N(e)^2 &= \mathbf{E} \int_{\Upsilon} \int_{\Upsilon} p^{-2N} \mathbf{1}_{\{\varphi_N, \psi_N \text{ are open}\}} \sum_{e \in E_G} \mathbf{1}_{\{e \in \varphi_N\}} \mathbf{1}_{\{e \in \psi_N\}} \, d\mu(\varphi) \, d\mu(\varphi) \\ &= \int_{\Upsilon} \int_{\Upsilon} |\varphi \cap \psi| \, p^{-|\varphi \cap \psi|} \, d\mu(\varphi) \, d\mu(\psi). \end{aligned}$$

Another application of exponential intersection tails allows us to bound the last integral by

$$\sum_{n=0}^{\infty} C\theta^n n p^{-n} \tag{10}$$

which is finite for $p > \theta$. For each edge e of G, the sequence $\{f_N(e)\}$ is bounded in L^2, so it has a weakly convergent subsequence. Using the diagonal method, we can find a single increasing sequence $\{N(k)\}_{k \geq 1}$ such that for every edge e, the sequence $f_{N(k)}(e)$ converges weakly as $k \to \infty$ to a limit, denoted $f(e)$. Recalling that Z_N is the strength of f_N, we deduce that $f_{N(k)}(e)$ converges weakly as $k \to \infty$ to a limit, denoted Z_∞. Since $r_N \to \infty$, the limit function f is a.s. a flow of strength Z_∞ on $C(v_0)$. Exhausting G by finite sets of edges, we conclude that the expected energy of f is also bounded by (10). Thus

$$\mathbf{P}[C(v_0) \text{ is transient}] \geq \mathbf{P}[Z_\infty > 0] > 0,$$

so the tail event $[\exists\ v\colon C(v)$ is transient$]$ must have probability 1 by Kolmogorov's zero–one law.

Finally, we remove the assumption (4). Any path tending to infinity contains a self-avoiding path, obtained by "loop-erasing" (erasing cycles as they are created); see Chapter 7 in Lawler (1991). Thus we may indeed assume that μ is supported on self-avoiding paths. Since all paths in Υ tend to infinity, by Egorov's theorem there is a closed subset Υ' of Υ on which this convergence is uniform, such that $\mu(\Upsilon') > \mu(\Upsilon)/2$. Restricting μ to Υ' and normalizing, we obtain a probability measure μ' on Υ' that satisfies (4) and (3) with $4C$ in place of C, so the proof given above applies. □

REMARK. Let $\Upsilon_1 = \Upsilon_1(G, v_0) \subset \Upsilon$ denote the set of *paths with unit speed*, that is, those paths such that the nth vertex is at distance n from v_0, for every n. In most applications of Proposition 1.2, the measure μ is supported on Υ_1. When that is the case, the flows f_N considered in the preceding proof converge a.s. to a flow f, so there is no need to pass to subsequences. Indeed, let B_N be the σ-field generated by the status (open or closed) of all edges on paths φ_N with $\varphi \in \Upsilon$. It is easy to check that $\{Z_N\}_{N\geq 1}$ is an L^2 martingale adapted to the filtration $\{B_N\}_{N\geq 1}$. Therefore $\{Z_N\}$ converges a.s. and in L^2 to a mean 1 random variable Z_∞. Moreover, for each edge e of G, the sequence $\{f_N(e)\}$ is a $\{B_N\}$-martingale which converges a.s. and in L^2 to a nonnegative random variable $f(e)$.

COROLLARY 2.1 [Grimmett, Kesten and Zhang (1993)]. *Consider ordinary bond percolation with parameter p on $\mathbf{Z}^d$, where $d \geq 3$. For all $p > p_c$, the unique infinite cluster is a.s. transient.*

PROOF. Transience of the unique infinite cluster is a tail event, so it has probability 0 or 1. Since $\mathbf{Z}^d$ admits random paths with EIT for $d \geq 3$ (see Theorem 1.3 and the discussion preceding it), it follows from Proposition 1.2 that the infinite cluster is transient if p is close enough to 1. As remarked before, this conclusion also applies to site percolation.

