CRITICAL RANDOM WALK IN RANDOM ENVIRONMENT ON TREES OF EXPONENTIAL GROWTH

ROBIN PEMANTLE [1] [2]

ABSTRACT:
This paper studies the behavior of RWRE on trees in the critical case left open in previous work. For trees of exponential growth, a random perturbation of the transition probabilities can change a transient random walk into a recurrent one. This is the opposite of what occurs on trees of sub-exponential growth.

1 Introduction

This paper is concerned with the problem of determining whether a random walk in a random environment (RWRE) on an infinite, exponentially growing tree is transient or recurrent. The problem was first studied in [8] as a way of analyzing another process called Reinforced Random Walk, and then in [7] where a more complete solution was obtained. It was shown there that the RWRE is transient when the size of the tree, as measured by the log of the *branching number*, is greater than the *backward push* of the random environment, and recurrent when the log of the branching number is smaller than the backward push. The case of equality was left open. For trees of sub-exponential growth, this critical case was almost completely settled in [9]. The present paper is a companion to [9] in that it attempts to settle the critical case for exponentially growing trees. The results here are less definitive than in the sub-exponential case, in that the sufficient

[1] Research supported in part by a National Science Foundation postdoctoral fellowship and be NSF grant # DMS9103738
[2] Department of Mathematics, University of Wisconsin-Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706

conditions for transience and for recurrence are in terms of capacity and growth respectively; these conditions are not quite complementary, leaving open a critical-within-critical case. A technical assumption on the random environment is also required; examples show that this assumption is often satisfied. On the positive side, it is proved here that a phase boundary occurs in an unusual place, namely when the growth rate of the tree, $\exp(\beta n + o(n))$, has the $o(n)$ term equal to a constant times $n^{1/3}$. Here follows a precise description of the problem.

Let Γ be any infinite, locally finite tree with no leaves (vertices of degree one). Designate a vertex ρ of Γ 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 [7] and the references therein, random environments studied in this paper satisfy the homogeneity condition

$$\text{The variables } X(\sigma) = \ln\left(\frac{q(\sigma', \sigma)}{q(\sigma', \sigma'')}\right) \text{ are i.i.d. for } |\sigma| \geq 2, \tag{1}$$

where $|\sigma|$ denotes the distance from σ to ρ. Here, and throughout, let X denote a random variable with this common distribution.

Before stating the main result, a few definitions and notations are required. Write $\sigma \leq \tau$ if σ is on the path connecting ρ and τ; in this paper, the term "path" always refers to a path without self-intersection. Write $\sigma \wedge \tau$ for the greatest lower bound of σ and τ; pictorially, this is where the paths from ρ to σ and τ diverge. Let $\partial\Gamma$, called the *boundary* of Γ, denote the set of infinite paths beginning at ρ. Let Γ_n denote the set $\{\sigma : |\sigma| = n\}$ of vertices at level n of Γ. Define the *backward push* of the random environment, denoted β or $\beta(X)$, by

$$\beta(X) = -\ln \min_{0 \leq \lambda \leq 1} \mathbf{E} e^{\lambda X}.$$

The size of an infinite tree is best discussed in terms of capacity.

Definition 1 *Let $\phi : \mathbf{Z}^+ \to \mathbf{R}^+$ be a nonincreasing function. Define the*

ϕ-energy of a probability measure μ on the boundary of Γ to be

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

Define the capacity of Γ in gauge ϕ by

$$Cap_\phi(\Gamma) = \left[\inf_\mu I_\phi(\mu)\right],$$

where the infimum is over all probability measures on $\partial\Gamma$., and $Cap_\phi(\Gamma) \neq 0$ if and only if there is some measure of finite energy.

Say that Γ is *spherically symmetric* if there is a *growth function* $f : \mathbf{Z}^+ \to \mathbf{Z}^+$ such that every vertex $\sigma \neq \rho$ has $1+f(|\sigma|)$ neighbors; in other words, the degree of a vertex depends only on its distance from the root. A spherically symmetric tree Γ has positive capacity in gauge ϕ if and only if

$$\sum \phi(n)|\Gamma_n|^{-1} < \infty.$$

Thus positive capacity in gauge $\phi(n) = e^{-kn}$ implies liminf exponential growth rate of at least k. In particular, the supremum of those k for which Γ has positive capacity in gauge $\phi(n) = e^{-kn}$ is the Hausdorff dimension, $dim(\Gamma)$; in the terminology of [5] and [7], $dim(\Gamma)$ is the log of the branching number.

The main result of [7] is that RWRE on Γ is a.s. transient if $dim(\Gamma) > \beta(X)$, and a.s. recurrent if $dim(\Gamma) < \beta(X)$. The use of gauges more general than e^{-kn} allows for finer distinctions of size to be made within the class of trees of the same dimension. In particular, when $\beta = dim(\Gamma) = 0$, it is shown in [9] that positive capacity in gauge $n^{1/2}$ is sufficient and almost necessary for transience of RWRE. Since the case $\beta = 0$ is in some sense a mean-zero perturbation of the deterministic environment of a simple random walk ($X \equiv 0$), and simple random walk is transient if and only if Γ has positive capacity in gauge n^{-1}, this shows that the perturbation makes the walk more transient. By contrast, the main result of this paper is as follows.

