The Annals of Probability
1995, Vol. 23, No. 1, 105–140

CRITICAL RANDOM WALK IN RANDOM ENVIRONMENT ON TREES

BY ROBIN PEMANTLE[1] AND YUVAL PERES

University of Wisconsin-Madison and University of California, Berkeley

We study the behavior of random walk in random environment (RWRE) on trees in the critical case left open in previous work. Representing the random walk by an electrical network, we assume that the ratios of resistances of neighboring edges of a tree Γ are i.i.d. random variables whose logarithms have mean zero and finite variance. Then the resulting RWRE is transient if simple random walk on Γ is transient, but not vice versa. We obtain general transience criteria for such walks, which are sharp for symmetric trees of polynomial growth. In order to prove these criteria, we establish results on boundary crossing by tree-indexed random walks. These results rely on comparison inequalities for percolation processes on trees and on some new estimates of boundary crossing probabilities for ordinary mean-zero finite variance random walks in one dimension, which are of independent interest.

1. Introduction. Precise criteria are known for the transience of simple random walk on a tree (in this paper, a tree is an infinite, locally finite, rooted acyclic graph and has no leaves, i.e., no vertices of degree one). See for example Woess (1986), Lyons (1990) or Benjamini and Peres (1992). How is the type of the random walk affected if the transition probabilities are randomly perturbed? Qualitatively we can say that if this perturbation has no "backward push" (defined below), then the random walk tends to become more transient; the primary aim of this paper is to establish a quantitative version of this assertion.

Designate a vertex ρ of the tree as its root. For any vertex $\sigma \neq \rho$, denote by σ' the unique neighbor of σ closer to ρ (σ' is also called the parent of σ). An *environment* for random walk on a fixed tree, Γ, is a choice of transition probabilities $q(\sigma, \tau)$ on the vertices of Γ, with $q(\sigma, \tau) > 0$ if and only if σ and τ are neighbors. When these transition probabilities are taken as random variables, the resulting mixture of Markov chains is called *random walk in random environment* (RWRE). Following Lyons and Pemantle (1992), we study random environments under the homogeneity condition

$$\text{(1.1)} \qquad \text{the variables } X(\sigma) = \log\left(\frac{q(\sigma', \sigma)}{q(\sigma', \sigma'')}\right) \text{ are i.i.d. for } |\sigma| \geq 2,$$

where $|\sigma|$ denotes the distance from σ to ρ. Let X denote a random variable with this common distribution.

Received September 1992; revised June 1994.
[1]Research supported in part by an NSF postdoctoral fellowship.
AMS 1991 *subject classifications*. Primary 60J15; secondary 60G60, 60G70, 60E07.
Key words and phrases. Tree, random walk, random environment, random electrical network, tree-indexed process, percolation, boundary crossing, capacity, Hausdorff dimension.

There are several motivations for studying RWRE under the condition (1.1):

1. For nearest-neighbour RWRE on the integers, the assumption (1.1) is equivalent to assuming that the transition probabilities themselves are i.i.d. The first result on RWRE was obtained by Solomon (1975), who showed that when $X(\sigma)$ have mean zero and finite variance, then RWRE on the integer line is recurrent, while it is transient if $E(X) > 0$. Thus in determining the type of the RWRE, X plays the primary role. The integer line is the simplest infinite tree, and Theorem 2.1 below determines almost exactly the class of trees for which the same criterion applies. An assumption that random variables analogous to $X(\sigma)$ are stationary (and in particular, identically distributed) is also crucial in the work of Durrett (1986), which extended the RWRE results of Sinai (1982) to the multidimensional integer lattice.
2. In terms of the associated resistor network, (1.1) means that the ratios of the resistances of adjacent edges in Γ are i.i.d.; such networks are useful for determining the Hausdorff measures of certain random fractals—see Falconer (1987) and Lyons (1990). The logarithms of the resistances in such a network form a tree-indexed random walk. (A precise definition of such walks is given below.) This structure appears in a variety of settings: as a generalization of branching random walk [Joffe and Moncayo (1973)]; in a model for "random distribution functions" [Dubins and Freedman (1967)]; in the analysis of game trees [Nau (1983)]; in studies of random polymers [Derrida and Spohn (1988)] and in first-passage percolation [Lyons and Pemantle (1992); Benjamini and Peres (1994b)].
3. The tools developed to analyse RWRE satisfying the assumption (1.1) are also useful when that assumption is relaxed, for example, to allow for some dependence between vertices which are "siblings". In Section 7 we describe an application to certain *reinforced random walks* which may be reduced to a RWRE.

The main result of Lyons and Pemantle (1992) is that RWRE on Γ is a.s. transient if the *Hausdorff dimension*, denoted $\dim(\Gamma)$, is strictly greater than the *backward push*

$$\beta(X) \stackrel{\text{def}}{=} -\log \min_{0\le\lambda\le1} \mathbf{E}\, e^{\lambda X} \tag{1.2}$$

and a.s. recurrent if $\dim(\Gamma) < \beta(X)$. [The definition of $\dim(\Gamma)$ will be given in Section 2; the quantity $e^{\dim(\Gamma)}$ is called the branching number of Γ in papers of R. Lyons; the backward push β is zero whenever X has mean zero.]

While subsuming previous results in Lyons (1990) and Pemantle (1988), these criteria leave some interesting cases unresolved. For instance if $\mathbf{E}\, X = 0$ (the random environment is "fair"), then one easily sees that $\beta(X) = 0$, so the above criteria yield transience of the RWRE only when Γ has positive Hausdorff dimension, which in particular implies exponential growth. In fact, much smaller trees suffice for transience of the RWRE in this case, at least if X has a finite second moment (Theorem 2.1 below). In particular, this RWRE is a.s. transient whenever simple random walk on the same tree is transient.

To illustrate the difference between old criteria such as exponential growth and the criteria set forth in this paper, we limit the discussion for the rest of the introduction to *spherically symmetric trees*, that is, trees determined by a growth function $f: \mathbf{Z}^+ \to \mathbf{Z}^+$ for which every vertex at distance n from the root has degree $1 + f(n)$. Note, however, that much of the interest in these results stems from their applicability to nonsymmetric trees; criteria for general trees involve the notion of capacity and are deferred to the next section.

Assume that Γ is spherically symmetric and that the variables $X(\sigma)$ in (1.1) have mean zero and finite variance [the assumption of finite variance is plausible since the $X(\sigma)$ are logs of ratios, so the ratios themselves may still have large tails]. Our first result, Theorem 2.1, is that the RWRE is almost surely transient if

$$\sum_n n^{-1/2} |\Gamma_n|^{-1} < \infty \tag{1.3}$$

and this condition is also necessary, provided that the regularity condition

$$\sum_n n^{-3/2} \log |\Gamma_n| < \infty \tag{1.4}$$

holds, where $|\Gamma_n|$ is the cardinality of the nth level of Γ. We conjecture that condition (1.3) is necessary as well as sufficient for transience of RWRE. To see why this is a natural conjecture, and to point out that the randomness makes the walk more transient, compare this to the following known result: simple random walk on a spherically symmetric tree is transient if and only if $\sum |\Gamma_n|^{-1} < \infty$.

The key to proving the above result is an analysis of *tree-indexed random walks*. These are random fields $\{S(\sigma): \sigma \in \Gamma\}$ defined from a collection of i.i.d. real random variables $\{X(\sigma): \sigma \in \Gamma, \sigma \neq \rho\}$ by letting $S(\sigma)$ be the sum of $X(\tau)$ over vertices τ on the path from the root to σ. Note that when Γ is a single infinite ray (identified with the positive integers), then this is just an ordinary random walk; when Γ is the family tree of a Galton–Watson branching process, this is a branching random walk. Random walks indexed by general trees first appeared in Joffe and Moncayo (1973). The motivating question for the study of tree-indexed random walks [cf. Benjamini and Peres (1994a, b)] is this: when is Γ large enough so that for a Γ-indexed random walk, the values of $S(\sigma)$ along at least one ray of Γ exhibit a prescribed behavior atypical for an ordinary random walk?

In this paper we prove several results in this direction, one of which we now describe, and apply them to RWRE on trees. In the special case where the variables $X(\sigma)$ take only the values ± 1 with equal probability, Benjamini and Peres (1994b) obtained conditions for the existence of a ray in Γ along which the partial sums $S(\sigma)$ tend to infinity. In particular, for spherically symmetric trees, (1.3) suffices for the existence of such a ray while the condition

$$\liminf_{n \to \infty} n^{-1/2} |\Gamma_n| > 0 \tag{1.5}$$

is necessary. In Theorem 2.2 below, this result is extended to variables $X(\sigma)$ with zero mean and finite variance, and also sharpened. For spherically symmetric trees, we show that (1.3) is necessary and sufficient for the existence of a ray along which $S(\sigma) \to \infty$.

As is well known, transience of a reversible Markov chain is equivalent to finite resistance of the associated resistor network, where the transition probabilities from any vertex are proportional to the conductances (reciprocal resistances); see for example Doyle and Snell (1984). For an environment satisfying (1.1), the conductance attached to the edge between σ' and σ is $e^{S(\sigma)}$, where $\{S(\sigma)\}$ is the Γ-indexed random walk with increments $\{X(\tau): \tau \neq \rho\}$. Since finite resistance is a tail event, transience of the environment satisfies a zero–one law. In particular, the network will have finite resistance whenever a ray exists along which $S(\sigma) \to \infty$ sufficiently fast so that $e^{-S(\sigma)}$ is summable. In this way Theorem 2.2 yields Theorem 2.1.

For completeness, we state here a result from Pemantle (1993) about the case where the i.i.d. random variables $\{X(\sigma)\}$ have negative mean and the backward push $\beta(X)$ is positive. For a spherically symmetric tree Γ, the result of Lyons and Pemantle (1992) yields recurrence of the RWRE if

$$\liminf_{n\to\infty} e^{-n\beta} |\Gamma_n| = 0$$

and transience if

$$|\Gamma_n| \geq Ce^{n(\beta+\varepsilon)}$$

for some $C, \varepsilon > 0$ and all n. Here, analyzing the critical case is more difficult, but assuming a regularity condition on the random environment it can be shown that the boundary between transience and recurrence occurs when

$$|\Gamma_n| \approx e^{\beta n + cn^{1/3}}.$$

Here, unlike in the mean zero case, randomness makes the RWRE more recurrent, since the known necessary and sufficient condition for transience of RWRE when $X(\sigma) = -\beta$ a.s. is that

$$\sum e^{n\beta} |\Gamma_n|^{-1} < \infty.$$

The rest of the paper is organized as follows. Precise statements of our main results are collected in the next section. Some estimates for ordinary, mean-zero, finite variance random walks that will be needed in the sequel are collected in Section 3. Some of these, along the lines of Woodroofe (1976), may be of independent interest. In particular, we determine the rate of growth of a mean-zero, finite variance random walk conditioned to remain positive; this sharpens considerably a result of Ritter (1981). Section 4 explains the second moment method for trees, developed by R. Lyons. Here we have some new results comparing dependent and independent percolation and comparing spherically symmetric trees to nonsymmetric trees of the same size (Theorem 4.3). After these preliminaries, tree-indexed random walks are discussed in Section 5, along with an example in which the increments are symmetric stable random variables. The example shows that the *sustainable speed* of a Γ-indexed random walk with a given increment distribution may not be determined by the dimension of Γ. Also, this answers a question of R. Lyons (personal communication) by providing an RWRE satisfying (1.1) which is transient on a tree of polynomial growth, even though $\mathbf{E}\, X(\sigma) < 0$. The RWRE with no backward push is discussed in Section 6 and an application to reinforced random walk is described in Section 7.

2. Statements of results. We begin with some definitions. Recall that all our trees are infinite, locally finite, rooted at some vertex ρ and have no leaves. We use the notation $\sigma \in \Gamma$ to mean that σ is a vertex of Γ.

1. An infinite path from the root of a tree Γ is called a *ray* of Γ. We refer to the collection of all rays as the *boundary*, $\partial\Gamma$, of Γ.
2. If a vertex τ of Γ is on the path connecting the root, ρ, to a vertex σ, then we write $\tau \le \sigma$. For any two vertices σ and τ, let $\sigma \wedge \tau$ denote their greatest lower bound, that is, the vertex where the paths from ρ to σ and τ diverge. Similarly, the vertex at which two rays ξ and η diverge is denoted $\xi \wedge \eta$.
3. A set of vertices Π of Γ which intersects every ray of Γ is called a *cutset.*
4. Let $\phi\colon \mathbf{Z}^+ \to \mathbf{R}$ be a decreasing positive function with $\phi(n) \to 0$ as $n \to \infty$. The Hausdorff measure of Γ in gauge ϕ is

$$\liminf_{\Pi} \sum_{\sigma\in\Pi} \phi(|\sigma|),$$

   where the lim inf is taken over Π such that the distance from ρ to the nearest vertex in Π goes to infinity. The supremum over α for which Γ has positive Hausdorff measure in gauge $\phi(n) = e^{-n\alpha}$ is called the *Hausdorff dimension* of Γ. Strictly speaking, this is the Hausdorff dimension of the boundary of Γ in the metric $d(\xi, \eta) = e^{-|\eta\wedge\xi|}$. For spherically symmetric trees, this is just the lim inf exponential growth rate; for general trees it may be smaller.
5. Hausdorff measure may be defined for Borel subsets $A \subseteq \partial\Gamma$ by only requiring the cutsets Π to intersect all rays in A. Say that Γ has σ-finite Hausdorff measure in gauge ϕ if $\partial\Gamma$ is the union of countably many subsets with finite Hausdorff measure in gauge ϕ.
6. Say that Γ has *positive capacity* in gauge ϕ if there is a probability measure μ on $\partial\Gamma$ for which the *energy*

$$I_\phi(\mu) = \int_{\partial\Gamma}\int_{\partial\Gamma} \phi(|\xi \wedge \eta|)^{-1}\, d\mu(\xi)\, d\mu(\eta)$$

   is finite. The infimum over probability measures μ of this energy is denoted by $1/\operatorname{Cap}_\phi(\Gamma)$.