Recall that a set of graphs $\mathbf{B}$ is called *increasing* if for any graph G that contains a subgraph in $\mathbf{B}$, necessarily G must also be in $\mathbf{B}$.

Consider now bond percolation with *any* parameter $p > p_c$ in $\mathbf{Z}^d$. Following Pisztora (1996), call an open cluster C contained in some cube Q a *crossing cluster* for Q if for all d directions there is an open path contained in C joining the left face of Q to the right face. For each v in the lattice $N\mathbf{Z}^d$, denote by $Q_N(v)$ the cube of side-length $5N/4$ centered at v, and let $A_p(N)$ be the set of $v \in N\mathbf{Z}^d$ with the following property: *the cube $Q_N(v)$ contains a crossing cluster C such that any open cluster in $Q_N(v)$ of diameter greater than $N/10$ is connected to C by an open path in $Q_N(v)$.*

Proposition 2.1 in Antal and Pisztora (1996), which relies on the work of Grimmett and Marstrand (1990), implies that $A_p(N)$ stochastically dominates site percolation with parameter $p^*(N)$ on the stretched lattice $N\mathbf{Z}^d$, where $p^*(N) \to 1$ as $N \to \infty$. [Related renormalization arguments can be found in Kesten and Zhang (1990) and Pisztora (1996); general results on domination by product measures where obtained by Liggett, Schonmann and Stacey (1996)]. This domination means that for any increasing Borel set of graphs $\mathbf{B}$, the probability that the subgraph of open sites under independent site percolation with parameter $p^*(N)$ lies in $\mathbf{B}$, is at most $\mathbf{P}[A_p(N) \in \mathbf{B}]$. If N is sufficiently large, then the infinite cluster determined by the site percolation with parameter $p^*(N)$ on the lattice $N\mathbf{Z}^d$, is transient a.s. By Rayleigh's monotonicity principle [see Doyle and Snell (1984)], the set of subgraphs of $N\mathbf{Z}^d$ that have a transient connected component is increasing, so $A_p(N)$ has a transient component $\hat{A}_n(N)$ with probability 1.

Recall from Doyle and Snell (1984) that the "k-fuzz" of a graph $\Gamma = (V, E)$ is the graph $\Gamma_k = (V, E_k)$ where the vertices $v, w \in V$ are connected by an edge in E_k iff there is a path of length at most k between them in Γ. The k-fuzz Γ_k is transient iff Γ is transient [See Section 8.4 in Doyle and Snell

(1984), or Lemma 7.5 in Soardi (1994).] By assigning to each $v \in N\mathbf{Z}^d$ a different vertex $F(v)$ in the intersection $C_p \cap Q_N(v)$, we see that $\hat{A}_p(N)$ is isomorphic to a subgraph of the $2(5N/4)^d$-fuzz of the infinite cluster C_p in the original lattice. It follows that C_p is also transient a.s. □

3. A summable predictability profile yields EIT. The following lemma will imply Theorem 1.3.

LEMMA 3.1. *Let $\{\Gamma_n\}$ be a sequence of random variables taking values in a countable set V. If the predictability profile [defined in (1)] of Γ satisfies $\sum_{k=1}^{\infty} \mathsf{PRE}_\Gamma(k) < \infty$, then there exist $C < \infty$ and $0 < \theta < 1$ such that for any sequence $\{v_n\}_{n \geq 0}$ in V and all $l \geq 1$,*

$$\mathbf{P}\big[\#\{n \geq 0: \Gamma_n = v_n\} \geq l\big] \leq C\theta^l. \tag{11}$$

PROOF. Choose m large enough so that $\sum_{k=1}^{\infty} \mathsf{PRE}_\Gamma(km) = \beta < 1$, whence for any sequence $\{v_n\}_{n \geq 0}$,

$$\mathbf{P}\big[\exists\ k \geq 1: \Gamma_{n+km} = v_{n+km} \big| \Gamma_0, \dots, \Gamma_n\big] \leq \beta \quad \text{for all } n \geq 0.$$