Definition 2 *Say that a real random variable X is* top-heavy *if the infimum of $\mathbf{E}e^{\lambda X}$ over $\lambda \in [0,1]$ is achieved at some $\lambda_0 \in (0,1)$ and $\mathbf{E}e^{\gamma X} < \infty$ for some $\gamma > \lambda_0$.*

Theorem 1 *Consider RWRE on a tree* Γ *with* $\beta(X) = dim(\Gamma) > 0$. *If* $X = -\beta$ *with probability one, then RWRE is transient if and only if* Γ *has positive capacity in gauge* $\phi(n) = e^{-n\beta}$. *On the other hand if* X *is nondeterministic, top-heavy, and is either a lattice distribution or has an absolutely continuous component with density bounded above and bounded away from zero in a neighborhood of zero, then*

(*i*) *there exists* $c_1(X)$ *for which the growth bound*

$$|\Gamma_n| \leq e^{\beta n + c_1 n^{1/3}} \text{ for all } n$$

implies that RWRE is recurrent;

(*ii*) *there exists* $c_2(X)$ *such that if* Γ *has positive capacity in gauge* $\phi(n) = e^{-n\beta - c_2 n^{1/3}}$ *then RWRE is transient.*

Remark: The requirement that X be top-heavy is enigmatic, but not overly restrictive. For example, it is satisfied by normal random variables with mean $-c$ and variance V whenever $c < 2V$. In the case where X takes only the values ± 1, with $\mathbf{P}(X = 1) = p < 1/2$, it is top-heavy if and only if $(1-p)/p < e^2$.

The remainder of this section outlines the the proof of this theorem and serves as a guide to the remaining sections. Theorem 1 is proved in the following three steps. First, in Section 2, a correspondence connection between random walks and electrical networks [2] reduces the problem to one of determining whether a random electrical network is transient or recurrent almost surely. After this reduction, the technical condition of top-heaviness comes in: top-heaviness implies that finite resistance, when achieved, will be due to a single random infinite path of finite resistance; searching for a single path with this property is easier than searching for some large collection of paths with a weaker property. Next, large deviation estimates are needed for the probability of an unusually small resistance along a fixed path of length n (Lemma 6 and Corollary 7). These are applied via a simple first-moment calculation to obtain Lemma 8, which is an upper bound tending to zero on the probability that any of the $|\Gamma_n|$ chains of resistances of length n stays small. Incidentally, this is where the "extra" factor of $e^{cn^{1/3}}$ comes in. Estimates with this same factor have been obtained by Kesten with much greater accuracy for branching Brownian motion [4]. Ours are a discrete analogue of Kesten's in the sense that continuous-time branching has been replaced by Γ-indexed branching; this analogy is explained more fully in [1] and [9]. Part (i) of Theorem 1 follows from this upper bound by computing the expected truncated conductance. The last step, which is needed only for the proof of part (ii), is a second-moment technique (Lemma 3) developed in [6] and [9] for proving the almost sure existence of an infinite path of finite resistance based on the two-dimensional marginals for finite paths (i.e. the probabilities for two paths of length n both to have large conductance if the paths share the first k resistors).

2 Reductions

Begin with the reduction of the recurrence/transience problem to an electrical problem. As is well known, transience of a reversible Markov chain is equivalent to finite resistance of the associated resistor network on the same graph, where the transition probabilities from any vertex are proportional to the conductances (reciprocal resistances) of the edges incident to that vertex; see for example [2]. For a random environment satisfying (1), the resistances in the associated random electrical network are easily seen to be given by

$$\text{Resistance along } \overline{\sigma'\sigma} = e^{-S(\sigma)}$$

where

$$S(\sigma) = \prod_{\rho < \tau \leq \sigma} X(\sigma).$$

Here, the values of $X(\sigma)$ for $|\sigma| \leq 1$ are assigned to make this relation hold for all $\sigma \neq \rho$, while the values for $|\sigma| \geq 2$ are i.i.d. by (1). Since finiteness of the total resistance is not affected by changing finitely many resistances, we alter the $X(\sigma)$ for $|\sigma| \leq 1$ so that the entire collection is i.i.d.

A sufficient condition for transience is the existence of an infinite path $\rho, \sigma_1, \sigma_2, \ldots$ along which $\sum e^{-S(\sigma)} < \infty$. Conversely, let

$$U(\sigma) = \min_{\rho < \tau \leq \sigma} e^{S(\sigma)}.$$

A useful sufficient condition for recurrence is given by the following lemma.

Lemma 2 *Let Γ be any tree with conductances $C(\sigma)$, and let Π be any cutset, i.e. any minimal set among those intersecting every infinite path from ρ. Then the conductance from ρ to Π is at most*

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

Consequently, if the conductances are random with $\sum_{|\sigma|=n} U(\sigma) \to 0$ in probability, then the random walk is recurrent with probability one.

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, 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 $U(\sigma)$. Putting $\Pi = \Gamma_n$ shows that the conductance from ρ to infinity is bounded above by $\liminf_n \sum_{|\sigma|=n} U(\sigma)$, proving the lemma. $\square$

In the next section estimates will be given on $\mathbf{P}(\sigma \in W)$ and $\mathbf{P}(\sigma, \tau \in W)$, where for fixed constants c and L, W is the set of vertices σ such that for every $\tau \leq \sigma$ with $|\tau| > L$,

$$1/10 \leq S(\tau)/c|\tau|^{1/3} \leq 1.$$

These estimates are then plugged into the following result of Lyons (see [9, Theorem 4.1]).

Lemma 3 *Let Γ be an infinite, locally finite tree without leaves, let $X(\sigma)$ be i.i.d. random variables indexed by the vertices of Γ and let B_n be a subset of $\mathbf{R}^n$ for each n. Let W be the set of vertices $\sigma \in \Gamma$ such that for every $\tau \leq \sigma$, the sequence $(X(\rho), \ldots, X(\tau))$ along the path from ρ to τ is in the set $B_{|\tau|}$. Suppose there is a positive, nonincreasing function $g : \mathbf{Z}^+ \to \mathbf{R}$ such that for any two vertices $\sigma, \tau \in \Gamma_n$ with $|\sigma \wedge \tau| = k$,*

$$\mathbf{P}(\sigma, \tau \in W) \leq \frac{\mathbf{P}(\sigma \in W)^2}{g(k)}. \tag{2}$$

Then the probability of W containing an infinite path is at least $Cap_g(\Gamma)$. $\square$

Remark: Of course any infinite path in W has $\sum e^{-S(\sigma)} < \infty$, implying transience. The reason that one looks for a path along which $S(\tau)$ is bounded above as well as below is so as to be able to apply this lemma, which is really a jazzed up second moment bound. It is important to find a random subset of Γ_n whose cardinality has a second moment not too much larger than the square of its first moment; then the LHS of (2) will not be too large.

3 Large deviation estimates

The first series of estimates concern the probabilities $\mathbf{P}(\sigma \in W)$ of the previous lemma. The routine proofs of the first two propositions in the series are omitted.

Proposition 4 *Let $\mathbf{P}_x$ be the law of a standard one-dimensional Brownian motion started at x. Let $-1 < c_1 < c_2 < 1$ and $-1 < c_3 < c_4 < 1$ be real constants. Then there exist positive constants K_1 and K_2 such that for any $L > 0$, the following two inequalities hold.*

$$\sup_x \mathbf{P}_x(|B_t| \leq 1 \text{ for all } t \leq L) \quad \leq \quad K_1 e^{-\frac{\pi^2}{8}L} \quad ;$$

$$\inf_{c_1 \leq x \leq c_2} \mathbf{P}_x(|B_t| \leq 1 \text{ for all } t \leq L \text{ and } c_3 \leq |B_L| \leq c_4) \quad \geq \quad K_2 e^{-\frac{\pi^2}{8}L}.$$

Proposition 5 *Let $\{X_n\}$ be a sequence of i.i.d. random variables with mean zero and variance $V < \infty$. Let $S_n = \sum_{i=1}^n X_i$ and for $0 < t < 1$, interpolate polygonally by letting $S_{n+t} = S_n + tX_{n+1}$. Then*

$$\limsup_{L\to\infty} \limsup_{n\to\infty} \sup_{|x|\leq 1} L^{-1} \ln \left(\mathbf{P}(|(Vn)^{-1/2} S_{nt} + x| \leq 1 \text{ for all } t \leq L)\right)$$

$$\leq -\frac{\pi^2}{8}. \tag{3}$$

If $-1 < c_1 < c_2 < 1$ and $-1 < c_3 < c_4 < 1$ are real constants, then

$$\liminf_{L\to\infty} \liminf_{n\to\infty} \inf_{c_1\leq x\leq c_2} L^{-1} \ln \mathbf{P}[A(L,n,x)]$$

$$\geq \quad -\frac{\pi^2}{8}, \tag{4}$$

where $A(L,n,x)$ *is the event*

$$\left\{|(Vn)^{-1/2}S_{nt}+x|\le 1 \textit{ for all } t\le L, \textit{ and } c_3\le S_{nL}+x\le c_4\right\}.$$

Lemma 6 *Suppose* $f,g:\mathbf{Z}^+\to\mathbf{R}$ *satisfy* $f>g$ *and* $\lim_{t\to\infty}f(t)-g(t)=\infty$*, and assume the following flatness hypothesis:*

$$\sup_L\lim_{t\to\infty}\frac{\sup_{0\le s\le L(f(t)-g(t))^2}\max(|f(t+s)-f(t)|,|g(t+s)-g(t)|)}{f(t)-g(t)} = 0.$$

Let $S_n=\sum_{i=1}^n X_i$ *be a random walk with* $\mathbf{E}X_1=0$ *and* $\mathbf{E}X_1^2=V<\infty$ *and suppose that for each* n*,* $\mathbf{P}(g(k)<S_k<f(k) \textit{ for all } k\le n)>0$*. Then*

$$\lim_{n\to\infty}\left(\sum_{k=1}^n(f(k)-g(k))^{-2}\right)^{-1}\ln\left(\mathbf{P}(g(k)<S_k<f(k) \textit{ for all } k\le n)\right)$$

$$= -\frac{\pi^2}{8}V. \qquad (5)$$

Corollary 7 *Let* Y_n *be i.i.d. mean zero random variables with partial sums* $T_n=\sum_{i=1}^n Y_i$ *and suppose that* $\mathbf{E}Y_1^2=V<\infty$*. Then for any real* $c_1<c_2$*,*

$$\lim_{n\to\infty}\frac{\ln\mathbf{P}\left(c_1k^{1/3}\le T_k\le c_2k^{1/3} \textit{ for all } k\le n\right)}{n^{1/3}}=-\frac{\pi^2}{8}\frac{3V}{(c_2-c_1)^2}$$

unless the probability is eventually zero.

Proofs: The first proposition is a standard eigenvalue estimate – see for example [3, p. 259]. The second follows from the first and the invariance principle by piecing together segments of path of length L; the two-page proof is entirely routine and is omitted.

To establish the lemma, prove first that the limsup is at most $-\frac{\pi^2}{8}V$. Extend f, g and S to $\mathbf{R}^+$ by linear interpolation with $f(0)=g(0)=S_0=0$; for any integer n, the event $\{g(k)\le S_k\le f(k):k=1,\ldots,n\}$ is the same when k is replaced by a real parameter t running from 0 to n. Fix $L>0$. For any positive integer m_0 inductively define $m_{k+1}=\lceil m_k+(1/4)V^{-1}L(f(m_k)-g(m_k))^2\rceil$. Let $M(n)$ denote $\max\{k:m_k<n\}$.

<u>Claim</u>: $\lim_{n\to\infty}M(n)^{-1}\sum_{k=1}^n 4VL^{-1}(f(k)-g(k))^{-2}=1$.

To prove this, first choose $\epsilon\in(0,1/8)$ and j large enough so that $t\ge m_j$ implies

$$\sup_s\frac{\max(|f(t+s)-f(t)|,|g(t+s)-g(t)|)}{f(t)-g(t)}<\epsilon, \qquad (6)$$

where the supremum is over s for which $0 \leq s \leq (1/4)V^{-1}L(f(t)-g(t))^2$. Summing the identity $\sum_{i=m_k}^{m_{k+1}}(m_{k+1}-m_k)^{-1} = 1$ over $j \leq k \leq M(n)+1$ yields

$$\begin{aligned} M(n) &\leq j+1+\sum_{i=m_0}^{n} 4VL^{-1}(f(m_{M(i)})-g(m_{M(i)}))^{-2} \\ &\leq j+1+\sum_{i=m_0}^{n} 4VL^{-1}(f(i)-g(i))^{-2}(1-2\epsilon)^{-2}. \end{aligned}$$

Similarly,

$$M(n) \geq j+\sum_{i=m_0}^{n} 4VL^{-1}(f(i)-g(i))^{-2}(1+2\epsilon)^{-2}.$$

Letting $n \to \infty$ proves that the limsup and liminf of

$$M(n)^{-1}\sum_{k=1}^{n} 4VL^{-1}(f(k)-g(k))^{-2}$$

are between $(1+2\epsilon)^{-1}$ and $(1-2\epsilon)^{-1}$. Letting $\epsilon \to 0$ proves the claim.

Continuing the proof of the lemma, pick $\epsilon > 0$ and N large enough so that for our (still fixed) value of L and $n \geq N$, the sup in (3) is at most $-\frac{\pi^2}{8}+\epsilon$. Pick m_0 large enough so that $t \geq m_0$ implies firstly (6) and secondly $f(t)-g(t) > (4VN)^{1/2}$. Let $y_k = (f(m_k)+g(m_k))/2$ and $d_k = (f(k)-g(k))/2$, so $m_{k+1}-m_k = \lceil LV^{-1}d_k^2 \rceil$. Let R_k be the rectangle

$$\{(x,y) : m_k \leq x \leq m_{k+1} \text{ and } y_k-(1+2\epsilon)d_k \leq y \leq y_k+(1+2\epsilon)d_k\}.$$

Then the union of the R_k covers the graphs of f and g after m_0, i.e. the region $\{(x,y) : x \geq m_0, g(x) \leq y \leq f(x)\}$. Write S_i' for a random walk independent of S_i and identically distributed. Using the Markov property of S_i, the stationarity of its increments, and then rescaling each R_k to have height one, gives

$$\begin{aligned} &\mathbf{P}(f(k) \leq S_k \leq g(k) : 1 \leq k \leq n) \\ \leq\ &\mathbf{P}((i,S_i) \in \bigcup R_k : m_0 \leq i \leq n) \\ \leq\ &\prod_{k=0}^{M(n)-1} \mathbf{P}\left((i,S_i) \in R_{k-1} : m_k \leq i \leq m_{k+1} \,|\, S_{m_k}\right) \end{aligned}$$

$$= \prod_{k=1}^{M(n)-1} \mathbf{P}(|S_{m_k} + S'_i - y_k| \leq (1+2\epsilon)d_k : i = 1, \ldots, m_{k+1} - m_k)$$

$$= \prod_{k=1}^{M(n)-1} \mathbf{P}\Big(|(1+2\epsilon)^{-1} d_k^{-1}[(S_{m_k} - y_k) + S'_{V^{-1}(1+2\epsilon)^2 d_k^2 t}]| \leq 1 \text{ for all } 0 \leq t \leq L(1+2\epsilon)^{-2}\Big)$$

$$\leq \sup_{w \geq N} \sup_{0 \leq x \leq 1} \left[\mathbf{P}(\frac{|x + S_{wt}|}{(Vw)^{1/2}} \leq 1 : 0 \leq t \leq L(1+2\epsilon)^{-2})\right]^{M(n)-1}$$

where the rescaling factor $w = (1+2\epsilon)^2 V^{-1} d_k^2$ is at least N by the previous choice of m_0. Taking the log, dividing by $L(1+2\epsilon)^{-2}$ and applying (3) gives

$$L^{-1}(1+2\epsilon)^2 \ln \mathbf{P}(f(k) \leq S_k \leq g(k) : k \leq n) \leq (M(n)-1)(-\frac{\pi^2}{8} + \epsilon)$$

by choice of N. Plugging in the asymptotic value of $M(n)$ from the claim above gives that for sufficiently large n,

$$\begin{aligned} & \ln \mathbf{P}(f(k) \leq S_k \leq g(k) : k \leq n) \\ \leq \quad & -\frac{\pi^2}{8}(M(n)-1)L(1+2\epsilon)^{-2} \\ \leq \quad & \epsilon + 4(-\frac{\pi^2}{8} + \epsilon)V(1+2\epsilon)^{-1} \sum_{k=1}^{n} (f(k) - g(k))^{-2} \end{aligned}$$

for large n. Letting $\epsilon \to 0$ proves that the limsup in (5) is at most $-\frac{\pi^2}{8}V$.

Proving that the liminf is at least $-\frac{\pi^2}{8}V$ is almost identical. Fixing $L, \epsilon > 0$, choose N and m_0 as before and this time define R_k to lie between f and g instead of covering them:

$$R_k = \{(x,y) : m_k \leq x \leq m_{k+1} \text{ and } y_k - (1-2\epsilon)d_k \leq y \leq y_k + (1-2\epsilon)d_k\}.$$

Let G_k be the event that $(i, S_i) \in R_k$ for $m_k \leq i \leq m_{k+1}$ and that $y_k - \frac{1}{4}d_k \leq S_{m_{k+1}} \leq y_k + \frac{1}{4}d_k$. Since $m_{M(n)+1} \geq n$, the probability we are trying to bound from below is at least $\mathbf{P}(g(k) \leq S_k \leq f(k) : k = 1, \ldots, m_{M(n)+1})$, which may be written as

$$\mathbf{P}(g(k) \leq S_k \leq f(k) : k = 1, \ldots, m_0) \prod_{i=0}^{M(n)} \mathbf{P}(G_k \,|\, S_{m_k}, G_{k-1}).$$

When G_{k-1} occurs, the value of S_k is certainly between $y_k - \frac{1}{2}d_k$ and $y_k + \frac{1}{2}d_k$, so

$$\mathbf{P}(G_k \mid S_{m_k}, G_{k-1}) \geq \inf_{y_k - \frac{1}{2}d_k \leq x \leq y_k + \frac{1}{2}d_k} \mathbf{P}(G_k \mid S_k = x).$$

Now rescaling each rectangle, applying (4) with $c_1 = -(1/2)(1-2\epsilon)^{-2}, c_2 = (1/2)(1-2\epsilon)^{-2}, c_3 = -1/4, c_4 = 1/4$, and taking limits establishes that the liminf in (5) is at least $-\frac{\pi^2}{8}V$, finishing the proof of the lemma.

Finally, the corollary is proved by letting $f(n) = c_2 n^{1/3}, g(n) = c_1 n^{1/3}$, and verifying the flatness hypothesis; summing $[f(k) - g(k)]^{-2}$ from 1 to n gives $(3 + o(1))(c_2 - c_1)^{-2} n^{1/3}$ and the desired conclusion follows. □

The next step is to apply these random walk estimates to prove a tree-indexed version of Kesten's result on branching Brownian motion staying above zero.

Lemma 8 *Let $\Gamma, X_\sigma, S_\sigma$ and β be as above. Assume that X is top-heavy and let λ_0 be the value of λ minimizing $\mathbf{E}e^{\lambda X}$ (which must exist and be strictly less than one, according to the definition of the term top-heavy). There exists a positive real number c, independent of Γ, such that if $|\Gamma_n| \leq e^{cn^{1/3} + n\beta}$ for all n, then*

$$\mathbf{P}(\max_{\sigma \in \Gamma_n} \min_{\tau \leq \sigma} S(\tau) \geq -2c(1-\lambda_0)^{-1} n^{1/3}) \to 0 \tag{7}$$

as $n \to \infty$. In other words with high probability, for each sufficiently large n, no path from ρ of length n stays above $-2c(1-\lambda_0)^{-1}n^{1/3}$.

To establish this, first record some elementary facts about large deviations.

Proposition 9 *Let X_n be i.i.d. and $S_k = \sum_{i=1}^k X_i$. Let $\beta(X)$ be the backwards push, and let $\lambda_0(X)$ be the $\lambda \in [0,1]$ minimizing $\mathbf{E}e^{\lambda X}$. Then the following three inequalities hold.*

(*i*) *For any real u,*

$$\mathbf{P}(S_n \geq u) \leq e^{-\beta n - \lambda_0 u};$$

(*ii*) *For any real y,*

$$\mathbf{E}e^{S_n} I(S_n \leq y) \leq (1-\lambda_0)^{-1} e^{(1-\lambda_0)y - \beta n}.$$

(*iii*) *For any real* y,

$$\mathbf{E}e^{S_n \wedge y} \leq [1 + (1 - \lambda_0)^{-1}] e^{(1-\lambda_0)y - \beta n}.$$

Proof: The first claim is just Markov's inequality:

$$\mathbf{P}(\mathcal{S}_n \geq u) \leq e^{-\lambda_0 u} \mathbf{E} e^{-\lambda_0 S_n}.$$

For the second claim, integrate the first by parts:

$$\begin{aligned}
\mathbf{E}(e^{S_n} I(S_n \leq y)) &= \int_{-\infty}^{y} e^u \mathbf{P}(S_n \in du) \\
&= \int_{-\infty}^{y} e^u \mathbf{P}(u \leq S_n \leq y) du \\
&\leq \int_{-\infty}^{y} e^u \mathbf{P}(u \leq S_n) du \\
&\leq \int_{-\infty}^{y} e^u e^{-\lambda_0 u - \beta n} du \\
&= (1 - \lambda_0)^{-1} e^{(1-\lambda_0)y - \beta n}.
\end{aligned}$$

Finally, the third claim follows from the first two, using

$$\mathbf{E}e^{S_n \wedge y} = \mathbf{E}e^{S_n} I(S_n \leq y) + e^y \mathbf{P}(S_n > y).$$

It should be remarked that more careful estimates give an extra factor of $(1 + o(1))(2\pi n \mathbf{E}Y^2)^{-1/2}$ in the RHS of each inequality which is then asymptotically sharp. $\square$

Next, plug this into a first moment calculation to establish:

Proposition 10 *Let* $\Gamma, X_\sigma, S_\sigma$ *and* β *be as above. Suppose that* $|\Gamma_n| \leq e^{cn^{1/3} + n\beta}$ *for some* $c > 0$. *Then for any* $\epsilon > 0$,

$$\mathbf{P}(S(\sigma) \geq (1 + \epsilon) c \lambda_0^{-1} n^{1/3} \textit{ for some } \sigma \textit{ with } |\sigma| \leq n) \to 0$$

as $n \to \infty$.

Proof: For each fixed L it is clear that

$$\mathbf{P}(S(\sigma) \geq (1 + \epsilon) c \lambda_0^{-1} n^{1/3} \text{ for some } \sigma \text{ with } |\sigma| \leq L) \tag{8}$$

goes to zero as $n \to \infty$. On the other hand,

$$\begin{aligned}
& \mathbf{P}(S(\sigma) \geq (1+\epsilon)c\lambda_0^{-1}n^{1/3} \text{ for some } \sigma \text{ with } n \geq |\sigma| > L) \\
\leq \; & \mathbf{P}(S(\sigma) \geq (1+\epsilon)c\lambda_0^{-1}|\sigma|^{1/3} \text{ for some } \sigma \text{ with } n \geq |\sigma| > L) \\
\leq \; & \sum_{m>L} \mathbf{P}(S(\sigma) \geq (1+\epsilon)c\lambda_0^{-1}m^{1/3} \text{ for some } \sigma \in \Gamma_m).
\end{aligned}$$

For $\sigma \in \Gamma_m$, Proposition 9 part (i) implies

$$\mathbf{P}(S(\sigma) \geq (1+\epsilon)c\lambda_0^{-1}m^{1/3}) \leq e^{-\lambda_0(1+\epsilon)c\lambda_0^{-1}m^{1/3}-\beta m}.$$

Multiplying by $|\Gamma_m|$ gives

$$\mathbf{P}(\mathcal{S}(\sigma) \geq (1+\epsilon)c\lambda_0^{-1}m^{1/3} \text{ for some } \sigma \in \Gamma_m) \leq e^{-\epsilon c n^{1/3}}.$$

This is summable in m, so the sum over $m > L$ goes to zero as $L \to \infty$, which together with (8) proves the proposition. □

Proof of Lemma 8: Let μ be the common distribution of the $X(\sigma)$ and let $Y_1, \ldots, Y_n$ be i.i.d. random variables whose law μ' satisfies

$$\frac{d\mu'}{d\mu}(x) = e^{\lambda_0 x}/\mathbf{E}e^{\lambda_0 X} = e^{\lambda_0 x + \beta(\mu)}. \tag{9}$$

Informally, μ' is μ tilted in the large deviation sense so as to have mean zero. The assumption that μ is top-heavy by definition implies that $\mathbf{E}e^{\lambda X} < \infty$ for λ in some neighborhood of λ_0, hence $\mathbf{E}e^{\lambda Y_1} < \infty$ for λ in some neighborhood of zero, and in particular $V \stackrel{def}{=} \mathbf{E}Y_1^2 < \infty$.

Choose a positive real c for which

$$c + 2\lambda_0(1-\lambda_0)^{-1}c - \frac{\pi^2}{8}\frac{3V}{(2c\lambda_0^{-1} + 2c(1-\lambda_0)^{-1})^2} < 0. \tag{10}$$

Let

$$A_n = \left\{ \max_{\sigma \in \Gamma_n} \min_{\tau \leq \sigma} S(\tau) \geq -2c(1-\lambda_0)^{-1}n^{1/3} \right\}$$

be the event in (7). Let G_n be the event

$$\bigcup_{|\sigma| \leq n} \{S(\sigma) \geq 2c\lambda_0^{-1}n^{1/3}\}$$

and let $H_n = A_n \setminus G_n$. Proposition 10 shows that $\mathbf{P}(G_n) \to 0$ so to show that $\mathbf{P}(A_n) \to 0$ it suffices to show that $\mathbf{P}(H_n) \to 0$.

To see this, fix $\sigma \in \Gamma_n$ and write $\mathbf{P}(H_n) \le |\Gamma_n| Q_n$ where

$$Q_n = \mathbf{P}(2c\lambda_0^{-1} n^{1/3} \ge S(\tau) \ge -2c(1-\lambda_0)^{-1} n^{1/3} \text{ for all } \tau \le \sigma).$$

Let ν be the law in $\mathbf{R}^n$ of the sequence $(S(\sigma_1), \ldots, S(\sigma_n))$, where $\sigma_1, \ldots, \sigma_n$ is the path from $\rho = \sigma_0$ to $\sigma = \sigma_n$; of course ν is just the law in $\mathbf{R}_n$ of a random walk whose steps have law μ. Recalling the tilted variables Y_n, write $T_n = \sum_{i=1}^n Y_i$. Let ν' denote the law in $\mathbf{R}_n$ of $(T_1, \ldots, T_n)$ and observe that

$$\frac{d\nu'}{d\nu}(s_1, \ldots, s_n) = e^{\lambda_0 s_n} / \mathbf{E} e^{\lambda_0 S_n} = e^{\lambda_0 s_n + n\beta}.$$

Use this to get an upper bound on Q_n as follows.

$$\begin{aligned}
& Q_n \\
= & \int I(2c\lambda_0^{-1} n^{1/3} \ge s_i \ge -2c(1-\lambda_0)^{-1} n^{1/3} \text{ for all } i \le n) \\
& d\nu(s_1, \ldots, s_n) \\
= & \int I(2c\lambda_0^{-1} n^{1/3} \ge s_i \ge -2c(1-\lambda_0)^{-1} n^{1/3} \text{ for all } i \le n) \\
& \qquad \frac{d\nu}{d\nu'}(s_1, \ldots, s_n) d\nu'(s_1, \ldots, s_n) \\
\le & \int I(2c\lambda_0^{-1} n^{1/3} \ge s_i \ge -2c(1-\lambda_0)^{-1} n^{1/3} \text{ for all } i \le n) \\
& \qquad \left[\sup_{s_n \ge -2c(1-\lambda_0)^{-1} n^{1/3}} \frac{d\nu}{d\nu'}(s_1, \ldots, s_n) \right] d\nu(s_1, \ldots, s_n) \\
= & \int I(2c\lambda_0^{-1} n^{1/3} \ge s_i \ge -2c(1-\lambda_0)^{-1} n^{1/3} \text{ for all } i \le n) \\
& \qquad \qquad \exp(2c\lambda_0(1-\lambda_0)^{-1} n^{1/3} - n\beta) \, d\nu(s_1, \ldots, s_n) \\
\le & \exp(2c\lambda_0(1-\lambda_0)^{-1} n^{1/3} - n\beta) \\
& \qquad \mathbf{P}(2c\lambda_0^{-1} n^{1/3} \ge T_i \ge -2c(1-\lambda_0)^{-1} \text{ for all } i \le n) \\
\le & \exp\left(-n\beta + n^{1/3} [\frac{2c\lambda_0}{1-\lambda_0} - \frac{\pi^2}{8} \frac{3V}{(2c\lambda_0^{-1} + 2c(1-\lambda_0)^{-1})^2} + o(1)] \right)
\end{aligned}$$

by Corollary 7. Thus $\mathbf{P}(H_n) \le |\Gamma_n| Q_n \le \exp(cn^{1/3} + n\beta) Q_n \le$

$$\exp\left(n^{1/3} \left[c + 2c\lambda_0(1-\lambda_0)^{-1} - \frac{\pi^2}{8} \frac{3V}{(2c\lambda_0^{-1} + 2c(1-\lambda_0)^{-1})^2} + o(1) \right] \right).$$

By choice of c, this is $\exp((K+o(1))n^{1/3})$ for some $K<0$, so $\mathbf{P}(H_n)\to 0$, proving the lemma. □

4 Proof of the main theorem

The case where $X\equiv -\beta$ is done in [5]. For part (i) of the nondegenerate case, use Lemma 2, showing that $\sum_{|\sigma|=n} U(\sigma)$ goes to zero in probability by computing a truncated expectation. Let c_1 be the constant c from Lemma 8 and let $U_n = \sum_{|\sigma|=n} U(\sigma)$. Let G_n be the event that $\max_{|\sigma|=n} U(\sigma) \geq \exp(-2c_1(1-\lambda_0)^{-1}n^{1/3})$. Then

$$
\begin{aligned}
P(U_n>\epsilon) &\leq \mathbf{P}(G_n)+\mathbf{P}(U_n>\epsilon \text{ and } G_n^c)\\
&\leq \mathbf{P}(G_n)+\epsilon^{-1}\mathbf{E}U_nI(G_n^c).
\end{aligned}
$$

Lemma 8 showed that $\mathbf{P}(G_n)\to 0$, so it remains to show that for any ϵ, $\mathbf{E}(U_nI(G_n^c))\to 0$.

Observe that $U_\sigma I(G_n^c) \leq \exp(S_\sigma \wedge -2c(1-\lambda_0)^{-1}n^{1/3})$. Hence for $\sigma\in\Gamma_n$,

$$
\begin{aligned}
&\mathbf{E}U_nI(G_n^c)\\
=\ & |\Gamma_n|\mathbf{E}U_\sigma I(G_n^c)\\
\leq\ & |\Gamma_n|\mathbf{E}\exp(S_\sigma\wedge -2c_1(1-\lambda_0)^{-1}n^{1/3})\\
\leq\ & |\Gamma|[1+(1-\lambda_0)^{-1}]\exp((1-\lambda_0)(-2c_1(1-\lambda_0)^{-1}n^{1/3})-n\beta)\\
& \text{by Proposition 9, part } (iii),\\
\leq\ & [1+(1-\lambda_0)^{-1}]\exp(-c_1n^{1/3})
\end{aligned}
$$

by the assumption on $|\Gamma_n|$. This goes to zero, thus $U_n\to 0$ in probability, proving part (i) of Theorem 1.

Part (ii) is proved by exhibiting an infinite path along which the resistances e^{-S_σ} are summable. In fact the proof finds an infinite path along which $S(\sigma)/c|\sigma|^{1/3}$ is bounded above and below.

Pick any $c, L, \epsilon > 0$ and any $c_2 = K + 2c\lambda_0 + (\frac{\pi^2}{8})\frac{3V}{(.9c)^2}$, where M shall be chosen later. Define W to be the random subset of vertices σ of Γ with the property that for every $\tau \leq \sigma$ with $|\tau| > L$,

$$\frac{c|\tau|^{1/3}}{10} \leq S_\tau \leq c|\tau|^{1/3},$$

where L is large enough so that W intersects each Γ_n with positive probability. To prove the theorem, it suffices to show that W is infinite with positive probability; this follows from Lemma 3 and the hypothesis of the theorem, provided that

$$\sup_n a(n,k)/a(n,n)^2 \leq e^{c_2 k^{1/3} + k\beta} \tag{11}$$

for all but finitely many k, where $a(n,k) = \mathbf{P}(\rho \leftrightarrow \sigma, \tau)$ for vertices $\sigma, \tau \in \Gamma_n$ with $|\sigma \wedge \tau| = k$.

To establish (11), begin with

$$a(n,k)/a(n,n)^2 = \mathbf{P}(\sigma \wedge \tau \in W)^{-1} \frac{\mathbf{P}(\sigma \in W \,|\, \tau \in W)}{\mathbf{P}(\sigma \in W \,|\, \sigma \wedge \tau \in W)}. \tag{12}$$

Recall the tilted random variables Y_n and T_n whose law μ' is defined by (9).

Fix any $k \leq n$ and $\sigma, \tau \in \Gamma_n$ with $|\sigma \wedge \tau| = k$. Let $C(a,b)$ denote the set of sequences $(s_a, \ldots, s_b) \in \mathbf{R}^{b-a+1}$ for which $cs_j^{1/3}/10 \leq s_j \leq cs_j^{1/3}$ for all $j \in [a,b]$. Let ν and ν' respectively denote the law of $(S_1, \ldots, S_k)$ and $(T_1, \ldots, T_k)$ and let ν_y and ν'_y denote the laws of $(S_k, \ldots, S_n)$ and $(T_k, \ldots, T_n)$ conditioned respectively on $S_k = y$ and $T_k = y$. Write the first factor on the RHS of (12) as

$$\left[\int I(\alpha \in C(1,k))\, d\nu(\alpha)\right]^{-1}.$$

Changing the integrating measure to ν' yields

$$\begin{aligned}
&\left[\int I(\alpha \in C(1,k))\, d\nu'(\alpha)\, \frac{d\nu(\alpha)}{d\nu'(\alpha)}\right]^{-1} \\
\leq\quad & \sup_{\alpha \in C(1,k)} \frac{d\nu'(\alpha)}{d\nu(\alpha)} \left[\int I(\alpha \in C(1,k))\, d\nu'(\alpha)\right]^{-1} \\
=\quad & e^{\lambda_0 c k^{1/3} + k\beta} (\nu'(C(1,k)))^{-1} \\
=\quad & \exp\left[\lambda_0 c k^{1/3} + k\beta + k^{1/3}(\frac{\pi^2}{8}\frac{3V}{(.9c)^2} + o(1))\right] \qquad (13)
\end{aligned}$$

by Corollary 7, since ν' is the law of a mean zero, finite variance random walk, and it has been assumed that $\nu'(C(1,k))$ never vanishes.

For the second factor on the RHS of (12), let μ_1 be the law of $S(\sigma \wedge \tau)$ conditional on $\tau \in W$ and let μ_2 be the law of $S(\sigma \wedge \tau)$ conditional on $\sigma \wedge \tau \in W$. The second factor is then

$$\frac{\int \left[\int I(\alpha \in C(k,n))\, d\nu_y(\alpha)\right]\, d\mu_1(y)}{\int \left[\int I(\alpha \in C(k,n))\, d\nu_y(\alpha)\right]\, d\mu_2(y)}\,. \tag{14}$$

Changing the integrating measure again, this becomes

$$\begin{aligned} & \frac{\int\int I(\alpha \in C(k,n))\, \frac{d\nu}{d\nu'}(\alpha)\, d\nu'_y(\alpha) d\mu_1(y)}{\int\int I(\alpha \in C(k,n))\, \frac{d\nu}{d\nu'}(\alpha)\, d\nu'_y(\alpha) d\mu_2(y)} \\ \leq\ & \frac{\int I(\alpha \in C(k,n)) d\nu'_y(\alpha) d\mu_1(y)}{\int I(\alpha \in C(k,n)) d\nu'_y(\alpha) d\mu_2(y)}\ \frac{\sup_{\alpha \in C(k,n)} \frac{d\nu}{d\nu'}(\alpha)}{\inf_{\alpha \in C(k,n)} \frac{d\nu}{d\nu'}(\alpha)} \\ \leq\ & \frac{\sup_y \nu'_y(C(k,n))}{\inf_y \nu'_y(C(k,n))}\ \frac{\sup_{\alpha \in C(k,n)} \frac{d\nu}{d\nu'}(\alpha)}{\inf_{\alpha \in C(k,n)} \frac{d\nu}{d\nu'}(\alpha)} \\ =\ & \frac{\sup_y \nu'_y(C(k,n))}{\inf_y \nu'_y(C(k,n))}\ e^{\lambda_0 c k^{1/3}}. \end{aligned}$$

The argument is then finished by establishing

$$\frac{\sup_y \nu'_y(C(k,n))}{\inf_y \nu'_y(C(k,n))} \leq e^{M k^{1/3}} \tag{15}$$

for some $M > 0$, since then multiplying inequalities (13) and (15) bounds $a(n,k)/a(n,n)^2$ from above by $\exp\left[2\lambda_0 c k^{1/3} + k\beta + k^{1/3}(\frac{\pi^2}{8}\frac{3V}{(.9c)^2} + M)\right]$, which is at most $e^{c_2 k^{1/3} + k\beta}$ by choice of c_2, yielding (11).

It remains to establish (15). An argument is given for the case where the distribution of the $X(\sigma)$'s has an absolutely continuous component near zero, the lattice case being similar. By hypothesis, the measure μ' has density at most A and is greater than some constant, $a > 0$, times Lebesgue measure on some interval $(-b, b)$. Call this latter measure π. Let $l = k + ck^{1/3}/b$. Write

$$\nu'_y(C(k,n)) = \int m_z(C(k+l,n))\, dm^y(z)$$

where m_z is the law of $(T_{k+l}, \ldots, T_n)$ conditioned on $T_{k+l} = z$ and m^y is the (deficient) law of T_{k+l} conditioned on $T_k = y$ and killed if $T_i \notin [ci^{1/3}/10, ci^{1/3}]$ for some $k \leq i \leq k+l$. It is easy to see that

$$m^y \geq \pi^y \geq a^l \text{ times Lebesgue measure on } [c(k+l)^{1/3}/10, c(k+l)^{1/3}]$$

where π^y is the measure on sequences $s_k, \ldots, s_{k+l}$ with $s_k = y$, having increments distributed as π and killed if $s_i \notin [ci^{1/3}/10, ci^{1/3}]$ for some $k \leq i \leq k+l$. Then for any $x, y \in [ck^{1/3}/10, ck^{1/3}]$,

$$\begin{aligned} & \nu'_y(C(k,n)) \\ = & \int m_z(C(k+l,n))\, dm^y(z) \\ \geq & a^l \int m_z(C(k+l,n))\, d\lambda(z) \\ \geq & (a/A)^l \int m_z(C(k+l,n))\, dm_x(z) \\ = & (a/A)^l \nu'_x(C(k,n)). \end{aligned}$$

Checking against the value of l proves (15) and the theorem. □

References

[1] Benjamini, I. and Peres, Y. (1993). Markov chains indexed by a tree. *Ann. Probab. to appear.*

[2] Doyle, P. and Snell, J. L. (1984). Random walks and electrical networks. Mathematical Association of America: Washington.

[3] Durrett, R. (1984). Brownian motion and martingales in analysis. Wadsworth: Monterey, CA.

[4] Kesten, H. (1978). Branching Brownian motion with absorption. *Stoch. Pro. Appl.* **7** 9 - 47.

[5] Lyons, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931 - 958.

[6] Lyons, R. (1992). Random walks, capacity and percolation on trees. *Ann. Probab. to appear.*

[7] Lyons, R. and Pemantle, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20** 125 - 136.

[8] Pemantle, R. (1988). Phase transition in Reinforced random walk and RWRE on trees. *Ann. Probab.* **16** 1229 - 1241.

[9] Pemantle, R. and Peres, Y. (1991). Critical random walk in random environment on a tree. *Preprint.*

Keywords: RWRE, critical RWRE, tree, tree-indexed random walk
Subject classification: Primary: 60J15. Secondary: 60G60, 60G70, 60E07.