An important fact about capacity and Hausdorff measure, proved by Frostman in 1935, is that σ-finite Hausdorff measure in gauge ϕ implies zero capacity in gauge ϕ; the converse just barely fails [cf. Carleson (1967), Theorem 4.1]. This gap is either the motivation or the bane of much of the present work, since many of our criteria would be necessary and sufficient if zero capacity were identical to σ-finite Hausdorff measure.

THEOREM 2.1 (Proved in Section 6). *Suppose that i.i.d. random variables* $\{X(\sigma): \rho \neq \sigma \in \Gamma\}$ *are used to define an environment on a tree* Γ *via* (1.1), *that is, the edge from* σ' *to* σ *is assigned the conductance*

$$\prod_{\rho<\tau\leq\sigma} e^{X(\tau)}.$$

Assume that $X(\sigma)$ *have zero mean and finite variance.*

(i) *If* Γ *has positive capacity in gauge* $\phi(n) = n^{-1/2}$, *then the resulting RWRE is transient.*

(ii) *If* Γ *has zero Hausdorff measure in the same gauge, then the RWRE is recurrent.*

(iii) *If* Γ *satisfies the regularity condition*

$$\sum_{n=1}^{\infty} n^{-3/2} \log|\Gamma_n| < \infty, \tag{2.1}$$

then $\sum_{n=1}^{\infty} n^{-1/2}|\Gamma_n|^{-1} = \infty$ *implies recurrence of the RWRE. In particular, if* Γ *is spherically symmetric and satisfies the regularity condition, then positive capacity in gauge* $n^{-1/2}$ *is necessary and sufficient for transience.*

REMARKS.

1. Lyons (1990) shows that simple random walk [$X(\sigma) = 0$ with probability one] is transient if and only if Γ has positive capacity in gauge $\phi(n) = n^{-1}$. Thus part (i) of the theorem justifies the assertion in the Introduction that a fair random environment makes the random walk more transient. For spherically symmetric trees the definitions of Hausdorff measure and capacity are simpler and the theorem reduces to the conditions in (1.3)–(1.5).
2. Any spherically symmetric tree to which this theorem does not apply must have zero capacity in gauge $n^{-1/2}$, but fail the regularity condition; this implies it grows in vigorous bursts, satisfying $|\Gamma_n| < n^{1/2+\varepsilon}$ infinitely often, and $|\Gamma_n| > \exp(n^{1/2-\varepsilon})$ infinitely often as well. One is not likely to encounter such a tree in the course of daily affairs.
3. If the variables $X(\sigma)$ in (1.1) have positive expectation, then (trivially) for any tree Γ the RWRE is transient, since the sum of the resistances along any fixed ray is almost surely finite.

Part (i) of the theorem is proved by showing that in the mean-zero case there exists a random ray with the same property. This in turn is deduced from the next theorem concerning tree-indexed random walks.

THEOREM 2.2 (Proved in Section 5). *Let $\{X(\sigma)\}$ be i.i.d. random variables indexed by the vertices of Γ and let $S(\sigma) = \sum_{\rho<\tau\leq\sigma} X(\tau)$. Suppose that $X(\sigma)$ have zero mean and finite variance. Then*

(i) *If Γ has positive capacity in gauge $\phi(n) = n^{-1/2}$, then*

$$\mathbf{P}(\exists \xi \in \partial\Gamma : \forall \sigma \in \xi \;\; S(\sigma) \geq 0) > 0.$$

Furthermore, under the same capacity condition, for every increasing positive function f satisfying

$$\sum_{n=1}^{\infty} n^{-3/2} f(n) < \infty, \tag{2.2}$$

there exists with probability one a ray ξ of Γ such that $S(\sigma) \geq f(|\sigma|)$ for all but finitely many $\sigma \in \xi$.

(ii) *If Γ has σ-finite Hausdorff measure in gauge $\phi(n) = n^{-1/2}$, then*

$$\mathbf{P}(\exists \xi \in \partial\Gamma : \forall \sigma \in \xi \;\; S(\sigma) \geq 0) = 0.$$

Furthermore, for any increasing f satisfying (2.2), *there is with probability one NO ray ξ such that $S(\sigma) \geq -f(|\sigma|)$ for all but finitely many $\sigma \in \xi$.*

(iii) *The conclusions of part* (ii) *also hold if we assume, instead of the Hausdorff measure assumption, that $\sum_n n^{-1/2}|\Gamma_n|^{-1} = \infty$.*

REMARK. For spherically symmetric trees, parts (i) and (iii) cover all cases, thus proving a sharp dichotomy for tree-indexed random walks: either there exist rays with $S(\sigma)$ tending to infinity faster than $n^{1/2-\varepsilon}$ for all $\varepsilon > 0$ or else along every ray $S(\sigma)$ must dip below $-n^{1/2-\varepsilon}$ infinitely often. We believe this dichotomy holds for all trees, but the proof eludes us. In general, the condition in part (iii) is not comparable to the condition in part (ii).

3. Estimates for mean-zero, finite variance random walk. Here we collect estimates for ordinary, one-dimensional, mean-zero, finite variance random walks which are needed in the sequel. Begin with a classical estimate whose proof may be found in Feller (1966), Section XII.8.

PROPOSITION 3.1. *Let $X_1, X_2, \ldots$ be i.i.d., nondegenerate, mean-zero, random variables with finite variance and let $S_n = \sum_{k=1}^{n} X_k$. Let T_0 denote the hitting time on the negative half-line: $T_0 = \min\{n \geq 1 : S_n < 0\}$. Then*

$$\lim_{n\to\infty} \sqrt{n}\, \mathbf{P}(T_0 > n) = c_1 > 0 \tag{3.1}$$

and, in particular, $c_1' \leq \sqrt{n}\, \mathbf{P}(T_0 > n) \leq c_1''$ for all n.

We now determine which boundaries $f(n)$ behave like the horizontal boundary $f(n) \equiv 0$ in that $\mathbf{P}(S_k > f(k),\ k = 1, \dots, n)$ is still asymptotically $cn^{-1/2}$.

THEOREM 3.2. *With X_n and S_n as in the previous proposition, let $f(n)$ be any increasing positive sequence. Then:*

(i) *The condition*

$$\sum_{n=1}^{\infty} n^{-3/2} f(n) < \infty \tag{3.2}$$

is necessary and sufficient for the existence of an integer n_f such that

$$\inf_{n \geq n_f} \sqrt{n}\, \mathbf{P}(S_k \geq f(k) \text{ for } n_f \leq k \leq n) > 0.$$

(ii) *The same condition* (3.2) *is necessary and sufficient for*

$$\sup_{n \geq 1} \sqrt{n}\, \mathbf{P}(S_k \geq -f(k) \text{ for } 1 \leq k \leq n) < \infty.$$

REMARKS.

1. Part (i) may be restated as asserting that if f satisfies (3.2), then random walk conditioned to stay positive to time n stays above f between times n_f and n with probability bounded away from zero as $n \to \infty$. This condition arises in the classical Dvoretzky–Erdös test for random walk in three-space to eventually avoid a sequence of concentric balls of radii $f(n)$. Passing to the continuous limit, the absolute value of random walk in three-space becomes a Bessel (3) process, which is just a (one-dimensional) Brownian motion conditioned to stay positive. Proving Theorem 3.2 via this connection (using Skorohod representation, say) seems more troublesome than the direct proof.
2. In the case where the positive part of the summands X_i is bounded and the negative part has a moment generating function, Theorem 3.2 (with further asymptotics) was proved by Novikov (1983). He conjectured that these conditions could be weakened. For the Brownian case, see Millar (1976). Other estimates of this type are given by Woodroofe (1976) and Roberts (1991). For their statistical ramifications, see Siegmund (1986) and the references therein. Our estimate can be used to calculate the rate of escape of a random walk conditioned to stay positive *forever*; this process has been studied by several authors; see Keener (1992) and the references therein.

The proof uses the following three-part lemma. Let

$$T_h = \min\{n \geq 1 : S_n < -h\}$$

denote the hitting time on $(-\infty, -h)$.

LEMMA 3.3.

(i) $\mathbf{P}(T_h > n) \le c_2 h n^{-1/2}$ *for all integers* $n \ge 1$ *and real* $h \ge 1$.
(ii) $\mathbf{E}(S_n^2 \mid T_0 > n) \le c_3 n$ *for* $n \ge 1$.
(iii) $\mathbf{P}(T_h > n) \ge c_4 h n^{-1/2}$ *for all integers* $n \ge 1$ *and real* $h \le \sqrt{n}$.

REMARKS (Corresponding to the assertions in the lemma). (i) This estimate is from Kozlov (1976); as we shall see, it follows immediately from Proposition 3.1.

(ii) In fact we shall verify that

$$\mathbf{E}(S_n^2 \mid T_0 > n) \le 2n\,\mathrm{Var}(X_1) + o(n),$$

where the constant 2 cannot be reduced in general.

(iii) Under the additional assumption that $\mathbf{E}\,|X_1|^3 < \infty$, this estimate is proved in Lemma 1 of Zhang (1991). Assuming only finite variance, as we do, the estimate was known to Kesten (personal communication) and is implicit in Lawler and Polaski (1992).

PROOF OF LEMMA 3.3. (i) We may assume that $h \le \sqrt{n}$ and that h is an integer. By the central limit theorem there exists an $r \ge 1$, depending only on the common distribution of the X_k, such that

$$\mathbf{P}(S_{r^2h^2} > h) > 1/3.$$

From Proposition 3.1 and the FKG inequality [or the Harris inequality; see Grimmett (1989), Section 2.2]:

$$\mathbf{P}(S_{r^2h^2} > h \text{ and } T_0 > r^2h^2) > \frac{c_1'(rh)^{-1}}{3} = ch^{-1}.$$

Consequently,

$$\begin{aligned} &ch^{-1}\mathbf{P}(T_h > n)\\ &\quad \le \mathbf{P}\left[S_{r^2h^2} > h \text{ and } T_0 > r^2h^2 \text{ and } \sum_{r^2h^2+1}^{r^2h^2+k} X_j \ge -h \text{ for } 1 \le k \le n\right]\\ &\quad \le \mathbf{P}(T_0 > r^2h^2 + n)\\ &\quad \le c_1'' n^{-1/2}, \end{aligned}$$

which yields the required estimate.

(ii) Consider the minimum of T_0 and n:

$$\mathbf{E}(T_0 \wedge n) = \sum_{k=1}^{n} \mathbf{P}(T_0 \ge k) = \sum_{k=1}^{n} (c_1 + o(1))k^{-1/2} = 2(c_1 + o(1))n^{1/2}.$$

Therefore, using Wald's identity,

$$\mathbf{E}(S_n^2 \mathbf{1}_{T_0>n}) \le \mathbf{E}\,S_{T_0\wedge n}^2 = \mathbf{E}\,X_1^2\,\mathbf{E}(T_0 \wedge n) = 2\,\mathbf{E}\,X_1^2 (c_1 + o(1))n^{1/2}.$$

Dividing both sides by $\mathbf{P}(T_0 > n)$ and using Proposition 3.1 gives

$$\mathbf{E}(S_n^2 \mid T_0 > n) \le (2 + o(1))n\,\mathbf{E}\,X_1^2.$$

Analyzing the proof, it is easy to see that this is sharp at least when the increments X_j are bounded.

(iii) First, we use (ii) to derive the estimate

$$\mathbf{E}(S_n^2 \mid T_h > n) \le cn \tag{3.3}$$

with c independent of h and n. By the invariance principle,

$$\inf\{\mathbf{P}(T_h > n): n \ge 1,\ h \ge \sqrt{n}\} > 0,$$

so (3.3) is immediate for $h \ge \sqrt{n}$. Assume then that $h < \sqrt{n}$. Let A_i denote the event that $T_h > n$ and S_i is the last minimal element among $0 = S_0, S_1, \dots, S_n$. Then

$$\mathbf{E}(S_n^2 \mid T_h > n) = \sum_{i=1}^{n} \mathbf{P}(A_i)\,\mathbf{E}(S_n^2 \mid A_i).$$

Conditioning further on S_i we see that this is at most

$$\sup\{\mathbf{E}(S_n^2 \mid A_i, S_i = y): 1 \le i \le n,\ -h \le y \le 0\},$$

but by the Markov property,

$$\mathbf{E}(S_n^2 \mid A_i, S_i = y) = \mathbf{E}((y + S_{n-i})^2 \mid S_k > 0 \text{ for } 1 \le k \le n - i).$$

Since $y < 0$ and $y^2 < n$, this gives

$$\mathbf{E}(S_n^2 \mid T_h > n) \le \sup_{0 \le i \le n} \{n + \mathbf{E}(S_{n-i}^2 \mid S_k > 0 \text{ for } 1 \le k \le n - i)\} \le (1 + c_3)n$$

with c_3 as in (ii), proving (3.3).