It follows by induction on ≥ 1 that for all $j \in [0, 1, \dots, m-1\}$,

$$\mathbf{P}\Big[\#\big\{k \geq 1: \Gamma_{j+km} = v_{j+km}\big\} \geq r\Big] \leq \beta^r. \tag{12}$$

If $\#\{n \geq 0: \Gamma_n = v_n\} \geq l$ then there must be some $j \in \{0, 1, \dots, m-1\}$ such that

$$\#\big\{k \geq 1: \Gamma_{j+km} = v_{j+km}\big\} \geq l/m - 1.$$

Thus the inequality (11), with $\theta = \beta^{1/m}$ and $C = m\beta^{-1}$, follows from (12). □

PROOF OF THEOREM 1.3. Let $\{S_n\}_{n \geq 0}$ be a nearest-neighbor process on $\mathbf{Z}$ starting from $S_0 = 0$, that satisfies (2) for some $\alpha > 1/2$ and $C_\alpha < \infty$. Denote $W_n = (n + S_n)/2$ and suppose that the processes $\{W_n\}$ and $\{W_n^\sharp\}$ are independent of each other and have the same distribution. Clearly $\{W_n\}$ and $\{S_n\}$ have the same predictability profile. We claim that *the random oriented path*,

$$\{\Gamma_n\}_{n \geq 0} = \Big\{\Big(W_{\lfloor n/2 \rfloor}, W^\sharp_{\lceil n/2 \rceil}, n - W_{\lfloor n/2 \rfloor} - W^\sharp_{\lceil n/2 \rceil}\Big)\Big\}_{n \geq 0}$$

in $\mathbf{Z}^3$ has exponential intersection tails.

First, observe that $\{\Gamma_n\}$ is indeed an oriented path; that is, $\Gamma_{n+1} - \Gamma_n$ is one of the three vectors $(1,0,0), (0,1,0), (0,0,1)$ for every n. Second, $\mathsf{PRE}_\Gamma(k) = \mathsf{PRE}_S(\lfloor k/2 \rfloor)^2 \leq C_\alpha^2 \lfloor k/2 \rfloor^{-2\alpha}$ is summable in k, so $\{\Gamma_n\}$ satisfies (11) for some $C < \infty$ and $0 < \theta < 1$. Denote the distribution of $\{\Gamma_n\}$ by μ, and let $\{\Gamma^*\}$ be an independent copy of $\{\Gamma_n\}$. Integrating (11) with respect to μ, we get

$$\forall\ l \qquad \mu \times \mu\big[\#\{n \geq 0: \Gamma_n = \Gamma_n^*\} \geq l\big] \leq C\theta^l.$$

For two oriented paths Γ and Γ' in $\mathbf{Z}^d$, the number of edges in common is at most the collision number $\#\{n \geq 0: \Gamma_n = \Gamma^*\}$ and hence μ has EIT(θ). □

4. Summing boundary spins yields an unpredictable path. In this section we prove Theorem 1.1. The engine of the proof is Lemma 4.1 concerning the distribution of the population in a two-type branching process. Let $l \geq 2$ and $r \geq 1$ be integers and write $h = l + r$. Denote by $\mathbf{T}_b$ the infinite rooted tree where each vertex has exactly b children. Consider the following labeling $\{\sigma(v)\}$ of the vertices of $\mathbf{T}_b$ by ± 1-valued random variables, called *spins* because of the analogy with the Ising model [cf. Moore and Snell (1979)]. Let $\sigma(\text{root}) = 1$. For any vertex v of $\mathbf{T}_b$ with children $w_1, \dots, w_b$, assign the first l children the same spin as their parent: $\sigma(w_j) = \sigma(v)$ for $j = 1, \dots, l$ and assign the other r children i.i.d. spins $\sigma(w_{l+1}), \dots, \sigma(w_b)$ that take the values ± 1 with equal probability and are independent of $\sigma(v)$. As N varies, the population vectors (Z_N^+, Z_N^-), which count the number of spins of each type at level N of $\mathbf{T}_b$, form a two-type branching process with mean offspring matrix

$$\begin{pmatrix} l + r/2 & r/2 \\ r/2 & l + r/2 \end{pmatrix}.$$

[See Athreya and Ney (1972) for background on branching processes.] The Perron eigenvalue of this matrix is b, but we are interested in the scalar product of the population vectors with the eigenvector $(1, -1)$ which corresponds to the second eigenvalue l of the mean offspring matrix.

LEMMA 4.1. *The sum of all spins at level N,*

$$Y_N = \sum_{|v| = N} \sigma(v) = Z_N^+ - Z_N^-, \tag{13}$$

satisfies the inequality

$$\mathbf{P}[Y_N = x] \leq C l^{-N} \tag{14}$$

for all $N \geq 1$ and all integer x, where the constant C depends only on l and r.

REMARK. A closer examination of the proof below shows that the inequality (14) is sharp iff $l > \sqrt{b}$. The significance of the condition $l > \sqrt{b}$ is explained by Kesten and Stigum (1966) in a more general setting. If $l < \sqrt{b}$ then the distributions of $Y_N b^{-N/2}$ converge to a normal law.

PROOF OF LEMMA 4.1. By decomposing the sum in the definition of Y_{N+1} into b parts according to the first level of $\mathbf{T}_b$, we see that

$$Y_{N+1} = \sum_{j=1}^{l} Y_N^{(j)} + \sum_{j=l+1}^{b} \sigma(w_j) Y_N^{(j)}, \tag{15}$$

where $\{\sigma(w_j)\}_{j=l+1}^{b}$ are r i.i.d. uniform spins, 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

$$\hat{Y}_N(\lambda) = \mathbf{E}\big(\exp(i\lambda Y_N)\big) \tag{16}$$

satisfy the recursion

$$\hat{Y}_{N+1}(\lambda) = \hat{Y}_N(\lambda)^l \left(\frac{\hat{Y}_N(\lambda) + \hat{Y}_N(-\lambda)}{2} \right)^r = \hat{Y}_N(\lambda)^l \left(\Re\, \hat{Y}_N(\lambda) \right)^r, \tag{17}$$

where $\Re$ denotes real part. Using the polar representation $\hat{Y}_N = |\hat{Y}_N| \cdot \exp(i\gamma_N(\lambda))$, the last equation implies that $\gamma_{N+1}(\lambda) \equiv l\gamma_N(\lambda) \bmod \pi$. [Note that $\Re\, \hat{Y}_N(\lambda)$ may be negative.] By definition $Y_0 = 1$, and therefore $\hat{Y}_0(\lambda) = e^{i\lambda}$, so $\gamma_0(\lambda) = \lambda$. Consequently, $\gamma_N(\lambda) \equiv l^N \lambda \bmod \pi$ for all N. Taking absolute values in (17) yields

$$\left| \hat{Y}_{N+1}(\lambda) \right| = \left| \hat{Y}_N(\lambda) \right|^b \left| \cos(l^N \lambda) \right|^r. \tag{18}$$

By induction on N, we obtain

$$\forall\, N \ge 0 \qquad \left| \hat{Y}_N(\lambda) \right| = \prod_{k=1}^{N} \left| \cos(l^{N-k} \lambda) \right|^{rb^{k-1}}. \tag{19}$$

Since $|\hat{Y}_N(\cdot)|$ is an even function with period π, changing variables $t = l^N \lambda$ gives

$$\begin{aligned} \int_{-\pi}^{\pi} \left| \hat{Y}_N(\lambda) \right| d\lambda &= 4 \int_0^{\pi/2} \left| \hat{Y}_N(\lambda) \right| d\lambda \\ &= 4 l^{-N} \int_0^{(l^N \pi)/2} \prod_{k=1}^{N} \left| \cos(l^{-k} t) \right|^{rb^{k-1}} dt. \end{aligned} \tag{20}$$