Now letting A denote the event $\{T_h > n\}$, we have

$$0 = \mathbf{E}\,S_{T_h \wedge n} = \mathbf{P}(A)\,\mathbf{E}(S_n \mid A) + \mathbf{P}(A^c)\,\mathbf{E}(S_{T_h} \mid A^c)$$

and therefore [using (3.3) in the final step],

$$\frac{\mathbf{P}(A)}{\mathbf{P}(A^c)} = \frac{\mathbf{E}(-S_{T_h} \mid A^c)}{\mathbf{E}(S_n \mid A)} \ge \frac{h}{(\mathbf{E}(S_n^2 \mid A))^{1/2}} \ge \frac{h}{(cn)^{1/2}}.$$

This establishes (iii). □

Proof of Theorem 3.2(i). First assume that $\sum n^{-3/2} f(n) < \infty$. Consider the events

$$V_m = \{S_k > f(2^m) \text{ for } 2^{m-1} < k \le 2^m\}.$$

We claim that for any $N \geq 2^{m-1}$,

$$\mathbf{P}(V_m^c \mid T_0 > 4N) \leq \tilde{c} f(2^m) 2^{-m/2} \tag{3.4}$$

with some constant $\tilde{c} > 0$. Indeed by conditioning on the first $k \in [2^{m-1}+1, 2^m]$ for which $S_k \leq f(2^m)$, one sees that

$$\begin{aligned}
&\mathbf{P}(V_m^c \cap \{T_0 > 4N\}) \\
&\quad \leq \mathbf{P}(T_0 > 2^{m-1}) \max_{k \leq 2^m} \mathbf{P}(S_j - S_k \geq -f(2^m) \text{ for all } j \in [k+1, 4N]) \\
&\quad \leq \mathbf{P}(T_0 > 2^{m-1}) \mathbf{P}(T_{f(2^m)} \geq N) \\
&\quad \leq c_1'' c_2 f(2^m)(N 2^{m-1})^{-1/2}.
\end{aligned}$$

Using Proposition 3.1, this establishes (3.4). Since f is nondecreasing, the hypothesis

$$\sum_n n^{-3/2} f(n) < \infty$$

is equivalent to $\sum_{m=1}^{\infty} 2^{-m/2} f(2^m) < \infty$. Choose an integer m_f such that

$$\tilde{c} \sum_{m=m_f}^{\infty} 2^{-m/2} f(2^m) < \tfrac{1}{2}.$$

By (3.4), for every $M > m_f$ we have

$$\sum_{m=m_f}^{M} \mathbf{P}(V_m^c \mid T_0 > 2^{M+1}) \leq \tfrac{1}{2}$$

and hence

$$\mathbf{P}\left(\bigcap_{m=m_f}^{M} V_m \mid T_0 > 2^{M+1} \right) \geq \tfrac{1}{2}.$$

Taking $n_f = 2^{m_f}$ and recalling the definition of V_m concludes the proof.

For the converse, first recall a known fact about mean-zero, finite variance, random walks, namely, that

$$\lim_{R \to \infty} \sup_{h \geq 0} \mathbf{P}(h + S_{T_h} \leq -R) = 0. \tag{3.5}$$

In other words, the amounts by which $\{S_n\}$ *overshoots* the boundary $-h$ are tight as h varies over $(0, \infty)$. Indeed the overshoots for the random walk $\{S_n\}$ are the same as for the associated renewal process $\{L_n\}$ of descending ladder random variables, where L_1 is the first negative value among $S_1, S_2, \ldots$, and in general, L_{n+1} is the first among $\{S_k\}$ which is less than L_n. The differences $L_{n+1} - L_n$ are i.i.d.; they have finite first moment if and only if $\mathbf{E}\, X_1^2 < \infty$ [see Feller (1966), Section XVIII.5]. In this case, the overshoots are tight since by the renewal theorem, they converge in distribution [see Feller (1966), Section XI.3]. This proves (3.5).

Some new notation will be useful:

DEFINITION 1. *For any function* $f(n)$ *and any random walk* $\{S_n\}$*, let* $A(f;a,b)$ *denote the event that* $S_n \geq f(n)$ *for all* $n \in [a,b]$*. Let* $A(f;b)$ *denote* $A(f;1,b)$.

Proceeding now to the proof itself, it is required to prove (3.2) from the assumption that for some $n_f \geq 1$,

$$\inf_{n \geq n_f} \mathbf{P}(A(f;n_f,n) \mid T_0 > n) > 0. \tag{3.6}$$

It may be assumed without loss of generality that $f(n) \to \infty$, since otherwise there is nothing to prove; also, by changing f at finitely many integers, it may be assumed, without affecting the condition (3.2) we are trying to prove, that (3.6) holds with $n_f = 1$.

Impose the restriction $f(n) \leq \sqrt{n}$; this restriction will be removed at the end of the proof. Let c_6 be the infimum of probabilities $\mathbf{P}(A(f;n) \mid T_0 > n)$, which is positive by (3.6). The key estimate to proving (3.2) is

$$\mathbf{P}(A(f;2n)^c \mid A(f;n) \text{ and } T_0 > N) \geq c_7 f(n) n^{-1/2} \tag{3.7}$$

for some $c_7 > 0$, all sufficiently large n and all $N \geq 2n$. Verifying this estimate involves several steps.

Step 1: *Controlling* S_n*, given* $A(f;n)$. From Lemma 3.3(ii),

$$\mathbf{E}(S_n^2 \mid A(f;n)) \leq c_6^{-1} \mathbf{E}(S_n^2 \mid T_0 > n) \leq (c_3/c_6)n.$$

Therefore,

$$\mathbf{P}(S_n \leq c_8\sqrt{n} \mid A(f;n)) \geq \tfrac{1}{2}, \tag{3.8}$$

where $c_8 = 2c_3/c_6 > 0$.

Step 2: *Securing a dip below the boundary.* By the central limit theorem there exists an integer n^* and a constant $c_9 > 0$ such that

$$\mathbf{P}(S_{2n} - S_n < -c_8\sqrt{n}) \geq 4c_9 \quad \text{for } n \geq n^*.$$

Let $t(n) = \min\{k > n : S_k < f(n)\}$. Then nonnegativity of f and the Markov property imply

$$\mathbf{P}(t(n) \leq 2n \mid A(f;n), S_n) \geq 4c_9 \mathbf{1}_{S_n \leq c_8\sqrt{n}},$$

and hence, by (3.8),

$$\mathbf{P}(t(n) \leq 2n \mid A(f;n)) \geq 2c_9 \tag{3.9}$$

whenever $n \geq n^*$.

Step 3: *Controlling the overshoot.* Use tightness of the overshoots to pick an $R > 0$ such that $\mathbf{P}(S_{t(n)} \geq f(n) - R \mid S_n = y) \geq 1 - c_9$ for any $y \geq f(n)$. Increase n^* if necessary to ensure that $f(n^*) > 2R$ and hence for all $n \geq n^*$, $\mathbf{P}(S_{t(n)} \geq f(n)/2 \mid A(f;n)) \geq 1 - c_9$. Combining this with (3.9) yields

$$\mathbf{P}\Big(t(n) \leq 2n \text{ and } S_{t(n)} \geq \frac{f(n)}{2} \,\Big|\, A(f;n)\Big) \geq c_9;$$

thus by Proposition 3.1 and the definition of c_6 there is some $c_{10} > 0$ such that for all n,

$$\mathbf{P}\Big(A(f;n) \cap \{t(n) \leq 2n\} \cap \Big\{S_{t(n)} \geq \frac{f(n)}{2}\Big\}\Big) \geq c_{10} n^{-1/2}. \tag{3.10}$$

Step 4: *Maintaining positivity.* From the strong Markov property and Lemma 3.3(iii), the event $\{S_k - S_{t(n)} \geq -f(n)/2 \text{ for } k \in [t(n)+1, t(n)+N]\}$ is independent of the random walk up to time $t(n)$ and has probability at least $(c_4/2)f(n)N^{-1/2}$. Multiplying by the inequality (3.10) proves that

$$\mathbf{P}(A(f;n) \cap A(f;2n)^c \cap \{T_0 > N\}) \geq c_{11} f(n)(nN)^{-1/2},$$

where $c_{11} = c_{10} \cdot c_4/2$. Now the key estimate (3.7) follows from Proposition 3.1.

From here the rest is easy sailing. If $n > n^*$ and $2^M \geq 2n$, then

$$\mathbf{P}(A(f;2n) \mid T_0 > 2^M) \leq (1 - c_7 f(n) n^{-1/2})\, \mathbf{P}(A(f;n) \mid T_0 > 2^M).$$

Therefore,

$$\mathbf{P}(A(f;2^M) \mid T_0 > 2^M) \leq \prod_{\log_2 n^* < m \leq M} 1 - c_7 f(2^m) 2^{-m/2}.$$

Recalling that the LHS is bounded away from zero, we infer that necessarily

$$\sum_m 2^{-m/2} f(2^m) < \infty.$$

Having proved summability from (3.6) when $f(n) \leq \sqrt{n}$, we now remove the restriction. If the restriction is violated finitely often, this is easily corrected where it is used in Step 2 by choosing n^* sufficiently large. If the restriction is violated infinitely often, then the above proof works for $g(n) = f(n) \wedge \sqrt{n}$ to show that

$$\sum_{m=1}^{\infty} 2^{-m/2} \min(f(2^m), 2^{m/2}) < \infty,$$

which contradicts infinitely many violations. Thus $2^{-m/2} f(2^m)$ is summable in any event, which is equivalent to (3.2). □

PROOF OF THEOREM 3.2(ii). We may assume that $f(n) \uparrow \infty$ since Lemma 3.3(i) covers bounded f. Retain the notation $A(f;a,b)$ from the previous proof. One of the two halves of the equivalence, $\sup_{n \geq 1} \sqrt{n}\, \mathbf{P}(S_k \geq -f(k) \text{ for } 1 \leq k \leq n) < \infty$ implies summability of (3.2), is easy. Assume that

$\mathbf{P}(A(-f;n)) \le Cn^{-1/2}$ for all $n \ge 1$. Under this assumption, we may repeat the proof of (3.7) substituting 0 for the upper boundary, f, and substituting $-f$ for the lower boundary, 0, to yield

$$\mathbf{P}(A(0;2n)^c \mid A(0;n) \cap A(-f;N)) \ge c_7 f(n) n^{-1/2} \tag{3.11}$$

for all large n and $N \ge 4n$. Our assumption implies that the products

$$\prod_{m=1}^{M} \mathbf{P}(A(0;2^{m+1}) \mid A(0;2^m) \cap A(-f;2^{M+2})),$$

being greater than $\mathbf{P}(A(0;2^{M+1}) \mid A(-f;2^{M+2}))$, must be bounded below by a positive constant. From (3.11) it follows that the product of $1 - c_7 n^{-1/2} f(n)$ is nonzero as n ranges over powers of two, which implies $\sum 2^{-m/2} f(2^m) < \infty$, completing the proof for this half of the equivalence.

The other direction is a consequence of an inequality which will require some work to prove: there is a constant $c_{12} \ge 1$ for which

$$\mathbf{P}(T_0 < 2n \mid T_0 \ge n \text{ and } A(-f;N)) \le c_{12} f(3n) n^{-1/2} \tag{3.12}$$

provided that $N \ge 4n$ [observe that when $f(3n)^2 \ge n$, the inequality is trivial].

Assuming (3.12) for the moment,

$$\begin{aligned} &\mathbf{P}(T_0 \ge 2^{M+1} \mid A(-f;2^{M+2})) \\ &\quad \ge \mathbf{P}(T_0 \ge 2^{m_0} \mid A(-f;2^{M+2})) \cdot \prod_{m=m_0}^{M} (1 - c_{12} f(3 \cdot 2^m) 2^{-m/2}) \\ &\quad \ge c_{m_0} \prod_{m=m_0}^{M} (1 - c_{12} f(3 \cdot 2^m) 2^{-m/2}), \end{aligned} \tag{3.13}$$

where m_0 is large enough so that all the factors on the right are positive. Since the summability of (3.2) is equivalent to

$$\sum f(3 \cdot 2^m) 2^{-m/2} < \infty,$$

the RHS of (3.13) is bounded from below by a positive constant c_{13} and, therefore,

$$\mathbf{P}(A(-f, 2^{M+2})) \le c^{-1} \mathbf{P}(T_0 \ge 2^{M+1}).$$

Proposition 3.1 then easily yields the inequality we are seeking (use the smallest M such that $2^M \ge n$).

Thus it suffices to establish (3.12). This is done by cutting and pasting portions of the random walk trajectory. To bound the LHS of (3.12) we will condition on the time $T_0 = k \in [n, 2n)$ of the first negative value of the random walk and on the overshoot $S_{T_0} = -y < 0$. This gives

$$\begin{aligned} &\mathbf{P}(n \le T_0 < 2n \text{ and } A(-f;N)) \\ &\quad \le \mathbf{P}(T_0 \ge n) \max_{\substack{n \le k < 2n \\ 0 < y < f(k)}} \mathbf{P}(A(-f;k,N) \mid S_k = -y). \end{aligned} \tag{3.14}$$

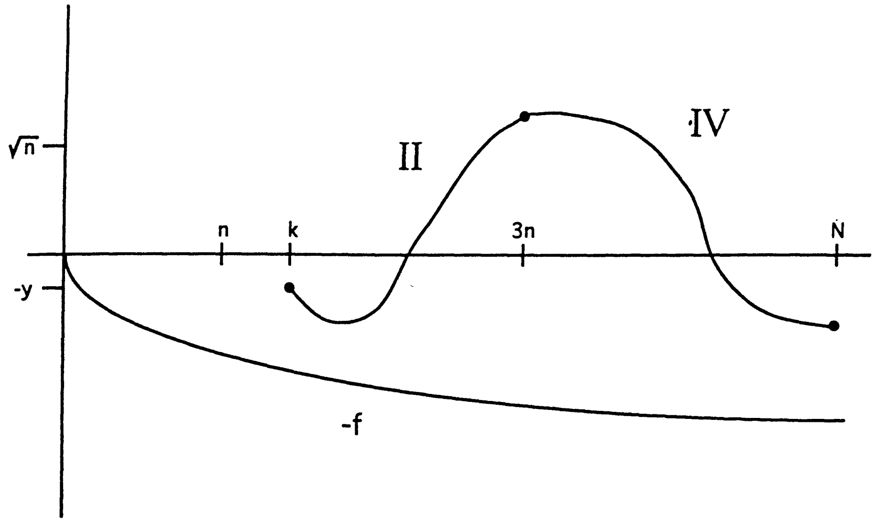


FIG. 1. *A trajectory in* $B_1 = A(-f; k, N) \cap \{S_{3n} > \sqrt{n}\}$, *given* $S_k = -y$.