Denote $\xi = \cos(\pi/2l)$. For $t \in [l^{k-1}\pi/2, l^k \pi/2]$, the kth factor in the rightmost integrand in (20) is bounded by $\xi^{rb^{k-1}}$, and the other factors are at most 1. Consequently,

$$\int_{-\pi}^{\pi} \left| \hat{Y}_N(\lambda) \right| d\lambda \le 4 l^{-N} \left(\frac{\pi}{2} + \sum_{k=1}^{N} \frac{l^k \pi}{2} \xi^{rb^{k-1}} \right) \le C l^{-N}, \tag{21}$$

where C depends only on l and r, since the sum $\sum_{k=1}^{\infty} l^k \xi^{rb^{k-1}}$ converges. Finally, Fourier inversion and (21) imply that for any integer x,

$$\mathbf{P}[Y_N = x] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{Y}_N(\lambda) e^{-i\lambda x}\, d\lambda \le \int_{-\pi}^{\pi} \left| \hat{Y}_N(\lambda) \right| d\lambda \le C l^{-N}. \qquad \square$$

PROOF OF THEOREM 1.1. (a) Given integers $l \ge 2$ and $r \ge 1$, let $b = l + r$ as above. Embed the regular tree $\mathbf{T}_b$ in the upper half-plane, with the root on the real line and the children of every vertex arranged from left to right above it. Label the vertices of $\mathbf{T}_b$ by ± 1-valued spins $\{\sigma(v)\}$ as described at the beginning of this section. Let $M > 1$ and suppose that $b^N \ge M$. Let $\{v_j\}_{j=1}^{b^N}$ be the vertices at level N of $\mathbf{T}_b$, enumerated from left to right. For $m \le M$, denote $S_m = \sum_{j=1}^{m} \sigma(v_j)$ and observe that the joint distribution of the M random variables $\{S_m\}_{m=1}^{M}$ does not depend on N. Using Kolmogorov's consistency theorem, we obtain an infinite process $\{S_m\}_{m=1}^{\infty}$. We claim that the predictability profile of this process satisfies

$$\mathrm{PRE}_S(k) \le (2b)^{\alpha} C k^{-\alpha} \quad \text{for all } k \ge 1, \tag{22}$$

where $\alpha = \log l / \log b$, and $C = C(l, r) \geq 1$ is given in (14). Since we can take l arbitrarily large and $r = 1$, establishing (22) for $k > 2$ will suffice to prove the theorem.

Given $n \geq 0$ and $k \geq 1$, choose n such that $b^N \geq n + k$, so the random variables $\{S_j\}_{j=0}^{n+k}$ may be obtained by summing spins along level N of $\mathbf{T}_b$. There is a unique $h \geq 0$ such that

$$2b^h \leq k < 2b^{h+1}. \tag{23}$$

For any vertex v, denote by $|v|$ its level in $\mathbf{T}_b$, and for $i \geq 1$ let $D_i(v)$ be the set of its b^i descendants at level $|v| + i$. By (23), there exists at least one vertex v at level $N - h$ of $\mathbf{T}_b$, such that $D_h(v)$ is contained in $\{v_{n+1}, v_{n+2} \dots, v_{n+k}\}$ in the left-to-right enumeration of level N. Denote by $D(v) = \bigcup_{i=1}^{\infty} D_i(v)$ the set of all descendants of v, and by F_v^* the sigma-field generated by all the spins $\{\sigma(w)\colon w \notin D(v)\}$. The random variable

$$\tilde{Y}_h(v) = \sigma(v) \sum_{w \in D_h(v)} \sigma(w)$$

is independent of F_v^* and has the same distribution as the variable Y_H defined by (13). Clearly, we can write

$$S_{n+k} = \sigma(v)\tilde{Y}_h(v) + S^*_{n+k},$$

where S^*_{n+k}, the sum of $n + k - b^h$ spins labeling vertices not in $D_h(v)$, is F_v^*-measurable. Consequently, for any integer x,