Next observe that, given $S_k = -y$, the events $A(-f; k, N)$ and $\{S_{3n} > \sqrt{n}\}$ are both increasing events in the conditionally independent variables $X_{k+1}, \dots, X_N$. Applying the Harris inequality (or FKG) yields

$$\begin{aligned}
&\mathbf{P}(A(-f; k, N) \text{ and } S_{3n} > \sqrt{n} \mid S_k = -y) \\
(3.15) \qquad &\geq \mathbf{P}(A(-f; k, N) \mid S_k = -y) \cdot \mathbf{P}(S_{3n} > \sqrt{n} \mid S_k = -y) \\
&\geq c_{14} \mathbf{P}(A(-f; k, n) \mid S_k = -y),
\end{aligned}$$

where the last inequality uses the fact that $3n - k > n$, that $0 < y \leq f(3n) < \sqrt{n}$, and the central limit theorem.

Now begins the cutting and pasting. We shall combine a trajectory $X_{k+1}^{(1)}, X_{k+2}^{(1)}, \dots, X_N^{(1)}$ in the event on the LHS of (3.15) (called B_1 in Figure 1) with two independent random walk trajectories $\{X_j^{(2)}: j \geq 1\}$ and $\{X_j^{(3)}: j \geq 1\}$ depicted in Figures 2 and 3, respectively.

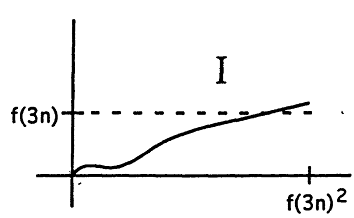


FIG. 2. *A trajectory in* $B_2 = A(0; f(3n)^2) \cap \{S_{f(3n)^2} > f(3n)\}$.

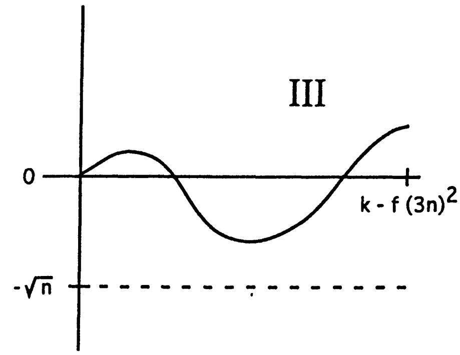


FIG. 3. *A trajectory in* $B_3 = A(-\sqrt{n}; k - f(3n)^2) \cap \{S_{k-f(3n)^2} > 0\}$.

Assume without loss of generality that $f(3n)^2$ is an integer. Define an event (i.e., a subset of sequence space) by

$$B(k, y) = A(y - f; 0, N - k) \cap \{S_{3n-k} > \sqrt{n} + y\}$$

and observe that $B_1 \cap \{S_k = -y\}$ may be written as the intersection of $\{S_k = -y\}$ with the event that the shifted sequence $X^{(1)}_{k+1}, X^{(1)}_{k+2}, \ldots$ is in $B(k, y)$. Define the mapping taking the three trajectories $\{X^{(i)}_j\}$, $i = 1, 2, 3$ into a trajectory $\{\tilde{X}_j : 1 \le j \le N\}$ as follows:

1. $\tilde{X}_j = X^{(2)}_j$ for $1 \le j \le f(3n)^2$,
2. $\tilde{X}_{f(3n)^2+j} = X^{(1)}_{k+j}$ for $1 \le j \le 3n - k$,
3. $\tilde{X}_{f(3n)^2+3n-k+j} = X^{(3)}_j$ for $1 \le j \le k - f(3n)^2$,
4. $\tilde{X}_j = X^{(1)}_j$ for $3n + 1 \le j \le N$.

We claim that the "pasted" trajectory $\{\tilde{X}_j : 1 \le j \le N\}$ lies in the event $B_4 \overset{\text{def}}{=} A(0; n) \cap A(-f; N)$ depicted in Figure 4 whenever the trajectories $\{X^{(1)}_j\}$, $\{X^{(2)}_j\}$ and $\{X^{(3)}_j\}$ lie in B_1, B_2 and B_3, respectively (see Figures 1–4 for the definitions of B_1, B_2 and B_3). Indeed, let $\tilde{S}_j = \tilde{X}_1 + \cdots + \tilde{X}_j$ and observe that for $1 \le j \le 3n - k$,

$$\sum_{i=k+1}^{k+j} X^{(1)}_i \ge -f(k + j) - (-y) \ge -f(3n).$$

Thus $\tilde{S}_{f(3n)^2+j} \ge \tilde{S}_{f(3n)^2} - f(3n) \ge 0$. This verifies that part II of the trajectory in Figure 4 satisfies the requirements to be in B_4; the other verifications are immediate.

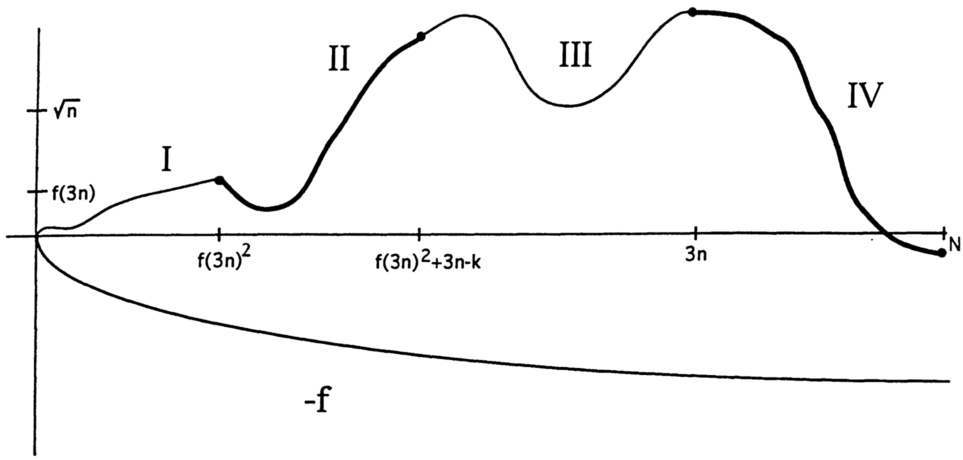


FIG. 4. *A trajectory in* $B_4 = A(0;n) \cap A(-f;N)$.

Since the sequence $\{\tilde{X}_j\}$ is a fixed permutation of the three sequences $\{X_j^{(1)}\}$, $\{X_j^{(2)}\}$ and $\{X_j^{(3)}\}$, it is still i.i.d. and hence

$$\mathbf{P}(B(k,y))\,\mathbf{P}(B_2)\,\mathbf{P}(B_3) = \mathbf{P}(B_1 \mid S_k = -y)\,\mathbf{P}(B_2)\,\mathbf{P}(B_3) \le \mathbf{P}(B_4). \tag{3.16}$$

Now the event B_2 is the intersection of two increasing events, so by the Harris inequality, Proposition 3.1 and the central limit theorem,

$$\mathbf{P}(B_2) \ge \mathbf{P}(A(0; f(3n)^2))\,\mathbf{P}(S_{f(3n)^2} \ge f(3n)) \ge \frac{c_{15}}{f(3n)} \tag{3.17}$$

for some positive c_{15} when n [and hence $f(3n)$] are sufficiently large.

Similarly, by Lemma 3.3(iii) and the Harris inequality,

$$\mathbf{P}(B_3) \ge \left(\tfrac{1}{2} - o(1)\right)\mathbf{P}(A(-\sqrt{n}; 3n-k)) \ge c_{16} > 0. \tag{3.18}$$

Together with (3.16) and (3.17) this yields

$$\mathbf{P}(B_1 \mid S_k = -y) \le c_{17} f(3n)\,\mathbf{P}(B_4). \tag{3.19}$$

By (3.15),

$$\mathbf{P}(A(-f;k,N) \mid S_k = -y) \le c_{18} f(3n)\,\mathbf{P}(B_4)$$

and since this is true for any choice of y, we integrate out y to get

$$\mathbf{P}(A(-f;k,N)) \le c_{18} f(3n)\,\mathbf{P}(B_4).$$

Finally, recalling (3.14) and the definition of B_4, we obtain

$$\begin{aligned}\mathbf{P}(n \le T_0 < 2n \text{ and } A(-f;N)) &\le c_{18}\,\mathbf{P}(T_0 \ge n) f(3n)\,\mathbf{P}(B_4)\\ &\le c_{19} n^{-1/2} f(3n)\,\mathbf{P}(T_0 \ge n \text{ and } A(-f;N)),\end{aligned}$$

which is equivalent to (3.12). The assumption [made after (3.17)] that n is large can be removed by taking c_{12} large enough. This completes the proof of Theorem 3.2. □

4. Target percolation on trees. In this section only, it will be convenient to consider finite as well as infinite trees. Among finite trees, we allow only those of constant height, that is, all maximal paths from the root (still called rays) have the same length N. Thus the set $\partial\Gamma$ of rays may be identified with Γ_N. The definitions of energy and capacity remain the same. The definition of Hausdorff measure fails, since the cutset Π cannot go to infinity; replacing the lim inf by an infimum defines the *Hausdorff content*.

Following Lyons (1992), we consider a very general *percolation process* on Γ: a random subgraph $W \subseteq \Gamma$ chosen from some arbitrary distribution on sets of vertices in Γ. The event that W contains the path connecting ρ and σ is denoted $\{\rho \leftrightarrow \sigma\}$; similarly write $\{\rho \leftrightarrow \partial\Gamma\}$ for the event that W contains a ray of Γ. A familiar example from percolation theory is when each edge e is retained independently with some probability $p(e)$. The random component W of this subgraph that contains the root is called a *Bernoulli percolation* on Γ. Another example that is general enough to include nearly all cases of interest is a *target* percolation. This is defined from a family of i.i.d. real random variables $\{X(\sigma): \sigma \ne \rho\}$ by choosing some closed set $B \subseteq \mathbf{R}^N$ and defining

$$W = \bigcup_{k=0}^{N} \{\sigma \in \Gamma_k : (X(\tau_1), \dots, X(\tau_k)) \in \pi_k B\},$$

where π_k is the projection of B on the first k coordinates, the sequence ρ, $\tau_1, \dots, \tau_k = \sigma$ is the path from the root to σ, and ρ is defined always to be in W. The set B is called the target set. Observe that since B is closed, a ray $\xi = (\rho, \sigma_1, \sigma_2, \dots)$ is in W if and only if $(X(\rho), X(\sigma_1), \dots) \in B$. Letting $\{X(\sigma)\}$ all be uniform on $[0,1]$ and $B = \{\prod_{j=1}^{\infty}[0, a_j]\}$ for some $a_j \in [0,1]$ recovers a class of Bernoulli percolations.

The following lemma, which will be sharpened below, is contained in the results of Lyons (1992). Because of notational differences, the brief proof is included.

LEMMA 4.1. *Consider a percolation in which $\mathbf{P}(\rho \leftrightarrow \sigma) = p(|\sigma|)$ for some strictly positive function p.*

(i) *First moment method: $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ is bounded above by the Hausdorff content of Γ in the gauge $p(n)$. If Γ is infinite and has zero Hausdorff measure in gauge $\{p(n)\}$, then $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) = 0$.*

(ii) *Second moment method: Suppose further that there is a positive, nonincreasing function* $g\colon \mathbf{Z}^+ \to \mathbf{R}$ *such that for any two vertices* $\sigma, \tau \in \Gamma_n$ *with* $|\sigma \wedge \tau| = k$,

$$\mathbf{P}(\rho \leftrightarrow \sigma \text{ and } \rho \leftrightarrow \tau) \le \frac{p(n)^2}{g(k)}. \tag{4.1}$$

Then $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \ge \mathrm{Cap}_g(\Gamma)$.

REMARK. For Bernoulli percolations, (4.1) holds with equality for $g(k) = p(k)$. More generally, when $g(k) = p(k)/M$ for some constant $M > 0$, the percolation is termed *quasi-Bernoulli* [Lyons (1989)]. In this case,

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \ge \frac{\mathrm{Cap}_p(\Gamma)}{M}. \tag{4.2}$$

Note also that any target percolation satisfies the condition in the lemma: $\mathbf{P}(\rho \leftrightarrow \sigma) = p(|\sigma|)$.

PROOF OF LEMMA 4.1. (i) For any cutset Π,

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \le \mathbf{P}(\rho \leftrightarrow \sigma \text{ for some } \sigma \in \Pi) \le \sum_{\sigma\in\Pi} p(|\sigma|).$$

The assertion follows by taking the infimum over cutsets.

(ii) First assume that Γ has finite height N. Let μ be a probability measure on $\partial\Gamma = \Gamma_N$. Consider the random variable

$$Y_N = \sum_{\sigma\in\Gamma_N} \mu(\sigma)\mathbf{1}_{\rho\leftrightarrow\sigma}.$$

Clearly $\mathbf{E}\,Y_N = p(N)$ and

$$\begin{aligned}\mathbf{E}\,Y_N^2 &= \mathbf{E}\sum_{\sigma,\tau\in\Gamma_N} \mu(\sigma)\mu(\tau)\mathbf{1}_{\rho\leftrightarrow\sigma \text{ and } \rho\leftrightarrow\tau}\\ &\le \sum_{\sigma,\tau\in\Gamma_N} \mu(\sigma)\mu(\tau)\frac{p(N)^2}{g(|\sigma\wedge\tau|)}\\ &= p(N)^2 I_g(\mu)\end{aligned}$$

by the definition of the energy I_g. The Cauchy–Schwarz inequality gives

$$p(N)^2 = \mathbf{E}(Y_N \mathbf{1}_{Y_n>0})^2 \le \mathbf{E}\,Y_N^2\,\mathbf{P}(Y_N > 0)$$

and dividing the previous inequality by $\mathbf{E}\,Y_N^2$ gives $1 \le I_g(\mu)\mathbf{P}(Y_N > 0)$. Since $\mathrm{Cap}_g(\Gamma)$ is the supremum of $I_g(\mu)^{-1}$ over probability measures μ, it follows that $\mathbf{P}(Y_N > 0) \ge \mathrm{Cap}_g(\Gamma)$. The case where $N = \infty$ is obtained from a straightforward passage to the limit. $\square$

For infinite trees, we are primarily interested in whether $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ is positive. The above lemma fails to give a sharp answer even for quasi-Bernoulli percolations, since the condition $\mathrm{Cap}_p(\Gamma) = 0$ does not imply zero Hausdorff measure. We believe but cannot prove the following.

CONJECTURE 1. *For any target percolation on an infinite tree with* $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$, *the tree* Γ *must have positive capacity in gauge* $p(n)$, *where* $p(|\sigma|) = \mathbf{P}(\rho \leftrightarrow \sigma)$.

To see why we restrict to target percolations, let Γ be the infinite binary tree, let ξ be a ray chosen uniformly from the canonical measure on $\partial\Gamma$ and let $W = \xi$. Then $\mathbf{P}(\rho \to \partial\Gamma) = 1$, but Γ has zero capacity in gauge $p(n) = 2^{-n}$. Evans (1992) gives a capacity criterion on the target set B necessary and sufficient for $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ in the special case where Γ is a homogeneous tree (every vertex has the same degree). His work was extended by Lyons (1992), who showed that $\mathrm{Cap}_p(\Gamma) > 0$ was necessary for $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ for all Bernoulli percolations and for non-Bernoulli percolations satisfying a certain condition.

Specific non-Bernoulli target percolations are used in Lyons (1989) to analyze the Ising model and in Lyons and Pemantle (1992), Benjamini and Peres (1994b) and Pemantle and Peres (1994) to determine the speed of first-passage percolation. In the present work, the special case

$$B = \left\{ \mathbf{x} \in \mathbf{R}^{\infty} : \sum_{i=1}^{n} x_i \geq 0 \text{ for all } n \right\}$$

will play a major role. Unfortunately, this set does not satisfy Lyons' (1992) condition. It is, however, an *increasing set*, meaning that if $\mathbf{x} \geq \mathbf{y}$ componentwise and $\mathbf{y} \in B$, then $\mathbf{x} \in B$. This motivates the next lemma.

LEMMA 4.2 (sharpened first moment method). *Consider a target percolation on a tree* Γ *in which the target set* B *is an increasing set. Assume that* $p(n) \stackrel{\text{def}}{=} p(\rho \leftrightarrow \sigma)$ *for* $\sigma \in \Gamma_n$ *goes to zero as* $n \to \infty$.

(i) *With probability one, the number of surviving rays (elements of* W*) is either zero or infinite.*

(ii) *If* $\partial\Gamma$ *has* σ*-finite Hausdorff measure in the gauge* $\{p(n)\}$, *then* $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) = 0$.

REMARK. With further work, we can show that the assumption that B is an increasing set may be dropped; since this is the only case we need, we impose the assumption to greatly simplify the proof. The above example shows that the "target" assumption cannot be dropped.

PROOF OF LEMMA 4.2. Assume that $\mathbf{P}(\text{ finitely many rays survive}) > 0$. Let A_k denote the event that exactly k rays survive and fix a k for which

$\mathbf{P}(A_k) > 0$. Let $\mathscr{F}_n$ denote the σ-field generated by $\{X(\sigma): |\sigma| \le n\}$. Convergence of the martingale $\mathbf{P}(A_k \mid \mathscr{F}_n)$ shows that for sufficiently large n the probability that $\mathbf{P}(\text{exactly } k \text{ surviving rays} \mid \mathscr{F}_n) > 0.99$ is positive. Let $\{x(\sigma): |\sigma| \le n\}$ be a set of values for which

$$\mathbf{P}(\text{exactly } k \text{ rays survive} \mid X(\sigma) = x(\sigma): |\sigma| \le n) > 0.99.$$

Totally order the rays of Γ in any way; since the probability of any fixed ray surviving is zero, it follows that there is a ray ξ_0 such that

$$\mathbf{P}(\text{some } \xi < \xi_0 \text{ survives} \mid X(\sigma) = x(\sigma): |\sigma| \le n) = \tfrac{1}{2}.$$

This implies that

$$\mathbf{P}(\text{at least } k \text{ rays } \xi > \xi_0 \text{ survive} \mid X(\sigma) = x(\sigma): |\sigma| \le n) \ge 0.49.$$

Since all the $X(\sigma)$ are conditionally independent given $X(\sigma) = x(\sigma)$ for $|\sigma| \le n$, and since the events of at least one ray less than ξ_0 surviving and at least k rays greater than ξ_0 surviving are both increasing, we may apply the FKG inequality to conclude that

$$\mathbf{P}(\text{at least } k+1 \text{ rays survive} \mid X(\sigma) = x(\sigma): |\sigma| \le n) \ge \frac{0.49}{2}.$$

This contradicts the choice of $x(\sigma)$, so we conclude that

$$\mathbf{P}(\text{finitely many rays survive}) = 0.$$

For part (ii), write $\Gamma = \bigcup_{n=1}^{\infty} \Gamma_k$ where each $\partial\Gamma_k$ has finite Hausdorff measure h_k in the gauge $\{p_n\}$. For each k there are cutsets $\Pi_k^{(j)}$ of Γ_k tending to infinity such that

$$\sum_{\sigma \in \Pi_k^{(j)}} p(|\sigma|) \to h_k$$

and therefore the expected number of surviving rays of Γ_k is at most h_k. Since the number of surviving rays of Γ_k is either 0 or infinite, we conclude it is almost surely 0. □

The remainder of this section makes progress toward Conjecture 1 by proving that positive capacity is necessary for $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ in some useful cases. We do this by comparing different target percolations, varying either Γ or the set B. The notation $p(n) = \mathbf{P}(\rho \leftrightarrow \sigma)$ for $\sigma \in \Gamma_n$ is written $p(B; n)$ when we want to emphasize dependence on B; similarly, $\mathbf{P}(B; \cdots)$ reflects dependence on B. It is assumed that the common distribution of the $X(\sigma)$ defining the target percolation never change, since this may always be accomplished through a measure-theoretic isomorphism. It will be seen below (and this makes the comparison theorems useful) that spherically symmetric trees are much easier to handle than general trees. This is partly because $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ may be calculated recursively by conditioning. In particular, if $f(n)$ is the growth function for Γ [i.e., each $\sigma \in \Gamma_{n-1}$ has $f(n)$ neighbors in Γ_n], $\Gamma(\sigma)$ is the subtree of Γ rooted at a vertex $\sigma \in \Gamma_1$, and $B/x \subseteq \mathbf{R}^{N-1} = \{y \in \mathbf{R}^{N-1}: (x, y_1, y_2, \dots, y_{N-1}) \in B\}$ is the cross section of B at x, then conditioning on $X(\sigma)$ for $\sigma \in \Gamma_1$ gives

$$\mathbf{P}(B; \rho \not\leftrightarrow \partial\Gamma) = [\mathbf{E}\,\mathbf{P}(B/X(\sigma); \sigma \not\leftrightarrow \partial\Gamma(\sigma))]^{f(1)}. \tag{4.3}$$

Notice that this recursion makes sense if $f(1)$ is any positive real, not necessarily an integer. As a notational convenience, we define $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ for *virtual* spherically symmetric trees with positive real growth functions, f, by (4.3) for trees of finite height and passage to the limit for infinite trees. Specifically, if B is any target set and Γ is any tree, we let $f(n) = |\Gamma_n|/|\Gamma_{n-1}|$ and define the symbol $\mathscr{S}(\Gamma)$ to stand for the spherical symmetrization of Γ in the sense that the expression $\mathbf{P}(B; \rho \not\leftrightarrow \partial\mathscr{S}(\Gamma))$ is defined to stand for the value of the function $\Psi(B; f(1), \dots, f(N))$, where Ψ is defined by the following recursion in which X is a random variable with the common distribution of the $X(\sigma)$:

$$\Psi(B; a) = \mathbf{P}(X \notin B)^a \quad \text{if } B \subseteq \mathbf{R} \text{ and } a > 0;$$

$$\Psi(B; a_1, \dots, a_n) = [\mathbf{E}\,\Psi(B/X; a_2, \dots, a_n)]^{a_1} \quad \text{if } B \subseteq \mathbf{R}^n \text{ and } a \in (\mathbf{R}^n)^+.$$

THEOREM 4.3. *Let Γ be any tree of height $N \leq \infty$, let $B \subset \mathbf{R}^N$ be any target set and let $\mathscr{S}(\Gamma)$ be the (virtual) spherically symmetric tree with $f(n) = |\Gamma_n|/|\Gamma_{n-1}|$. Let $\mathscr{S}(B)$ be a Cartesian product target set $\{y \in \mathbf{R}^N: y_i \leq b_i$ for all finite $i \leq N\}$ with b_i chosen so that*

$$\prod_{i=1}^{n} \mathbf{P}(X(\sigma) \leq b_i) = p(B; n).$$

Then

$$\mathbf{P}(B; \rho \leftrightarrow \partial\Gamma) \leq \mathbf{P}(B; \rho \leftrightarrow \partial\mathscr{S}(\Gamma)) \tag{4.4}$$

$$\leq \mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\mathscr{S}(\Gamma)) \tag{4.5}$$

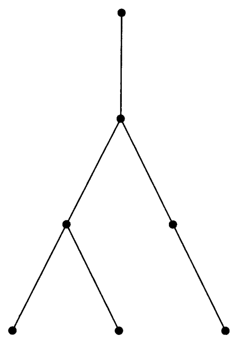
$$\leq 2\left[p(B; N)^{-1}|\Gamma_N|^{-1} + \sum_{k=0}^{N-1} p(B; k)^{-1}(|\Gamma_k|^{-1} - |\Gamma_{k+1}|^{-1})\right]^{-1}, \tag{4.6}$$

where the $p(B; N)^{-1}|\Gamma_N|^{-1}$ term appears only if $N < \infty$. If $\mathscr{S}(\Gamma)$ exists as a tree, that is, $f(n)$ is an integer for all n, then expression (4.6) *is precisely* $2\,\mathrm{Cap}_p(\Gamma)$.

When Γ is spherically symmetric, $\mathscr{S}(\Gamma) = \Gamma$ and we get:

COROLLARY 4.4. *If Γ is spherically symmetric, then $\mathrm{Cap}_p(\Gamma) > 0$ is necessary for $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ in any target percolation.*

REMARK AND COUNTEREXAMPLE. The inequality $\mathbf{P}(B; \rho \leftrightarrow \partial\Gamma) \leq \mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\Gamma)$ holds for spherically symmetric trees by (4.5) and also for certain types of target sets B (see Theorem 4.6 below), but fails in general. Whenever this inequality holds, Lyons' (1992) result that $\mathrm{Cap}_p(\Gamma) > 0$ is necessary for $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ in Bernoulli percolation implies Conjecture 1 for that case, since $\mathscr{S}(B)$ defines a Bernoulli percolation. A counterexample to the general inequality is the tree:



Γ

Let $X(\sigma)$ be uniform on $[0,1]$ and define

$$(B = [0,1/2] \times [2\varepsilon, 1] \times [0,1]) \cup ([1/2,1] \times [0,1] \times [4\varepsilon, 1]).$$

Then

$$\mathscr{S}(B) = [0,1] \times [0, 1-\varepsilon] \times \left[0, \frac{1-3\varepsilon}{1-\varepsilon}\right]$$

and

$$\mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\Gamma) \le 1 - 3\varepsilon^2 < 1 - 2\varepsilon^2 - 32\varepsilon^3 = \mathbf{P}(B; \rho \leftrightarrow \partial\Gamma)$$

for sufficiently small ε.

QUESTION. Is there an infinite tree, Γ, and a target set, B, for which

$$\mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\Gamma) = 0 < \mathbf{P}(B; \rho \leftrightarrow \partial\Gamma)?$$

The proof of Theorem 4.3 is based on the following convexity lemma.

LEMMA 4.5. *For any tree Γ of height $n < \infty$, define the function $h_\Gamma(z_1, \ldots, z_n)$ for arguments $1 \ge z_1 \ge \cdots \ge z_n \ge 0$ by*

$$h_\Gamma(z_1, \ldots, z_n) = \mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \Gamma_n),$$

where B defines a target percolation with $\mathbf{P}(B; \rho \leftrightarrow \sigma) = z_{|\sigma|}$. If Γ is spherically symmetric, then h_Γ is a convex function. The same holds for virtual trees, under the restriction that the growth function $f(n)$ is always greater than or equal to one.

PROOF. Proceed by induction on n. When $n = 1$, certainly $h_\Gamma(z_1) = (1 - z_1)^{|\Gamma_1|}$ is convex. Now assume the result for trees of height $n - 1$ and let $\Gamma(\sigma)$ denote the subtree of Γ rooted at σ: $\{\tau: \tau \geq \sigma\}$. Since Γ is spherically symmetric, all subtrees $\Gamma(\sigma)$ with $\sigma \in \Gamma_1$ are isomorphic, spherically symmetric trees of height $n - 1$. By definition of h_Γ,

$$h_\Gamma(z_1, \dots, z_n) = \left[(1 - z_1) + z_1 h_{\Gamma(\sigma)}\left(\frac{z_2}{z_1}, \dots, \frac{z_n}{z_1}\right)\right]^{|\Gamma_1|},$$

where $\sigma \in \Gamma_1$. By induction, the function $h_{\Gamma(\sigma)}$ is convex. This implies that

$$g(z_1, \dots, z_n) = z_1 h_{\Gamma(\sigma)}\left(\frac{z_2}{z_1}, \dots, \frac{z_n}{z_1}\right)$$

is convex: since g is homogeneous of degree one, it suffices to check convexity on the affine hyperplane $\{z_1 = 1\}$, where it is clear. Adding a linear function to g and taking a power of at least one preserves convexity, so h_Γ is convex, completing the induction. □

PROOF OF THEOREM 4.3. The first inequality is proved in Pemantle and Peres (1994). For the second inequality it clearly suffices to consider trees of finite height, N. When $N = 1$ there is nothing to prove, so fix $N > 1$ and assume for induction that the inequality holds for trees of height $N - 1$. Define $\Gamma(\sigma)$, h_Γ, B/x and $p(B; k)$ as previously, and observe that for every $j < N$,

$$\mathbf{E}\, p(B/X; j - 1) = p(B; j),$$

where X has the common distribution of the $\{X(\sigma)\}$. Use the induction hypothesis and the fact that $X(\sigma)$ are independent to get

$$\begin{aligned}
&\mathbf{P}(B; \rho \not\leftrightarrow \partial\Gamma)\\
&\quad = \prod_{\sigma\in\Gamma_1}\left[1 - p(B;1) + p(B;1)\, \mathbf{E}\mathbf{P}\left(\frac{B}{X}; \sigma \not\leftrightarrow \partial\Gamma(\sigma)\right)\right]\\
&\quad \geq \left[1 - p(B;1) + p(B;1)\, \mathbf{E}\, h_{\Gamma(\sigma)}\left(\frac{p(B/X;1)}{p(B;1)}, \dots, \frac{p(B/X;N-1)}{p(B;1)}\right)\right]^{|\Gamma_1|}.
\end{aligned}$$

Utilizing the convexity of $h_{\Gamma(\sigma)}$ and Jensen's inequality, the last expression is at least

$$\begin{aligned}
&\left[1 - p(B;1) + p(B;1)\, h_\Gamma\left(\frac{p(B;2)}{p(B;1)}, \dots, \frac{p(B;N)}{p(B;1)}\right)\right]^{|\Gamma_1|}\\
&\qquad = h_\Gamma(p(B;1), \dots, p(B;N))\\
&\qquad = \mathbf{P}(\mathscr{S}(B); \rho \not\leftrightarrow \partial\Gamma),
\end{aligned}$$

completing the induction and the proof of the second inequality.