$$\mathbf{P}\left[S_{n+k} = x \mid F_v^*\right] = \mathbf{P}\left[\tilde{Y}_h(v) = \sigma(v)(x - S^*_{n+k}) \middle| F_v^*\right] \leq Cl^{-h}, \tag{24}$$

by Lemma 4.1. The definition of α and h imply that $l^{-h} = b^{-h\alpha}$ and $b^h > k/2b$, so we infer from (24) that

$$\forall\, x \in \mathbf{Z} \qquad \mathbf{P}\left[S_{n+k} = x \mid F_v^*\right] < (2b)^{\alpha} Ck^{-\alpha}.$$

Since $S_0, S_1, \dots S_n$ are F_v^*-measurable, this yields (22) and completes the proof of part (a) of the theorem.

(b) The property (2) is stable under shifts, mixtures, weak limits and passing to ergodic components, so it is possible to obtain the desired stationary process as an ergodic component of a weak limit point of the averages $(1/n)(S + \Theta S + \cdots + \Theta^{n-1} S)$, where Θ is the left shift.

We now describe such a process more explicitly, by modifying the construction in part (a). Let $\sigma(\text{root})$ be a uniform random spin; define the other spins from it as in part (a). Given $N > 1$, choose U uniformly in $\{1, \dots, b^N\}$ and define

$$\tilde{S}_n = \sum_{j=U}^{U+n-1} \sigma(v_j) \quad \text{for } n \leq b^N - U + 1,$$

where $\{v_j\}_{j=1}^{b^N}$ is the left-to-right enumeration of level N to $\mathbf{T}_b$. To extend the sequence $\tilde{S}$ further, we consider the root of $\mathbf{T}_b$ as the Jth child w_J of a new vertex ρ, where J is chosen uniformly in $\{1, \dots, b\}$. If $J \leq l$, let $\sigma(\rho) = \sigma(w_J)$, and if $J > l$ let $\sigma(\rho)$ be a uniform random spin, independent of the spins on the original tree. We can view the original tree $\mathbf{T}_b$ as a subtree of a new b-tree $\tilde{\mathbf{T}}_b$ rooted at ρ. Since $(J-1)b^N + U$ is uniformly distributed in $\{1, \dots, b^{N+1}\}$, the vertex v_U is uniformly distributed in level $N+1$ of $\tilde{\mathbf{T}}_b$. Repeating this rerooting procedure and enlarging N as needed, yields the desired process $\{\tilde{S}_j\}_{j=1}^{\infty}$. The proof given in part (a) also shows that this process has the unpredictability property (2). Stationarity and ergodicity of the increments can be derived from the invariance and ergodicity of the Haar measure on the b-adic integers under the operation of adding 1; we omit the details. □

5. Concluding remarks and questions.

1. Consider the following three properties that an infinite connected graph G may have:
   (i) G admits random paths with EIT.
   (ii) There exists $p < 1$ such that simple random walk is transient on a percolation cluster of G for bond percolation with parameter p.
   (iii) A random walk in random environment on G defined by i.i.d. resistances with any common distribution is almost surely transient.
   In Pemantle and Peres (1996) it is shown that properties (ii) and (iii) are equivalent. Proposition 1.2 of the present paper shows that (i) implies (ii); does (ii) imply (i)? Note that there exist transient trees of polynomial growth [see, e.g., Lyons (1990)], and these cannot admit random paths with EIT since they have $p_c = 1$.
2. Does $\mathbf{Z}^d$ with $d \geq 3$ admit random paths with EIT(θ) for *all* $\theta > p_c$? (This question was suggested to us by Rick Durrett.) A similar equation can be asked for other graphs in place of $\mathbf{Z}^d$, for example, for transient Cayley graphs. A positive answer to this question when the graph in question is a *tree* follows from the work of Lyons (1990). Indeed, a flow from the root of the tree can be identified with a measure on paths, and the energy of the flow μ in the kernel $p^{-|x \wedge y|}$ can be identified with an exponential moment of the number of intersections of two paths chosen independently according to μ.
3. Lyons (1995) finds a tree with $p_c < 1$ contained as a subgraph in the Cayley graph of any group of exponential growth. It follows that such Cayley graphs admit paths with EIT.
4. It is easy to adapt the proofs of Theorem 1.3 and Corollary 2.1 to show that for any $\varepsilon > 0$, the cone $\{(x, y, z) \in \mathbf{Z}^3 \colon |z| \leq \varepsilon|x|\}$ in $\mathbf{Z}^3$ admits random paths with EIT and contains a transient percolation cluster for all $p > p_c$. Does the subgraph $\{(x, y, z) \in \mathbf{Z}^3 \colon |z| \leq |x|^{\varepsilon}\}$ share these properties? (This subgraph is sometimes viewed as a model for a "$2 + \varepsilon$ dimensional lattice".)
5. Does oriented percolation in $\mathbf{Z}^d$ admit transient infinite clusters $C(v)$ for all parameters $p > p_c^{\mathrm{or}}$? The challenge here is to adapt the renormalization argument used in the proof of Corollary 2.1 to the oriented setting.