When Γ is spherically symmetric, Theorem 2.1 of Lyons (1992) asserts that

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \leq 2\,\mathrm{Cap}_p(\Gamma).$$

The measure μ that minimizes $I_p(\mu)$ for spherically symmetric trees is easily seen to be uniform, with

$$I_p(\mu) = \int\int p(|\xi \wedge \eta|)^{-1} d\mu^2;$$

summing by parts shows that the RHS of (4.6) is equal to $2I_p(\mu)^{-1}$. When Γ is not spherically symmetric, the final inequality is proved by induction, as follows.

Fix B and let $p_n = \mathbf{P}(B; \rho \leftrightarrow \sigma) = \mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \sigma)$ for $|\sigma| = n$. Let $\psi_p(\lambda_1, \dots, \lambda_N)$ denote $\mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\Gamma)$ for a (possibly virtual) spherically symmetric tree Λ whose growth numbers $f(n)$ satisfy $\prod_{i=0}^{k-1} f(i) = \lambda_k$. Write $R_p(\lambda_1, \dots, \lambda_N)$ for the "electrical resistance" of this tree when edges at level i are assigned resistance $p_i^{-1} - p_{i-1}^{-1}$ and an additional unit resistor is attached to the root. Explicitly, define

$$R_p(\lambda_1, \dots, \lambda_N) = p_N^{-1}\lambda_N^{-1} + \sum_{i=0}^{N-1} p_i^{-1}(\lambda_i^{-1} - \lambda_{i+1}^{-1}),$$

where $\lambda_0 \stackrel{\text{def}}{=} 1$, so the inequality to be proved is

$$\psi_p(\lambda_1, \dots, \lambda_N) \le 2R_p(\lambda_1, \dots, \lambda_N)^{-1}. \tag{4.7}$$

Proceed by induction, the case $N = 0$ boiling down to $1 \le 2$. Letting $p_i' = p_i/p_1$, we have

$$\psi_p(\lambda_1, \dots, \lambda_N) = 1 - \left[1 - p_1\psi_{p'}\left(\frac{\lambda_2}{\lambda_1}, \dots, \frac{\lambda_N}{\lambda_1}\right)\right]^{\lambda_1}.$$

Using the elementary inequality

$$\frac{1 - x^{\lambda_1}}{1 + x^{\lambda_1}} \le \lambda_1 \frac{1 - x}{1 + x},$$

valid for all $x \in [0, 1]$ and $\lambda_1 \ge 1$, we obtain

$$\begin{aligned}\frac{\psi_p(\lambda_1, \dots, \lambda_N)}{2 - \psi_p(\lambda_1, \dots, \lambda_N)} &= \frac{1 - [1 - p_1\psi_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)]^{\lambda_1}}{1 + [1 - p_1\psi_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)]^{\lambda_1}} \\ &\le \frac{\lambda_1 p_1 \psi_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)}{2 - p_1\psi_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)}.\end{aligned}$$

Applying the inductive hypothesis, this is at most

$$\frac{2p_1\lambda_1 R_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)^{-1}}{2 - 2p_1 R_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)^{-1}} = \frac{p_1\lambda_1}{R_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1) - p_1}.$$

The last expression may be simplified using

$$R_p(\lambda_1, \dots, \lambda_N) = 1 - \lambda_1^{-1} + p_1^{-1}\lambda_1^{-1} R_{p'}(\lambda_2/\lambda_1, \dots, \lambda_N/\lambda_1)$$

to get

$$\frac{\psi_p(\lambda_1,\dots,\lambda_N)}{2-\psi_p(\lambda_1,\dots,\lambda_N)} \le \frac{p_1\lambda_1}{p_1\lambda_1(R_p(\lambda_1,\dots,\lambda_N)-1)} = \frac{2R_p(\lambda_1,\dots,\lambda_N)^{-1}}{2-2R_p(\lambda_1,\dots,\lambda_N)^{-1}}.$$

This proves (4.7) and the theorem. □

The last result of this section gives a condition on the target set B, sufficient to imply that $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ increases when B is replaced by $\mathscr{S}(B)$. The condition is rather strong; however, it can be applied in two very natural cases; see Theorem 5.1 below.

THEOREM 4.6. *Let Γ be any tree of height $N \le \infty$, let $X(\sigma)$ be i.i.d. real random variables and let B be any target set. For integers $k \le j \le N$ and real numbers $x_1,\dots,x_k$, define*

$$p_j(x_1,\dots,x_k) = \mathbf{P}((X_{k+1},\dots,X_j) \in \pi_j(B)/x_1\cdots x_k)$$

to be the measure of the cross section at $x_1,\dots,x_k$ of the projection onto the first j coordinates of B. Suppose that for every fixed $x_1,\dots,x_{k-1}$, the matrix M whose (y,j)-entry is $p_j(x_1,\dots,x_{k-1},y)$ is totally positive of order two (TP_2), that is, $M_{xi}M_{yj} \ge M_{yi}M_{xj}$ when $k \le i < j$ and $x < y$. Then

$$\mathbf{P}(B;\rho \leftrightarrow \partial\Gamma) \le \mathbf{P}(\mathscr{S}(B);\rho \leftrightarrow \partial\Gamma). \tag{4.8}$$

We shall require the following version of Jensen's inequality.

LEMMA 4.7. *Let $\trianglelefteq$ be a partial order on $\mathbf{R}^n$ such that for every $w \in \mathbf{R}^n$ the set $\{z: z \trianglelefteq w\}$ is convex. Let μ be a probability measure supported on a bounded subset of $\mathbf{R}^n$, which is totally ordered by $\trianglelefteq$, and let $h:\mathbf{R}^n \to \mathbf{R}$ be a continuous function. If h is convex on any segment connecting two comparable points $z \trianglelefteq w$, then*

$$h\Big(\int x\,d\mu\Big) \le \int h(x)\,d\mu.$$

PROOF. It is enough to prove this in the case where μ has finite support, since we may approximate any measure by measures supported on finite subsets and use continuity of h and bounded support to pass to the limit. Letting $z_1 \trianglelefteq \cdots \trianglelefteq z_m$ denote the support of μ and letting $a_i = \mu\{z_i\}$, we proceed by induction on m. If $m = 2$, the desired inequality $h(a_1z_1 + a_2z_2) \le a_1h(z_1)+a_2h(z_2)$ is a direct consequence of the assumption on h. If $m > 2$, let ν be the measure which puts mass a_1+a_2 at $(a_1z_1+a_2z_2)/(a_1+a_2)$ and mass a_i at z_i for each $i \ge 3$. The support of ν is a totally ordered set of cardinality $n - 1$, so the induction hypothesis implies

$$h\Big(\int x\,d\mu\Big) = h\Big(\int x\,d\nu\Big) \le \int h(x)\,d\nu.$$

Applying the convexity assumption on h at z_1 and z_2 then gives

$$\int h(x)\,d\nu \le \int h(x)\,d\mu,$$

completing the induction. □

PROOF OF THEOREM 4.6. For each n, let Δ_n denote the space of points $\{(z_1,\dots,z_n) \in \mathbf{R}^n : 1 \ge z_1 \ge \cdots \ge z_n \ge 0\}$ and for $\mathbf{z},\mathbf{w} \in \Delta_n$, define $\mathbf{z} \preceq \mathbf{w}$ if and only if the matrix with rows $\mathbf{z}$ and $\mathbf{w}$ and first column $(1,1)$ is TP_2 (equivalently, z_i/w_i is at most one and nonincreasing in i). Define h_Γ on $\bigcup \Delta_n$ as in Lemma 4.5 so that

$$h_\Gamma(z_1,\dots,z_n) = \prod_{\sigma\in\Gamma_1}\left[(1-z_1) + z_1 h_{\Gamma(\sigma)}\left(\frac{z_2}{z_1},\dots,\frac{z_n}{z_1}\right)\right].$$

In order to use Lemma 4.7 for h_Γ, $\preceq$ and Δ_n, we observe first that $\preceq$ is closed under convex combinations in either argument. Observe also that Δ_n is compact and h_Γ is continuous; we now establish by induction on n that h_Γ is convex along the line segment joining $\mathbf{z}$ and $\mathbf{w}$ whenever $\mathbf{z} \preceq \mathbf{w} \in \Delta_n$.

The initial step is immediate: $h_\Gamma(z_1) = (1-z_1)^{|\Gamma_1|}$, which is convex for $z_1 \in [0,1]$. When $n > 1$, observe that for each $\sigma \in \Gamma_1$, the function $1 - z_1 + z_1 h_{\Gamma(\sigma)}(z_2/z_1,\dots,z_n/z_1)$ is decreasing along the line segment from $\mathbf{z}$ to $\mathbf{w}$ when $\mathbf{z} \preceq \mathbf{w}$. The product of decreasing convex functions is again convex, so it suffices to check that each

$$\phi(z_1,\dots z_n) \stackrel{\text{def}}{=} z_1 h_{\Gamma(\sigma)}(z_2/z_1,\dots,z_n/z_1)$$

is convex along such a line segment. Pictorially, we must show that the graph of ϕ in $\Delta_n \times \mathbf{R}$ defined by $\{(z_1,\dots,z_{n+1}) : z_{n+1} = z_1 h_{\Gamma(\sigma)}(z_2/z_1,\dots,z_n/z_1)\}$ lies below any chord $\overline{(\mathbf{z},\phi(\mathbf{z}))(\mathbf{w},\phi(\mathbf{w}))}$ whenever $\mathbf{z} \preceq \mathbf{w}$. Observe that the graph of ϕ is the cone of the set $\{(1,z_2,\dots,z_{n+1}) : z_{n+1} = h_{\Gamma(\sigma)}(z_2,\dots,z_n)\}$ with the origin. In other words, viewing Δ_{n-1} as embedded in Δ_n by $(z_2,\dots,z_n) \mapsto (1,z_2,\dots,z_n)$, the graph of ϕ is the cone of the graph of the $n-1$-argument function $h_{\Gamma(\sigma)}$. To check that the chord of the graph of ϕ between $\mathbf{z}$ and $\mathbf{w}$ lies above the graph, it then suffices to see that the chord of the graph of $h_{\Gamma(\sigma)}$ between $(z_2/z_1,\dots,z_n/z_1)$ and $(w_2/w_1,\dots,w_n/w_1)$ lies above the graph. However, $\mathbf{z} \preceq \mathbf{w} \in \Delta_n$ implies $(z_2/z_1,\dots,z_n/z_1) \preceq (w_2/w_1,\dots,w_n/w_1) \in \Delta_{n-1}$, so this follows from the induction hypothesis.

We now prove the theorem for trees of finite height, the infinite case following from writing $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ as the decreasing limit of $\mathbf{P}(\rho \leftrightarrow \Gamma_n)$. Let $N < \infty$ and proceed by induction on N, the case $N = 1$ being trivial since $B = \mathscr{S}(B)$. Assume therefore that $N > 1$ and that the theorem is true for smaller values of N.

The induction is then completed by justifying the following chain of identities and inequalities. By conditioning on the independent random variables

$\{X(\sigma): \sigma \in \Gamma_1\}$ and using the induction hypothesis we get:

$$
\begin{aligned}
\mathbf{P}(B; \rho \not\leftrightarrow \partial\Gamma) &= \prod_{\sigma \in \Gamma_1} \mathbf{E}\,\mathbf{P}(B/X; \sigma \not\leftrightarrow \partial\Gamma(\sigma)) \\
&\geq \prod_{\sigma \in \Gamma_1} \mathbf{E}\,\mathbf{P}(\mathscr{S}(B)/X; \sigma \not\leftrightarrow \partial\Gamma(\sigma)).
\end{aligned} \tag{4.9}
$$

Recalling the definition of $h_{\Gamma(\sigma)}$ and $p_j(x)$, this is equal to

$$
\begin{aligned}
&\prod_{\sigma \in \Gamma_1} \mathbf{E}[h_{\Gamma(\sigma)}(p_2(X), \dots, p_N(X))] \\
&\quad = \prod_{\sigma \in \Gamma_1} \Big[(1 - p_1) + p_1 \int h_{\Gamma(\sigma)}(p_2(x), \dots, p_N(x))\, d\mu(x)\Big],
\end{aligned} \tag{4.10}
$$

where μ is the conditional distribution of X given $X \in \pi_1(B)$ and $p_1 = \mathbf{P}(X \in \pi_1(B))$.