6. Consider the stationary processes $\{\tilde{S}_n\}$ constructed at the end of the previous section, and let $\alpha = l/b$ with $\sqrt{b} < l < b$. Do the rescaled step functions $t \mapsto n^{-\alpha}\tilde{S}_{\lfloor nt \rfloor}$ on $[0,1]$ converge in law? Is the limit a Gaussian process? It is easily verified that $\mathbf{E}|\tilde{S}_n - \tilde{S}_m|^2 \asymp |n-m|^{2\alpha}$, which is reminiscent of fractional Brownian motion. The proof of Lemma 4.1 implies that $b^{-\alpha N} Y_N$ converges in law to a (non-Gaussian) distribution with characteristic function $s \mapsto e^{is}\prod_{k=1}^{\infty}[\cos(l^{-k}s)]^{rb^{k-1}}$. (Recall that $l > \sqrt{b}$).
7. How fast can the predictability profile (1) of a nearest-neighbor process on $\mathbf{Z}$ decay? By Theorem 1.1, a decay rate of $O(k^{-\alpha})$ is possible for any $\alpha < 1$. On the other hand, a decay rate of $O(1/k)$ is impossible. Indeed, if there exists a nearest-neighbor process with predictability profile bounded by $\{g(k)\}$, then there exists such a process with stationary ergodic increments; then $g(k) = O(1/k)$ is ruled out by the ergodic theorem.
8. Among nearest-neighbor processes $\{S_n\}$ on $\mathbf{Z}$, clearly simple random walk has the most unpredictable *increments*, in any conceivable sense. Heuristically, there is a tradeoff here: when the increments are very unpredictable (e.g., their predictability profile tends rapidly to $1/2$), cancellations dominate, and the partial sums becomes more predictable. Our construction in Section 4 sacrificed the independence of the increments to make their partial sums less predictable. It would be quite interesting to establish a precise quantitative form of this tradeoff.
9. Is there a construction of a measure on paths in $\mathbf{Z}^3$ with exponential intersection tails, which is simpler than that given in Section 3 an 4?

5.1. *Recent developments*. After a previous version of the present paper was circulated, some of the problems raised above were solved, and several further extensions of the fundamental transience theorem of Grimmett, Kesten and Zhang (1993) were obtained.

1. Häggström and Mossel (1998) constructed processes with predictability profiles bounded by $C/[k(f(k)]$, for any decreasing f such that $\sum_j f(2^j) < \infty$. They gave two different constructions, one based on the Ising model on trees and the other via a random walk with a random drift that varies in time. Häggström and Mossel also answered affirmatively question 4 above, by constructing paths with exponential intersection tails in "$2+\varepsilon$" dimensions. Remarkably, they were able to show that for a class of "trumpet-shaped" subgraphs G of $\mathbf{Z}^3$, transience of G implies a.s. transience of an infinite percolation cluster in G for any $p > p_c$.
2. In a brief but striking paper, Hoffman (1998) improved the bounds in question 7 above and showed that the constructions of Häggström and Mossel are optimal. Specifically, he used a novel renormalization argument to prove that if f satisfies $\sum_j f(2^j) = \infty$, then there is no nearest-neighbor process on $\mathbf{Z}$ with predictability profile bounded by $C\,[kf(k)]$.
3. Hiemer (1998) proved a renormalization theorem for oriented percolation that allowed him to extend our result on transience of oriented percolation