Now observe that the vectors $(p_2(x), \dots, p_N(x))$ with $x \in \pi_1(B)$ are totally ordered by $\preceq$ according to the $k = 1$ case of the $T\mathbf{P}_2$ assumption of the theorem. Since $\int p_j(x)\, d\mu(x) = p_j/p_1$ for $2 \leq j \leq N$, Lemma 4.7 applied to $h_{\Gamma(\sigma)}$ and $\preceq$ on the set $\triangle_{N-1}$ shows that (4.10) is at least

$$
\begin{aligned}
&\prod_{\sigma \in \Gamma_1} \left[(1 - p_1) + p_1 h_{\Gamma(\sigma)}\left(\frac{p_2}{p_1}, \dots, \frac{p_N}{p_1}\right)\right] \\
&\quad = h_\Gamma(p_1, \dots, p_N) \\
&\quad = \mathbf{P}(\mathscr{S}(B); \rho \not\leftrightarrow \partial\Gamma).
\end{aligned} \tag{4.11}
$$

Comparing (4.9) to (4.11) we see that the theorem is established. □

5. Positive rays for tree-indexed random walks. This section puts together results from the previous two sections in order to prove Theorem 2.2 and a few related corollaries and examples.

PROOF OF THEOREM 2.2. (i) We are given a tree Γ with positive capacity in gauge $\phi(n) = n^{-1/2}$ and i.i.d. real random variables $\{X(\sigma)\}$ with mean zero and finite variance. Consider the target percolation with target set $B^{(0)} = \{\mathbf{x} \in \mathbf{R}^\infty: \sum_{i=1}^n x_i \geq 0 \text{ for all } n\}$. By Proposition 3.1,

$$
c_1'|\sigma|^{-1/2} \leq \mathbf{P}(\rho \leftrightarrow \sigma) \leq c_1''|\sigma|^{-1/2}.
$$

Now we verify that this percolation is quasi-Bernoulli, that is to say,

$$
\mathbf{P}(\rho \leftrightarrow \sigma \text{ and } \rho \leftrightarrow \tau \mid \rho \leftrightarrow \sigma \wedge \tau) \leq c\, \frac{|\sigma \wedge \tau|}{|\sigma|^{1/2}|\tau|^{1/2}} \tag{5.1}
$$

for $\sigma, \tau \in \Gamma$. Assume without loss of generality that $|\sigma \wedge \tau| \le (1/2)\min(|\sigma|, |\tau|)$, since otherwise the claim is immediate. The LHS of (5.1) is equal to

$$\begin{aligned}(5.2)\qquad &\int_0^\infty \mathbf{P}(S(\sigma\wedge\tau)\in dy \mid \rho \leftrightarrow \sigma\wedge\tau)\\ &\quad\times \mathbf{P}(\rho\leftrightarrow\sigma \mid \rho\leftrightarrow\sigma\wedge\tau, S(\sigma\wedge\tau)=y)\\ &\quad\times \mathbf{P}(\rho\leftrightarrow\tau \mid \rho\leftrightarrow\sigma\wedge\tau, S(\sigma\wedge\tau)=y).\end{aligned}$$

Recalling the definition of T_y as the first time a trajectory is less than $-y$, we may write

$$\begin{aligned}&\mathbf{P}(\rho\leftrightarrow\sigma \mid \rho\leftrightarrow\sigma\wedge\tau, S(\sigma\wedge\tau)=y)\\ &\quad= \mathbf{P}(T_y \ge |\sigma| - |\sigma\wedge\tau|) \le 4c_2\frac{y}{|\sigma|^{1/2}}\end{aligned}$$

by part (i) of Lemma 3.3. A similar bound holds for the last factor in (5.2). Now use the the second part of Lemma 3.3 to show that (5.2) is at most

$$\begin{aligned}&4c_2^2\int_0^\infty \mathbf{P}(S(\sigma\wedge\tau)\in dy \mid \rho\leftrightarrow\sigma\wedge\tau)\frac{y^2}{|\sigma|^{1/2}|\tau|^{1/2}}\\ &\quad= \frac{4c_2^2}{|\sigma|^{1/2}|\tau|^{1/2}}\mathbf{E}(S(\sigma\wedge\tau)^2 \mid \rho\leftrightarrow\sigma\wedge\tau)\\ &\quad\le 4c_2^2c_3\frac{|\sigma\wedge\tau|}{|\sigma|^{1/2}|\tau|^{1/2}},\end{aligned}$$

verifying the claim.

Putting $|\sigma| = |\tau| = n$ and $|\sigma\wedge\tau| = k$, it immediately follows that $\mathbf{P}(\rho \leftrightarrow \sigma \text{ and } \rho\leftrightarrow\tau) \le c\sqrt{k}/n$, and the second moment method [Lemma 4.1 part (ii)] implies that with positive probability a ray exists along which $S(\sigma)$ remains nonnegative. To obtain the full assertion of the theorem, let $f(n)$ be any increasing sequence satisfying $\sum n^{-3/2}f(n) < \infty$ and define a new target percolation with target set

$$B^{(f)} = \left\{\mathbf{x}\in\mathbf{R}^\infty: \sum_{i=1}^n x_i \ge f(n)\mathbf{1}_{n\ge n_f} \text{ for all } n\ge 1\right\},$$

where n_f is as in Theorem 3.2, part (i). The conclusion of Theorem 3.2, part (i), shows that $\mathbf{P}(B^{(f)};\rho\leftrightarrow\sigma)$ is of order $|\sigma|^{-1/2}$; since $\mathbf{P}(B^{(f)};\rho\leftrightarrow\sigma \text{ and } \rho\leftrightarrow\tau) \le \mathbf{P}(B^{(0)};\rho\leftrightarrow\sigma \text{ and } \rho\leftrightarrow\tau)$, the new percolation $B^{(f)}$ is still quasi-Bernoulli. Thus with positive probability, Γ contains a ray along which $\mathcal{S}(\sigma) \ge f(|\sigma|)$ for sufficiently large $|\sigma|$. To see that this event actually has probability one, observe that it contains the tail event

$$\{\exists\xi\in\partial\Gamma\;\exists C>0\;\forall\sigma\in\xi,\; S(\sigma)\ge f(|\sigma|)+|\sigma|^{1/4}-C\},$$

which has positive probability by the preceding argument, hence probability one.

(ii) Assume that Γ has σ-finite Hausdorff measure in gauge $\phi(n) = n^{-1/2}$. For any nondecreasing function, f, consider the random subgraph $W_{-f} = \{\sigma \in \Gamma: S(\sigma) \geq -f(|\sigma|)\}$. By the summability assumption on f and by Theorem 3.2 part (ii), there is a constant c for which

$$p(|\sigma|) = \mathbf{P}(\rho \leftrightarrow \sigma) \leq c|\sigma|^{-1/2}.$$

The sharpened first moment method (Lemma 4.2) implies that W_{-f} almost surely fails to contain a ray of Γ. This easily implies the stronger statement that the subgraph W_{-f} has no infinite components, almost surely.

(iii) Define W_{-f} and $p(|\sigma|)$ as above. From Theorem 4.3 we get

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \leq 2\left[\sum_{k=0}^{\infty} p(k)^{-1}\left(|\Gamma_k|^{-1} - |\Gamma_{k+1}|^{-1}\right)\right]^{-1}. \tag{5.3}$$

Since $p(k) \leq ck^{-1/2}$, summation by parts shows that

$$\begin{aligned}
&\sum_{k=0}^{\infty} p(k)^{-1}(|\Gamma_k|^{-1} - |\Gamma_{k+1}|^{-1}) \\
&\quad \geq c^{-1} \sum_{k=0}^{\infty} k^{1/2}(|\Gamma_k|^{-1} - |\Gamma_{k+1}|^{-1}) \\
&\quad = c^{-1} \sum_{k=1}^{\infty} [k^{1/2} - (k-1)^{1/2}]|\Gamma_k|^{-1} \\
&\quad \geq (2c)^{-1} \sum_{k=1}^{\infty} k^{-1/2}|\Gamma_k|^{-1},
\end{aligned}$$

so if the last sum is infinite, then the RHS of (5.3) is zero, completing the proof. □

For certain special distributions of the step sizes $\{X(\sigma)\}$, the class of trees for which some ray stays positive with positive probability may be sharply delineated.

THEOREM 5.1. *Let Γ be any infinite tree and let the i.i.d. random variables $\{X(\sigma)\}$ have common distribution F_1 or F_2, where F_1 is a standard normal and F_2 is the distribution putting probability $1/2$ each on ± 1. Then the probability that $S(\sigma) \geq 0$ along some ray of Γ is nonzero if and only if Γ has positive capacity in gauge $\phi(n) = n^{-1/2}$.*

REMARK. The usual variants also follow. When Γ has positive capacity in gauge ϕ, the probability is one that some ray of Γ has $S(\sigma) < 0$ finitely often. This is equivalent to finding, almost surely, a ray for which $S(\sigma) \geq f(|\sigma|)$ all but finitely often, for any monotone f satisfying $\sum n^{-3/2}|f(n)| < \infty$.

PROOF OF THEOREM 5.1. Let $B \subseteq \mathbf{R}^\infty$ be the target set $\{\mathbf{x} \in \mathbf{R}^\infty : \sum_{i=1}^n x_i \geq 0 \text{ for all } n\}$. One half of the theorem, namely, that positive capacity implies $\mathbf{P}(B; \rho \leftrightarrow \partial\Gamma) > 0$, follows immediately from part (i) of Theorem 2.2. For the other half, observe that zero capacity implies $\mathbf{P}(\mathscr{S}(B); \rho \leftrightarrow \partial\Gamma) = 0$, since $\mathscr{S}(B)$ is Bernoulli with the same values of $p(n)$, and $\mathrm{Cap}_p(\Gamma) > 0$ is known to be necessary and sufficient for percolation (cf. remarks after Conjecture 1). The present theorem then follows from (4.8) once the conditions of Theorem 4.6 are verified. It suffices to establish $M_{yj}/M_{xj} > M_{yi}/M_{xi}$ for $y > x$, in the case where $j = i + 1$.

To verify this, pick any j, k and any $x_1, \ldots, x_k \geq 0$ and observe that $p_j(x_1, \ldots, x_k) = \mathbf{P}(x_1 + \cdots + x_k + S_i \geq 0 \text{ for all } i \leq j)$, where $\{S_i\}$ is a random walk with step sizes distributed as the $\{X(\sigma)\}$. It suffices then to show that for $0 \leq x < y$,

$$\mathbf{P}(y + S_{j+1} \geq 0 \mid y + S_i \geq 0 : i \leq j) \geq \mathbf{P}(x + S_{j+1} \geq 0 \mid x + S_i \geq 0 : i \leq j).$$

To see this in the case of F_1, use induction on j. For $j = 0$ one pointmass obviously dominates the other. Assuming it now for $j - 1$, write

$$\begin{aligned}&\mathbf{P}(y + S_{j+1} \geq 0 \mid y + S_i \geq 0 : i \leq j) \\ &\quad = \int d\nu_y(z)\, \mathbf{P}(z + S_j \geq 0 \mid z + S_i \geq 0 : i \leq j - 1),\end{aligned}$$

where $\nu_y(z)$ is the conditional measure of $y + S_1$ given $S_i \geq -y : i \leq j$. The Radon–Nikodym derivative $d\nu_y/d\nu_x$ at $z \geq y$ is $dF_1(z - y)/dF_1(z - x)$ times a normalizing constant. This is an increasing function of z, by the increasing likelihood property of the normal distribution. Thus ν_y stochastically dominates ν_x. By induction, the integrand is increasing in z, which, together with the stochastic domination, establishes the inequality. The same argument works for F_2, noting that $y - x$ is always an even integer. □

COROLLARY 5.2. *Suppose the edges of an infinite tree* Γ *are labeled by i.i.d., mean-zero, finite variance, random variables* $\{X(\sigma)\}$ *with partial sums* $\{S(\sigma)\}$. *Assume that*

(∗) *either* Γ *is spherically symmetric*
or
$\{X(\sigma)\}$ *are normal or take values* ± 1.

If there is almost surely a ray along which $\inf S(\sigma) > -\infty$, *then there is almost surely a ray along which* $\lim S(\sigma) = +\infty$.

PROOF. Both are equivalent to Γ having positive capacity in gauge $n^{-1/2}$. □

PROBLEM. Remove the assumption (∗).

If the moment generating function of X fails to exist in a neighborhood of zero, it is possible that $\mathbf{E}\,X < 0$, but still some trees of polynomial growth have rays along which $S(\sigma) \to \infty$. The critical growth exponent need not be 1/2 in this case. We conclude this section with such an example.

Suppose that the common distribution of the $X(\sigma)$ is a symmetric, stable random variable with index $\alpha \in (1,2)$. Fix $c > 0$ and consider the target set

$$B = \left\{\mathbf{x} \in \mathbf{R}^{\infty}: \sum_{i=1}^{n} x_i > cn \text{ for all } n\right\}$$

and

$$B' = \left\{\mathbf{x} \in \mathbf{R}^{\infty}: cn^2 > \sum_{i=1}^{n} x_i > cn \text{ for all } n\right\}.$$

The following estimates may be proved.

PROPOSITION 5.3. *For $\sigma, \tau \in \Gamma_n$, let $k = |\sigma \wedge \tau|$. Then*

$$\mathbf{P}(B'; \rho \leftrightarrow \sigma) \le \mathbf{P}(B; \rho \to \sigma) \le c_1 n^{-\alpha}, \tag{5.4}$$

$$\mathbf{P}(B'; \rho \leftrightarrow \sigma \text{ and } \rho \leftrightarrow \tau) / \mathbf{P}(B'; \rho \to \sigma)^2 \le c_2(\varepsilon) k^{1+4\alpha+\varepsilon} \tag{5.5}$$

for any $\varepsilon > 0$ and some constants $c_i > 0$.

Suppose that Γ is a spherically symmetric tree with growth rate $|\Gamma_n| \approx n^{\beta}$. Plugging (5.4) into Lemma 4.1, we see that $\mathbf{P}(B'; \rho \leftrightarrow \partial\Gamma)$ is zero when $\beta < \alpha$, while plugging in (5.5) shows that $\mathbf{P}(B'; \rho \leftrightarrow \partial\Gamma)$ is positive when $\beta > (1+4\alpha)$. If one then considers the tree-indexed random walk whose increments are distributed as $X - c$, one sees that $\beta < \alpha$ implies that with probability one $S(\sigma) < 0$ infinitely often on every ray, whereas $\beta > 1 + 4\alpha$ implies that with probability one $S(\sigma) \to \infty$ with at least linear rate along some ray. For $\beta > 1 + 4\alpha$, RWRE with this distribution of X is therefore transient even though $\mathbf{E}\, X < 0$.

Defining the *sustainable speed* of a tree-indexed random walk to be the almost surely constant value

$$\sup_{\xi} \liminf_{\sigma \in \xi} \frac{S(\sigma)}{|\sigma|},$$

Lyons and Pemantle (1992) have shown that the Hausdorff dimension of Γ and the distribution of X together determine the sustainable speed of the tree-indexed random walk, as long as X has a moment generating function in a neighborhood of zero. When the increments are symmetric stable random variables, the moment hypothesis is violated, and the analysis above shows that the sustainable speed of the tree-indexed random walk can be different for different polynomially growing trees of Hausdorff dimension zero.

6. Critical RWRE: proofs. The following easy lemma will be useful.

LEMMA 6.1. *If Γ is any tree with conductances $C(\sigma)$, let $U(\sigma) = \min_{\rho<\tau\leq\sigma} C(\tau)$. Then the conductance from ρ to a cutset Π is at most*

$$\sum_{\sigma\in\Pi} U(\sigma). \tag{6.1}$$

PROOF. For each $\sigma \in \Pi$, let $\gamma(\sigma)$ be the sequence of conductances on the path from ρ to σ, and let Γ' be a tree consisting of disjoint paths for each $\sigma \in \Pi$, each path having conductances $\gamma(\sigma)$. Γ is a contraction of Γ', so by Rayleigh's monotonicity law [Doyle and Snell (1984)], the conductance to Π in Γ is less than or equal to the conductance of Γ', which is the sum over $\sigma \in \Pi$ of conductances bounded above by $U(\sigma)$. □

PROOF OF THEOREM 2.1. (i) This is almost immediate from Theorem 2.2, which was proved in the previous section. Since Γ has positive capacity in gauge $n^{-1/2}$, that theorem guarantees the almost sure existence of a ray ξ along which the partial sums $S(\sigma) = \sum_{\rho<\tau\leq\sigma} X(\tau)$ satisfy $S(\sigma) \geq 2\log|\sigma|$ when σ is sufficiently large. The total resistance along ξ is then $\sum_{\sigma\in\xi} e^{-S(\sigma)} < \infty$, so the resistance of the entire tree is finite and the RWRE is transient.

(ii) First we calculate a bound for the expected conductance along the path from ρ to σ in terms of $|\sigma|$, then use the Hausdorff measure assumption to bound the net conductance of the tree from ρ to $\partial\Gamma$. In this calculation it is convenient to attach an auxiliary unit resistor, thought of as an edge $\rho'\rho$ comprising the -1 level of Γ. After this addition, the minimal conductance in any edge in the path from ρ' to σ is $e^{m(\sigma)}$, where $m(\sigma) = \min\{S(\tau): \rho \leq \tau \leq \sigma\}$ is nonpositive since $S(\rho) = 0$. Applying the first part of Lemma 3.3 gives

$$\mathbf{P}(m(\sigma) \geq -y) \leq c|\sigma|^{-1/2}\max\{y, 1\}$$

for some constant c. Plug this into the identity

$$\mathbf{E}\, e^{m(\sigma)} = \int_0^1 \mathbf{P}(e^{m(\sigma)} \geq u)\, du,$$

changing variables to $y = \log(1/u)$ to get

$$\begin{aligned}\mathbf{E}\, e^{m(\sigma)} &= \int_0^\infty \mathbf{P}(m(\sigma) \leq -y)e^{-y}\, dy\\ &\leq c|\sigma|^{-1/2}\Big(1 + \int_1^\infty ye^{-y}\, dy\Big)\\ &\leq 2c|\sigma|^{-1/2}.\end{aligned}$$

The Hausdorff measure assumption implies the for any $\varepsilon > 0$ there is a cutset $\Pi(\varepsilon)$ for which $\sum_{\sigma\in\Pi(\varepsilon)} |\sigma|^{-1/2} < \varepsilon$, hence by Lemma 6.1 the expected conductance from ρ' to $\Pi(\varepsilon)$ is at most $2c\varepsilon$. Thus the net conductance from ρ' to $\partial\Gamma$ vanishes almost surely, so the RWRE is almost surely recurrent.

(iii) Set $f(n) = \log|\Gamma_n| + 2\log(n+1)$. The assumptions in (iii) imply that f is increasing and $\sum n^{-3/2} f(n) < \infty$, so Theorem 2.2 implies that with probability one, every ray of Γ has $S(\sigma) < -f(|\sigma|)$ infinitely often. Thus for every $N \geq 1$ there exists almost surely a random cutset Π such that for every $\sigma \in \Pi$, $|\sigma| > N$ and $S(\sigma) < -f(|\sigma|)$. The net conductance from ρ to this Π is at most

$$\begin{aligned} \sum_{\sigma\in\Pi} e^{S(\sigma)} &\leq \sum_{\sigma\in\Pi} e^{-f(|\sigma|)} \\ &\leq \sum_{n=N}^{\infty} |\Gamma_n| e^{-f(n)} \\ &\leq \sum_{n=N}^{\infty} n^{-2}. \end{aligned}$$

Taking N large shows that the net conductance to $\partial\Gamma$ vanishes almost surely. □

7. Reinforced random walk. Reinforced RW is a process introduced by Coppersmith and Diaconis (unpublished) to model a tendency of the random walker to revisit familiar territory. The variant of this process analysed in Pemantle (1988) has an inherent bias toward the root, that is, a positive backward push; here we consider the following "unbiased" variant which fits into the general framework of reinforcement described in Davis (1990), and may be analysed using the tools developed in the previous sections of this paper.

Let Γ be an infinite rooted tree, with (dynamically changing) positive weights $w_n(e)$ for $n \geq 0$ attached to each edge e. At time zero, all weights are set to one: $w_0(e) = 1$ for all e. Let Y_0 be the root of Γ, and for every $n \geq 0$, given $Y_1, \dots, Y_n$ let Y_{n+1} be a randomly chosen vertex adjacent to Y_n, so that each edge e emanating from Y_n has conditional probability proportional to $w_n(e)$ to be the edge connecting Y_n to Y_{n+1}. Each time an edge is traversed *back and forth*, its weight is increased by 1, that is, $w_k(e) - 1$ is the number of "return trips taken on e by time k". Call the resulting process $\{Y_n\}$ an "unbiased reinforced random walk."

THEOREM 7.1. (i) *If Γ has positive capacity in gauge $\phi(n) = n^{-1/2}$, then the resulting reinforced RW is transient, that is,* $\mathbf{P}(Y_n = Y_0$ *infinitely often*$) = 0$.

(ii) *If Γ has zero Hausdorff measure in the same gauge, then the reinforced RW is recurrent, that is,* $\mathbf{P}(Y_n = Y_0$ *infinitely often*$) = 1$.

PROOF. Fix a vertex σ in Γ, of degree d. As explained in Section 3 of Pemantle (1988), for every vertex σ of Γ, the sequence of edges by which the walk leaves σ is equivalent to "Polya's urn" stopped at a random time. Initially the urn contains d balls, one of each color. (The colors correspond to the edges emanating from σ.) Each time the walk leaves σ, a ball is picked at random from the urn, and returned to the urn along with another ball of the same color. (This corresponds to increasing the weight of the relevant edge.)

From Section VII.4 of Feller (1966) we find that the sequence of edges taken from σ is stochastically equivalent to a mixture of sequences of i.i.d. variables, where the mixing measure is uniform over the simplex of probability vectors of length d. A standard method to generate a uniform random vector on the simplex is to pick d independent identically distributed exponential random variables, and normalize them by their sum. This leads to the following *RWRE description of the reinforced RW*:

Assign to each edge e in Γ two exponential random variables $U(\overrightarrow{e})$ and $U(\overleftarrow{e})$, one for each orientation, so that all the assigned variables are i.i.d. These labels are then used to define an environment for a random walk on Γ such that the transition probability from a vertex σ to a neighboring vertex τ is

$$q(\sigma,\tau) = \frac{U(\overrightarrow{\sigma\tau})}{\sum\{U(\overrightarrow{e})\colon \overrightarrow{e} \text{ emanates from } \sigma\}}.$$

Thus the log-ratios $\{X(\sigma)\}_{\sigma\in\Gamma}$ defined in (1.1) are identically distributed, and any subcollection of these variables where no two of the corresponding vertices are siblings, is independent. Clearly the variables $\{X(\sigma)\}$ have mean 0 and finite variance.

The proof of Theorem 2.1(ii) goes over unchanged to prove part (ii) of the present theorem, since in order to apply the "first moment method", Lemma 4.1(i), it suffices that for any ray ξ in Γ, the variables $\{X(\sigma)\}$ for σ on ξ be independent.

To prove part (ii) via the second moment method, Lemma 4.1(ii), consider the percolation process defined by retaining only vertices for which the partial sum from the root of $X(\sigma)$ is positive. It suffices to verify that this percolation is quasi-Bernoulli, that is, it satisfies (5.1); this involves only a minor modification (which we omit) of the proof of Theorem 2.2(i) given in Section 5. □

Acknowledgments. We are indebted to the referee for a remarkably careful reading of the paper. We gratefully acknowledge Russell Lyons and the Institute for Iterated Dining for bringing us together.

REFERENCES

BENJAMINI, I. and PERES, Y. (1992). Random walks on a tree and capacity in the interval. *Ann. Inst. H. Poincaré Probab. Statist.* **28** 557–592.

BENJAMINI, I. and PERES, Y. (1994a). Markov chains indexed by trees. *Ann. Probab.* **22** 219–243.

BENJAMINI, I. and PERES, Y. (1994b). Tree-indexed random walks on groups and first-passage percolation. *Probab. Theory Related Fields* **98** 91–112.

CARLESON, L. (1967). *Selected Problems on Exceptional Sets. Mathematical Studies* **13** Van Nostrand, Princeton, NJ.

DAVIS, B. (1990). Reinforced random walk. *Probab. Theory Related Fields* **84** 203–229.

DERRIDA, B. and SPOHN, H. (1988). Polymers on disordered trees, spin glasses, and traveling waves. *J. Statist. Phys.* **51** 817–841.

DOYLE, P. and SNELL, J. L. (1984). *Random Walks and Electrical Networks*. Math. Assoc. Amer., Washington.

DUBINS, L. and FREEDMAN, D. (1967). Random distribution functions. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* Univ. California Press, Berkeley.

DURRETT, R. (1986). Multidimensional random walks in random environments with subclassical limit behavior. *Comm. Math. Phys.* **104** 87–102.

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

FALCONER, K. J. (1987). Cut-set sums and tree processes. *Proc. Amer. Math. Soc.* **101** 337–346.

FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

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

JOFFE, A. and MONCAYO, A. R. (1973). Random variables, trees, and branching random walks. *Adv. in Math.* **10** 401–416.

KEENER, R. (1992). Limit theorems for random walks conditioned to stay positive. *Ann. Probab.* **20** 801–824.

KOZLOV, M. (1976). On the asymptotic probability of nonextinction for critical branching processes in a random environment. *Theory Probab. Appl.* **21** 791–804.

LAWLER, G. and POLASKI, T. (1992). Harnack inequalities and difference estimates for random walks with infinite range. Preprint.

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

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

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

LYONS, R. and PEMANTLE, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20** 125–136.

MILLAR, P. W. (1976). Sample functions at almost exit time. *Z. Wahrsch. Verw. Gebiete* **34** 91–111.

NAU, D. S. (1983). Pathology on game trees revisited, and an alternative to minimaxing. *Artif. Intell.* **21** 221–244.

NOVIKOV, A. (1983). A martingale approach to problems on first crossing time of nonlinear boundaries. *Proc. Steklov Inst. Math.* **4** 141–163.

PEMANTLE, R. (1988). Phase transition in reinforced random walk and RWRE on trees. *Ann. Probab.* **16** 1229–1241.

PEMANTLE, R. (1993). Critical random walk in random environment on trees of exponential growth. In *Proceedings of the Seminar on Stochastic Processes 1992* (E. Çinlar, K. Chung and M. Sharpe, eds.) 221–240. Birkhäuser, Boston.

PEMANTLE, R. and PERES, Y. (1994). Domination between trees and application to an explosion problem. *Ann. Probab.* **22** 180–194.

RITTER, G. (1981). Growth of random walk conditioned to stay positive. *Ann. Probab.* **9** 699–704.

ROBERTS, G. O. (1991). Asymptotic expansions for Brownian motion hitting times. *Ann. Probab.* **19** 1689–1731.

SIEGMUND, D. (1986). Boundary crossing probabilities and statistical applications. *Ann. Statist.* **14** 361–404.

SINAI, YA. G. (1982) Limit behaviour of one-dimensional random walks in random environment. *Theory Probab. Appl.* **27** 247–258.

SOLOMON, S. (1975). Random walk in a random environment. *Ann. Probab.* **3** 1–31.

WOESS, W. (1986). Transience and volumes of trees. *Arch. Math.* **46** 184–192.

WOODROOFE, M. (1976). A renewal theorem for curved boundaries and moments of first passage times. *Ann Probab.* **4** 67–80.

ZHANG, Y. (1991). A power law for connectedness of some random graphs at the critical point. *Random Structures and Algorithms* **2** 101–119.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
VAN VLECK HALL
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720