clusters in $\mathbf{Z}^d$ for $d \geq 3$, from the case of high p to the whole supercritical phase $p > p_c^{\text{or}}$.

4. Consider supercritical percolation on $\mathbf{Z}^d$ for $d \geq 3$. The transience result of Grimmett, Kesten and Zhang (1993) is equivalent to the existence of a nonzero flow f on the infinite cluster such that the two-energy $\sum_e f(e)^2$ is finite. Using the method of unpredictable paths, Levin and Peres (1998) sharpened this result and showed that the infinite cluster supports a nonzero flow f such that the q-energy $\sum_e |f(e)|^q$ is finite for all $q > d/(d-1)$. Thus the infinite cluster has the same "parabolic index" as the whole lattice. [See the last chapter of Soardi (1994) for the definition.]

Acknowledgments. We are grateful to Geoffrey Grimmett for a helpful remark on renormalization. We also thank Chris Hoffman, David Levin, Elhanan Mossel and the referee for their useful comments on the presentation.

REFERENCES

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

ATHREYA, K. and NEY, P. (1972). *Branching Processes*. Springer, New York.

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

DOYLE, P. G. and SNELL, E. J. (1984). *Random Walks and Electrical Networks*. Math. Assoc. Amer., Washington, D.C.

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

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

HÄGGSTRÖM, O. MOSSEL, E. (1998). Nearest-neighbor walks with low predictability profile and percolation in $2+\varepsilon$ dimensions. *Ann. Probab.* **26** 1212–1231.

HIEMER, P. (1998). Dynamical renormalisation in oriented percolation. Preprint.

HOFFMAN, C. (1998). Unpredictable nearest neighbor processes. Preprint.

KESTEN, H. and STIGUM, B. P. (1966). Additional limit theorems for indecomposable multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37** 1463–1481.

KESTEN, H. and ZHANG, Y. (1990). The probability of a large finite cluster in supercritical Bernoulli percolation. *Ann. Probab.* **18** 537–555.

LAWLER, G. (1991). *Intersections of Random Walks*. Birkhäuser, Boston.

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. Picardello and W. Woess eds.). To appear.

LIGGETT, T. M. SCHONMANN, R. H. and STACEY, A. M. (1996). Domination by product measures. *Ann. Probab.* **24** 1711–1726.

LYONS, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931–958.

LYONS, R. (1995). Random walks and the growth of groups. *C. R. Acad. Sci. Paris* **320** 1361–1366.

MOORE, T. and SNELL, J. L. (1979). A branching process showing a phase transition. *J. Appl. Probab.* **16** 252–260.

PEMANTLE, R. and PERES, Y. (1996). On which graphs are all random walks in random environments transient? In *Random Discrete Structures* (D. Aldous and R. Pemantle, eds.) 207–211. Springer, New York.

PISZTORA, A. (1996). Surface order large deviation for Ising, Potts and percolation models. *Probab. Theory Related Fields* **104** 427–466.
SOARDI, P. M. 1994. *Potential Theory on Infinite Networks*. Springer, Berlin.

ITAI BENJAMINI
MATHEMATICS DEPARTMENT
WEIZMANN INSTITUTE
REHOVOT 76100
ISRAEL

ROBIN PEMANTLE
DEPARTMENT OF MATHEMATICS
VAN VLECK HALL
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706
E-MAIL: pemantle@math.wisc.edu

YUVAL PERES
MATHEMATICS INSTITUTE
HEBREW UNIVERSITY
JERUSALEM
ISRAEL
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA