Probab. Theory Relat. Fields 122, 567–592 (2002)
Digital Object Identifier (DOI) 10.1007/s004400100179

Rick Durrett · Harry Kesten · Vlada Limic

Once edge-reinforced random walk on a tree

Received: 6 March 2001 / Revised version: 16 July 2001
Published online: 15 March 2002 –

Abstract. We consider a nearest neighbor walk on a regular tree, with transition probabilities proportional to weights or conductances of the edges. Initially all edges have weight 1, and the weight of an edge is increased to $c > 1$ when the edge is traversed for the first time. After such a change the weight of an edge stays at c forever. We show that such a walk is transient for all values of $c \geq 1$, and that the walk moves off to infinity at a linear rate. We also prove an invariance principle for the height of the walk.

1. Introduction

Let $\mathcal{G}$ be an infinite connected graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$. Consider a Markov chain $\{X_n, w(e,n), e \in \mathcal{E}\}_{n\geq 0}$ which starts with $X_0 = v_0 \in \mathcal{G}$, $w(e, 0) = 1$ for all $e \in \mathcal{E}$. We think of the $w(e,n)$ as weights (or conductances) and the transitions for X will be to a nearest neighbor, with probabilities proportional to the weights of the incident edges. Formally,

$$P\{X_{n+1} = u | X_n = v, w(\cdot, n)\} = \frac{w(\{v,u\},n)}{\sum_{v'\sim v} w(\{v,v'\},n)}, \tag{1.1}$$

where $v' \sim v$ means that v' is adjacent to v in $\mathcal{G}$, and $\{v, v'\}$ denotes the (unoriented) edge between v and v'. After X_n has changed to X_{n+1} the weights are updated. Such systems were introduced by Coppersmith and Diaconis (1987) and are called *reinforced random walks*. For $\mathcal{G}$ a regular tree they were first studied by Pemantle (1988). His paper studied linear reinforcement. That is, Pemantle took $w(e,n) = 1 + k(c-1)$ after the edge e had been traversed k times by the random walk, for some fixed $c > 1$. Here we shall study a problem raised by Davis (1990), in which the weights are updated by the following rule:

$$w(e, n+1) = \begin{cases} c & \text{if } \{X_n, X_{n+1}\} = e \\ w(e,n) & \text{otherwise.} \end{cases} \tag{1.2}$$

Thus, the weight of the edge e is raised to c at the first jump across e, but then stays at the value c thereafter. This system is called a once-reinforced random walk.

R. Durrett, H. Kesten, V. Limic: Department of Mathematics, Malott Hall, Cornell University, Ithaca, NY 14853, USA.
e-mail: rtd1@cornell.edu; kesten@math.cornell.edu; limic@math.cornell.edu

Mathematics Subject Classification (2000): 60J15, 60K99

Key words or phrases: Reinforced random walk – Trees

The question is whether $\{X_n\}$ (which by itself is non-Markovian) is recurrent or transient. We only consider the case when $\mathcal{G}$ is a regular b-ary tree. On such a tree one expects the once-reinforced random walk $\{X_n\}$ to be transient, since for any fixed choice of $w(\cdot, n)$, which does not vary with n and is bounded away from 0 and ∞, the random walk with transition probabilities (1.1) is transient (this is easiest seen by interpreting the tree as an electrical network and bounding its resistance; see Lyons and Peres (1997), Ch. 2, or Doyle and Snell (1984), Ch. 6). Similarly one may conjecture that $\{X_n\}$ is transient on $\mathbb{Z}^d$ for $d \geq 3$ and recurrent for $d = 1, 2$. On the b-ary tree it is trivial to see that $\{X_n\}$ is indeed transient if $c < b$, for then the distance from X_n to a fixed vertex v_0 always has a positive drift. In this note we prove transience of the once-reinforced random walk on a regular tree for any choice of c. This contrasts with the results of Pemantle (1988) in the linearly reinforced case; he found transience for $c < c_0$ and recurrence for $c > c_0$ for some critical c_0.

Theorem 1. *If $\mathcal{G}$ is a rooted b-ary tree with $b \geq 2$ (that is, a tree in which all vertices have degree $b+1$ and some vertex v_0 has been singled out as the root) and $c > 1$, then the process $\{X_n\}$ is transient.*

Remark 1. *A slight extension of our proof shows that one still has transience on the rooted b-ary tree when each edge is reinforced up to k_0 times for some fixed $k_0 < \infty$, that is, if one takes $w(e, n) = 1 + (c - 1)\min(k, k_0)$ when the edge e has been traversed k times.*

It will be a simple consequence of some of the lemmas used to prove Theorem 1, that $X_.$ drifts off to infinity at a positive speed S. More specifically, let $h(v)$ be the graph distance of the vertex v from the root of the tree. $h(v)$ is also called the *height* of v. Then we also prove the following theorem.

Theorem 2. *Under the conditions of Theorem 1 there exists a constant $S = S(b, c) > 0$ such that*

$$\lim_{n\to\infty} \frac{1}{n} h(X_n) = S \text{ a.s.}. \tag{1.3}$$

Remark 2. *It is easy to show (see Section 3) that*

$$S \leq \frac{b}{b+c}. \tag{1.4}$$

This shows that $S \to 0$ as $c \to \infty$. However, we have no simple lower bound for S which is always strictly positive.

Our last theorem shows that $h(X_{nt})$ satisfies an invariance principle.

Theorem 3. *Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion and for noninteger values of nt, let $h(X_{nt})$ be the linear interpolation between $h(X_{\lfloor nt \rfloor})$ and $h(X_{\lfloor nt \rfloor} + 1)$. Under the conditions of Theorem 1 there exists a constant $\sigma = \sigma(b, c) > 0$ such that*

$$\left\{\frac{1}{\sqrt{n}}[h(X_{nt}) - nS(b,c)]\right\}_{0\le t\le 1} \Longrightarrow \{\sigma B(t)\}_{0\le t\le 1}. \tag{1.5}$$

Outline of proof of Theorem 1. Since the proof of Theorem 1 is surprisingly long we give here a brief outline. Assume that $h(X_n) = h$ at a certain time n. To show that $\{X_n\}$ is transient we estimate the probability that $h(X_.)$ will reach the level $h+1$ before $X_.$ returns to the root of the tree. We will prove a lower bound on this probability, uniformly in the history up till time n, which is such that the product of these bounds over $h = 1, 2, \dots$ is strictly positive. First we observe that for the X-process to return to the root from a position v with $h(v) = h$, it has to go back through the unique self-avoiding path from the root to v. Let this path be $r = (v_0, v_1, \dots, v_h)$ with v_0 the root and $v_h = v$. By considering the imbedded random walk on r we first find a crude lower bound for the probability of the height to increase by 1 to $h+1$ before X goes back to the root. By multiplying these crude estimates one easily obtains a crude lower bound of order $h^{-c/b}$ for the probability of reaching height h from the root, before returning to the root.

The main step is to improve this crude bound by noting that while $X_.$ goes back through r, each time it visits v_j, it may move into the subtree of descendants of v_j not containing v_{j+1}. In this subtree it can then reach height $h+1$ without returning to the root. Of course we do not know the probability to reach height $h+1$ in this side tree, but we have gained that there are typically many visits to the v_j before reaching the root, and therefore many occasions to move up into a side tree and not to return to the root before reaching level $h+1$.

Nevertheless we need some lower bound for the probability to move up in the side tree to height $h+1$ before returning to v_0. Such a bound is derived by first showing that in the many visits to v_j a good piece of the subtree of descendants of v_j is visited. In fact, we show that in m visits to v_j, typically all descendants of v_j in the next $m^{1/4}$ generations are visited. Moreover, the average number of visits to v_j by the imbedded random walk on r as it moves from v to v_0 is of order h. This means that all edges within distance $h^{1/4}$ of v_j in the subtree of its descendants (other than v_{j+1}) are reinforced. In this piece, where all edges are reinforced, $X_.$ behaves like simple random walk. By the known behavior of simple random walk on a tree this means that once $X_.$ moves into the subtree of descendants of v_j, it tends to move away at least distance $h^{1/4}$ from v_j and to spend a lot of time there before it returns to v_j. During this time it has many chances (in fact typically at least $\exp[Kh^{1/4}]$ many chances, for some constant $K > 0$) to move distance $h+1$ away from v_j. For each of these occasions to move away distance h we use the previously derived crude lower bound of order $h^{-c/b}$ that this will actually happen. With $\exp[Kh^{1/4}]$ many chances we end up with an excellent lower bound that the X-process will move from height h to height $h+1$ before returning to the root.

2. Proof of Theorem 1

Throughout we use T to denote a rooted b-ary tree. The *root* of the tree is denoted by v_0. The distance from a vertex v to the root is also called the height of v and is denoted by $h(v)$. A vertex v' is called a *descendant* of v if the unique self-avoiding path in T from the root to v' passes through v. v is regarded as a descendant of itself. If v' is a descendant of v which is adjacent to v, then v' is called a *child* of v. In this case we also call v the *parent* of v'. The subtree of T whose vertex set is the collection of descendants of v is denoted by $T(v)$ (so that $T = T(v_0)$. We further define the σ-field

$$\mathcal{F}_n = \sigma(X_k, w(e,k),\ e \in \mathcal{E}, 0 \le k \le n).$$

We use K_i for various strictly positive, finite constants.

The proof is broken up into a sequence of lemmas. To start we need a very simple zero-one law which makes the terms "recurrent" and "transient" unambiguous. Note that the proof of even this simple lemma already relies on T being a tree.

Let A_0 be the condition

$$A_0 = \{X_0 = v_0, w(e,0) = 1, e \in \mathcal{E}\}.$$

We remind the reader that if we do not indicate otherwise, then we start our process with this initial condition.

Lemma 1. *(i) If*

$$P\{X_n \textit{ never returns to } v_0 | A_0\} > 0, \tag{2.1}$$

then for all $v \in \mathcal{V}$ and all k,

$$P\{X_n \textit{ visits } v \textit{ only finitely often} | \mathcal{F}_k\} = 1. \tag{2.2}$$

(ii) If

$$P\{X_n \textit{ never returns to } v_0 | A_0\} = 0, \tag{2.3}$$

then for all $v \in \mathcal{V}$ and all k,

$$P\{X_n \textit{ visits } v \textit{ infinitely often} | \mathcal{F}_k\} = 1. \tag{2.4}$$

Proof. (i) Define

$$\alpha := P\{X_n \text{ never returns to } v_0 | A_0\}. \tag{2.5}$$

Further let

$$A_m(v) = \{X_m = v \text{ but } X_n \neq v \text{ for all } 0 \le n < m\}.$$

On this event the random walk visits v, and hence also $T(v)$, for the first time at time m. Consequently if we started in A_0, then all edges in $T(v)$ still have weight 1 at time m. Therefore, on the event $A_m(v)$, for any child $\hat{v}$ of v, the subtree $T(\hat{v})$ looks exactly like the subtree $T(v_1)$ for any child v_1 of v_0 on the event A_0. Thus, on $A_m(v)$,

$$\begin{aligned} &P\{X_n \neq v_0 \text{ for all } n \ge m | \mathcal{F}_m\} \\ &\quad \ge P\{X_n \in T(v) \setminus \{v\} \text{ for all } n > m | \mathcal{F}_m\} \\ &\quad \ge P\{X_{m+1} \text{ is some child } \hat{v} \text{ of } v \text{ and } X_n \in T(\hat{v}) \text{ for } n \ge m+1 | \mathcal{F}_m\} \\ &\quad \ge \frac{b}{b+c}\alpha. \end{aligned} \tag{2.6}$$

It is easy to see that almost surely there exist infinitely many m and vertices v_m such that $A_m(v_m)$ occurs. Thus if (2.1) holds, then

$$\limsup_{m\to\infty} P\{X_n \neq v_0 \text{ for } n \geq m | \mathcal{F}_m\} \geq \frac{b}{b+c}\alpha > 0 \text{ a.s.}. \tag{2.7}$$

It is a well known consequence of the martingale convergence theorem (compare Breiman (1968), Problem 5.6.9) that this implies that

$$P\{X_n = v_0 \text{ for only finitely many } n | A_0\} = 1. \tag{2.8}$$

Now one easily sees that for any vertex v there exists some N, and a constant $K_0 > 0$ such that for all n

$$P\{X_{n+N} = v_0 | A_0, \mathcal{F}_n\} \geq K_0 \text{ a.e. on } \{X_n = v\}.$$

Thus, again by Breiman (1968), Problem 5.6.9, (2.8) remains valid if v_0 is replaced by any v. Hence, (2.2) follows from (2.1).

To prove (ii) we first show that (2.3) implies

$$P\{X_n \text{ visits } v_0 \text{ infinitely often} | A_0\} = 1. \tag{2.9}$$

To this end we consider the events A of the form

$$A = \{X_0 = v_0, w(e,0) = w_e\}, w_e \in \{1, c\}, e \in \mathcal{E}, \tag{2.10}$$

with the further requirement that

$$\text{the subgraph } T^r = T^r(A) \text{ spanned by the collection of (reinforced) edges, } \{e : w_e = c\}, \text{ is connected and contains } v_0. \tag{2.11}$$

Note that this last condition is equivalent to the condition that T^r is a tree containing v_0. If we start on A_0, then at each time n an A of the form (2.10), (2.11) must occur. To prove (2.9) it suffices to show that $P\{X_n \text{ never returns to } v_0 | A\} = 0$, for all A of the form (2.10) and satisfying (2.11). Assume to the contrary that there exists some A^* of the form (2.10) and satisfying (2.11) and with the property

$$P\{X_n \text{ never returns to } v_0 | A^*\} > 0. \tag{2.12}$$

If $X_n \neq v_0$ for $n \geq 1$, then X_n must stay in some subtree $T(v_1)$ for $n \geq 1$, where v_1 is one of the children of v_0. Thus (2.12) implies

$$P\{X_n \in T(v_1), n \geq 1 | A^*\} > 0 \tag{2.13}$$

for some child v_1 of v_0. Let $w^*(e)$ be the weights in the definition of A^*. The last probability can be changed only by a bounded factor if we change the weights of the edges outside $T(v_1) \cup \{v_0, v_1\}$; in fact such a change can have an effect only on the probability that X moves into $T(v_1)$ at the first step. We may further assume that $w^*(\{v_0, v_1\}) = c$, because X has to move from v_0 to v_1 in the first step for the event in (2.13) to occur, and hence the weight of $\{v_0, v_1\}$ has to be c after the first step anyway. (2.13) therefore remains valid if we replace w^* by

$$w^{**}(e) = \begin{cases} w^*(e) & \text{if } e \in T(v_1) \\ c & \text{if } e = \{v_0, v_1\} \\ 1 & \text{for all other } e. \end{cases}$$

Let $B_1(m) = \{X_m = v_1, X_i \neq v_0, 1 \leq i \leq m\}$, and $B_2(m) = \{w(e,m) = w_e^{**}\}$. It is easy to see that

$$P\{B_1(m) \cap B_2(m) | A_0\} > 0, \tag{2.14}$$

for some finite m. In words, it is possible for X to traverse $\{v_0, v_1\}$ plus all the edges of $T(v_1)$ which are reinforced in A^* and no other edge before returning to vertex v_1. Now (2.14) and (2.12) show that

$$\begin{aligned} &P\{B_1(m) \cap B_2(m) \text{ and } X \text{ never returns to } v_0 | A_0\} \\ &\quad \geq P\{B_1(m) \cap B_2(m) | A_0\} P\{X_n \in T(v_1) | A^*\} > 0, \end{aligned}$$

which contradicts (2.3). Thus (2.3) implies (2.9).

As in the lines following (2.8) one now sees that (2.9) remains valid if v_0 is replaced by an arbitrary v. In turn, this implies (2.4). □

Of course the walk $\{X_n\}$ is called *recurrent* if (2.3), and hence (2.4), holds. Otherwise it is called *transient*. In the transient case $\alpha > 0$, by definition. *For the remainder of this section we assume that $\{X_n\}$ is recurrent and will derive a contradiction from this.*

We next derive a very crude estimate for the probability that $\{h(X_n)\}$ reaches the value $h+1$ from a vertex v of height h, before it goes down by an amount $1 \leq k \leq h$.

Lemma 2. *Define*

$$\tau(n, j) = \inf\{t > n : h(X_t) = j\}, \tag{2.15}$$

and for any vertex v of T with $h(v) = h$,

$$\beta(k, v) = \sup_n \text{ ess } \sup_\omega P\{\tau(n, h(v) - k) < \tau(n, h(v) + 1) | \mathcal{F}_n\},$$

where the ess sup is over all sample points ω with $X_n(\omega) = v$. Then

$$\beta(k, v) \leq \frac{c}{c + bk}, \quad h(v) = h \geq 1, 1 \leq k \leq h. \tag{2.16}$$

Moreover, for some constant $K_1 = K_1(c/b) > 0$, and all vertices v

$$\begin{aligned} P\{X_{n+1} \textit{ is a child of } X_n \textit{ and } \quad &\tau(n, h(v) + q) < \tau(n, h(v)) | \mathcal{F}_n\} \geq K_1 q^{-c/b} \\ &\textit{a.s. on the event} \{X_n = v\}, \quad q \geq 1. \end{aligned} \tag{2.17}$$

Proof. Let $h \geq 1$ and let $X_n = v$ with $h(v) = h$ and $0 \leq k \leq h$. Let $r = (v_0, v_1, \dots, v_h = v)$ be the unique self-avoiding path in T from the root to v. We now consider the imbedded random walk on r. That is, we only look at the successive positions of $\{X_j\}$ which lie in r and which differ from the last preceding position in r. Since $X_n = v$, and the random walk started at $X_0 = v_0$, X_j must have traversed all the edges $\{v_i, v_{i+1}\}, 0 \leq i \leq n-1$, by time n. Therefore, at time n all these edges have been reinforced already and have weight c. Thus, when the imbedded random walk is at v_j for some $0 < j < h$ at some time $\geq n$, then the next position is v_{j-1} or v_{j+1}, each possibility having probability 1/2. Thus on the interior of r, the imbedded random walk is just symmetric simple random walk. Note that by our assumption of recurrence the imbedded random walk will take with probability 1 infinitely many steps. Now for $\tau(n, h(v)-k) < \tau(n, h(v)+1)$ to occur, it must be that at every time t when the walk returns to $v = X_n$, it moves to v_{h-1} at the next step . The conditional probability of a step from $v = v_h$ to v_{h-1}, given $\mathcal{F}_t$, is at most $c/(b+c)$. When the walk is at v_{h-1} it has a conditional probability of reaching $v_h - k$ before v_h of $(1 - 1/k)$ (because it behaves like symmetric simple random walk till it hits v_{h-k} or v_h). Thus decomposing the event $\tau(n, h-k) < \tau(n, h+1)$ with respect to the number of times the walk moves from v to v_{h-1} before $\tau(n, h-k)$, we obtain

$$P\{\tau(n, h-k) < \tau(n, h+1)|\mathcal{F}_n\} \leq \sum_{j=1}^{\infty} \Big[\frac{c}{b+c}\Big]^j \big(1 - \frac{1}{k}\big)^{j-1} \frac{1}{k} = \frac{c}{c+bk}. \quad (2.18)$$

This, of course, implies (2.16).

For (2.17) note that if X_{n+1} is a child of $X_n = v$, then $\tau(n, h(v))$ is just the time of the first return after time n to v. The bound (2.17) now follows from the fact that almost everywhere on $\{X_n = v\}$,

$$\begin{aligned}
&P\{X_{n+1} \text{ is a child of } v \text{ and } h(X_.) \text{ reaches } h(v)+q \text{ before } X_. \text{ reaches } v_0|\mathcal{F}_n\} \\
&\quad \geq P\{X_{n+1} \text{ equals a child } \hat{v} \text{ of } v|\mathcal{F}_n\} \\
&\quad \times \prod_{j=1}^{q-1} \text{ess inf } P\{h(X_.) \text{ reaches } h(v)+j+1 \text{ before } X_. \text{ reaches } v|\mathcal{F}_n, X_{n+1} = \hat{v} \\
&\qquad\qquad \text{and } h(X_.) \text{ has reached } h(v)+\ell \text{ before } \tau(n, h(v)) \text{ for } 1 \leq \ell \leq j\} \\
&\quad \geq \frac{b}{b+c} \prod_{j=1}^{q-1} [1 - \sup_{v: h(v)=n+j} \beta(j, v)] \geq \frac{b}{b+c} \prod_{j=1}^{q-1} \frac{bj}{c+bj} \\
&\quad = \frac{b}{b+c} \exp\Big[-\sum_{j=1}^{q-1} \big(\frac{c}{bj} + O(j^{-2})\big)\Big] \\
&\quad (\text{since } 1-x = \exp[\log(1-x)] = \exp[-x + O(x^2)]) \\
&\quad \geq K_1 q^{-c/b}.
\end{aligned} \quad (2.19)$$

□

The estimates in the last proof only relied on the imbedded random walk on the path r, and ignored "excursions" into the rest of the tree. We shall use these excursions to improve the bound (2.17). To deal with such excursions we denote the children of v_0 in some arbitrary order as $v_{0,1}, v_{0,2}, \ldots, v_{0,b+1}$. Fix some z and write $z \le t_1 < t_2 < \ldots$ for the successive times $t \ge z$ for which $X_t = v_0$, $X_{t+1} = v_{0,b+1}$ and let $s_p = \inf\{s \ge t_p : X_s = v_0\}$. Then we regard $[t_p, s_p]$ as the time interval of the p-th excursion after time z into $T(v_{0,b+1})$. We next define the maximal height reached during the first L of these excursions, that is

$$H(z, L) = \max\{h(X_\ell) : t_p \le \ell \le s_p \text{ for some } 1 \le p \le L\}.$$

We also define

$$\gamma(L, q) = \inf_z \operatorname{ess\,inf}_\omega P\{H(z, L) \ge q | \mathcal{F}_z\}, \tag{2.20}$$

For the next lemma we again assume that $X_n = v$ for a v with $h(v) = h$ and that $r = (v_0, \ldots, v_h = v)$ is the unique self-avoiding path from v_0 to v. For $1 \le j < h$ we denote the children of v_j as $v_{j,1} = v_{j+1}, v_{j,2}, \ldots, v_{j,b}$. We further introduce

$$L = L(n, j, k) = \text{number of } t \in (n, \tau(n, 0)) \text{ for which } X_t = v_j, X_{t+1} = v_{j,k}. \tag{2.21}$$

Lemma 3.

$$\beta(h, v)$$
$$\le \sup_n \operatorname{ess\,sup}_\omega E\left\{I[\tau(n, 0) < \tau(n, h+1)] \prod_{j=1}^{h-1} \prod_{k=2}^{b} [1 - \gamma(L(n, j, k), h-j+1)] | \mathcal{F}_n\right\}. \tag{2.22}$$

The expectation here is over the conditional distribution of the imbedded random walk on r and the random variables $L(n, j, k)$.

Proof. This lemma is almost obvious after one deciphers the notation. We now have to consider "excursions" into the various subtrees $T(v_{j,k})$. Let $t_1(j, k) < t_2(j, k) < \ldots < t_L(j, k)$ be the successive times in $(n, \tau(n, 0))$ for which $X_t = v_j$, $X_{t+1} = v_{j,k}$ and let $s_p(j, k) = \inf\{m > t_p(j, k) : X_m = v_j\}$. Thus $[t_p(j, k), s_p(j, k)]$ is the time interval of the p-th excursion into $T(v_{j,k})$. The maximal heights above j reached during these excursions are the quantities

$$\widehat{H}(n, j, k, L(n, j, k))$$
$$:= \max\{h(X_m) - j : t_p(j, k) \le m \le s_p(j, k) \text{ for some } 1 \le p \le L(n, j, k)\}.$$

If the walk reaches height $j + (h - j + 1) = h + 1$ during any of the intervals $[t_p(j, k), s_p(j, k)]$, then $h(X_.)$ reaches the value $h + 1$ during $[n, \tau(n, 0))$, and this excludes the occurrence of $\tau(n, 0) < \tau(n, h + 1)$. Thus,

$$\beta(h, v)$$
$$\le \sup_\omega E\left\{I[\tau(n, 0) < \tau(n, h+1)] \prod_{j=1}^{h-1} \prod_{k=2}^{b} I[\widehat{H}(n, j, k, L(n, j, k)) < h-j+1] | \mathcal{F}_n\right\}.$$

To complete the proof we observe that the imbedded random walk and the $L(n, j, k)$ only determine the numbers of excursions into the various $T(v_{j,k})$, but the behaviors of $X_.$ during these excursions and the weights in the subtrees $T(v_{j,k})$ are independent of the imbedded random walk on r and the $L(n, j, k)$. We now condition on $\mathcal{F}_n$ and the imbedded random walk, $I[\tau(n,0) < \tau(n, h+1)]$ and the $L(n, j, k)$. These data also determine the weights in the $T(v_{j,k})$ at time n. These weights remain unchanged till the time of the first excursion into $T(v_{j,k})$. Moreover, the excursions into one such subtree $T(v_{j,k})$ do not influence the weights in the other subtrees $T(v_{j',k'})$ with $(j', k') \neq (j, k)$. Consequently, the conditional expectation of

$$\prod_{j=1}^{h-1}\prod_{k=2}^{b} I[\widehat{H}(n, j, k, L(n, j, k)) < h - j + 1],$$

given $\mathcal{F}_n$, the imbedded random walk on r, and the $L(n, j, k)$, is at most

$$\prod_{j=1}^{h-1}\prod_{k=2}^{b}[1 - \gamma(L(n, j, k), h - j + 1)].$$

The bound (2.22) follows. □

What is needed to complete the proof of Theorem 1 is a good lower bound for $\gamma(L, q)$. We already have one lower bound, which is a simple consequence of (2.17). Indeed, almost everywhere on $\{X_m = v_{0,b+1}\}$

$$\begin{aligned}P\{H(m, L) \geq q|\mathcal{F}_m\} &= 1 - P\{h(X_\ell) < q \text{ for all } t_p \leq \ell \leq s_p, 1 \leq p \leq L|\mathcal{F}_m\}\\ &\geq 1 - \Big[\operatorname*{sup\,ess}_{n} \sup_{\omega: X_n(\omega)=v_{0,b+1}} P\{\tau(n, q) > \tau(n,0)|\mathcal{F}_n\}\Big]^L\\ &\geq 1 - \big[1 - K_1 q^{-c/b}\big]^L,\end{aligned}$$

so that

$$\gamma(L, q) \geq 1 - \big[1 - K_1 q^{-c/b}\big]^L. \tag{2.23}$$

For large c this bound is not strong enough and we now set about improving (2.23).

Lemma 4. *Let v be an arbitrary vertex of T and u a descendant of v such that*

$$h(u) - h(v) = k.$$

Let σ_0 be an $\{\mathcal{F}_n\}$-stopping time such that

$$X_{\sigma_0} = v \text{ a.s.}, \tag{2.24}$$

and let

$$\sigma_j = \inf\{n > \sigma_{j-1} : X_n = v\}, \quad j \geq 1. \tag{2.25}$$

Then for some constants $0 < K_2 = K_2(b, c) < 1$ and $K_3 = K_3(b, c) > 0$, and for

$$k^2 \geq K_2 \text{ and } m \geq \left\lceil \frac{2k^2}{K_2} \right\rceil \tag{2.26}$$

it holds that

$$P\{X_{.}\ \textit{visits } u \textit{ during } [\sigma_0, \sigma_m] | \mathcal{F}_{\sigma_0}\} \geq 1 - \exp\left[-\frac{2K_3 m}{k^2}\right]. \tag{2.27}$$

Proof. Let $s = (u_0 = v, u_1, u_2, \ldots, u_k = u)$ be the unique self-avoiding path in T from v to u. In fact, this path necessarily lies in $T(v)$. Assume that for some i and some $0 \leq \ell < k$,

$$X_{.} \text{ has visited } u_0, u_1, \ldots, u_\ell \text{ during } [\sigma_0, \sigma_i]. \tag{2.28}$$

We claim that there exists some constant $0 < K_2 = K_2(b, c) < 1$, such that on this event one has

$$P\{X_{.} \text{ visits } u_{\ell+1} \text{ during } (\sigma_i, \sigma_{i+1}] | \mathcal{F}_{\sigma_i}\} \geq \frac{K_2}{\ell + 1}. \tag{2.29}$$

This follows immediately by considering the imbedded random walk on s. As in the proof of Lemma 2 this imbedded random walk is just a symmetric simple random walk, as long as it is on $s \setminus \{u_0, u_\ell\}$. By our choice of the σ_j we may take $X_{\sigma_i} = v = u_0$. The probability on the left of (2.29) is therefore at least

$$\begin{aligned} P\{X_{\sigma_i+1} = u_1 | \mathcal{F}_{\sigma_i}\} \cdot P\{&\text{imbedded random walk reaches } u_\ell \\ &\text{before returning to } u_0 | \text{imbedded random walk starts at } u_1\} \\ &\times \inf_m P\{X_{m+1} = u_{\ell+1} | \mathcal{F}_m, X_m = u_\ell\} \geq \frac{1}{1+bc} \frac{1}{\ell} \frac{1}{1+bc}. \end{aligned}$$

This implies the claimed (2.29) with $K_2 = [1 + bc]^{-2}$ when $1 \leq \ell \leq k - 1$, and the case $\ell = 0$ is trivial.

By iterating the bound (2.29) we get that on the event (2.28),

$$P\{X_{.} \text{ visits } u_{\ell+1} \text{ during } (\sigma_i, \sigma_{i+q}] | \mathcal{F}_{\sigma_i}\} \geq 1 - \left[1 - \frac{K_2}{\ell + 1}\right]^q. \tag{2.30}$$

We apply this in the following form. First we define

$$i_\ell = \inf\{m : X_{.} \text{ reaches } u_\ell \text{ during } [\sigma_0, \sigma_m]\}.$$

Then (2.30) implies that almost everywhere on (2.28)

$$P\{i_{\ell+1} - i_\ell > q | \mathcal{F}_{\sigma_{i_\ell}}\} \leq \left[1 - \frac{K_2}{\ell + 1}\right]^q,$$

and consequently,

$$E\{i_{\ell+1} - i_\ell | \mathcal{F}_{\sigma_{i_\ell}}\} \leq \frac{\ell + 1}{K_2}.$$

But then also

$$E\{i_k - i_0 | \mathcal{F}_{\sigma_0}\} \leq \sum_{\ell=0}^{k-1} \frac{\ell + 1}{K_2} \leq \frac{k^2}{K_2}$$

and

$$P\{X_{.} \text{ does not visit } u_k = u \text{ during } (\sigma_0, \sigma_{\lceil 2k^2/K_2 \rceil}] | \mathcal{F}_{\sigma_0}\}$$
$$\leq P\Big\{i_k - i_0 \geq \frac{2k^2}{K_2} \Big| \mathcal{F}_{\sigma_0}\Big\} \leq \frac{1}{2}. \tag{2.31}$$

Finally we consider in succession the time intervals $(\sigma_{j\lceil 2k^2/K^2 \rceil}, \sigma_{(j+1)\lceil 2k^2/K_2 \rceil}]$. In each of these intervals there is a conditional probability of at least 1/2 for $X_{.}$ to visit u, so that the left hand side of (2.27) is at least

$$1 - 2^{\{-\lfloor m/\lceil 2k^2/K_2 \rceil \rfloor\}},$$

and (2.27) follows. □

The next lemma improves the preceding one. It shows that $X_{.}$ can reach descendants of v at much greater height than indicated by (2.27). We introduce the following events:

$$A = A(m, v) = \{X_{.} \text{ visits all descendants } u \text{ of } v \text{ with } h(u) - h(v) \leq m^{1/4} \text{ during } [\sigma_0, \sigma_m]\},$$

and

$$B = B(m, v) = \{X_{.} \text{ visits some descendant } u \text{ of } v \text{ with } h(u) - h(v) \geq m^2 \text{ during } [\sigma_0, \sigma_{2m}]\}.$$

Lemma 5. *There exists some constant $K_4 = K_4(b, c) < \infty$ such that with v and σ_i as in Lemma 4, it holds for $m \geq K_4$,*

$$P\{A(m, v) | \mathcal{F}_{\sigma_0}\} \geq 1 - \exp[-K_3 m^{1/2}]. \tag{2.32}$$

Moreover, for $m \geq K_4$,

$$P\{B(m, v) | \mathcal{F}_{\sigma_0}\} \geq 1 - 2\exp[-K_3 m^{1/2}]. \tag{2.33}$$

Proof. There are at most

$$b^{m^{1/4}+1}$$

descendants u of v with $h(u) - h(v) \leq m^{1/4}$. Apply (2.27) with k replaced by $\lfloor m^{1/4} \rfloor$ to each of these descendants u to obtain (2.32).

To derive (2.33) we assume that $A(m, v)$ occurred, and we find a lower bound for the conditional probability of $B(m, v)$, given $\mathcal{F}_{\sigma_m}$, on the event A. To this end we first note that (2.17) and the strong Markov property give for any vertex $\widehat{v}$ and any $\{\mathcal{F}_n\}$-stopping time ρ for which $X_\rho = \widehat{v}$,

$$P\{X_{\rho+1} \text{ is a child of } X_\rho = \widehat{v} \text{ and } h(X_{.}) \text{ reaches } h(\widehat{v}) + m^2 \text{ before returning to } \widehat{v} | \mathcal{F}_\rho\} \geq K_1 m^{-2c/b}. \tag{2.34}$$

Now consider one of the times σ_i with $i \geq m$. Recall that $X_{\sigma_i} = v$ by our assumption on the σ_i. Suppose next that X_{σ_i+1} is a child of $X_{\sigma_i} = v$ and assume that $X_{.}$ continues to move in $T(v) \setminus \{v\}$ till some stopping time ρ_1 at which it reaches a vertex v_1 with $h(v_1) = h(v) + \lfloor m^{1/4} \rfloor$. Conditionally on this event there is a probability of at least $K_1 m^{-2c/b}$ that $X_.$ will move to a child of v_1 at the next step and then will continue to move in $T_{v_1} \setminus \{v_1\}$ until it reaches height $h(v_1) + m^2 \geq h(v) + m^2$. In this case $X_.$ actually reaches a vertex of height $h(v) + m^2$ before time σ_{i+1} (recall that this is the time of the first return to v after σ_i). If X_{ρ_1+1} is not a child of v_1, or $X_.$ returns to v_1 before it reaches height $h(v_1) + m^2$ in T_{v_1}, then we try again to reach height $h(v) + m^2$ from some vertex v_2 of height $h(v_2) = h(v) + \lfloor m^{1/4} \rfloor$, etc. We continue this till $X_.$ returns to v at time σ_{i+1}.

We now give a more formal description of the procedure in the preceding paragraph, which also lets us estimate the probability that $h(X_.)$ reaches $h(v) + m^2$ before σ_{m+1} by this procedure. We define

$$\rho_1 = \inf\{n > \sigma_i : h(X_n) = h(v) + \lfloor m^{1/4} \rfloor\},$$
$$\rho_{j+1} = \inf\{n > \rho_j : h(X_n) = h(v) + \lfloor m^{1/4} \rfloor\}.$$

We further define

$$v_j = X_{\rho_j}$$

and

$$\nu = \begin{cases} 0 & \text{if } \rho_1 \geq \sigma_{i+1} \\ \max\{j : \rho_j < \sigma_{i+1}\} & \text{if } \rho_1 < \sigma_{i+1}. \end{cases}$$

We also use the notation

$$a \wedge b = \min(a, b).$$

Then, by (2.34), as explained in the preceding paragraph, on the event $\{\rho_j < \sigma_{i+1}\}$,

$$P\{h(X_.) \text{ does not reach } h(v)+m^2 \text{ during } [\rho_j, \rho_{j+1} \wedge \sigma_{i+1}) | \mathcal{F}_{\rho_j}\} \leq 1 - K_1 m^{-2c/b}.$$

Now assume that $A(m, v)$ occurs and let $i \geq m$. Then, since $\{\nu \geq j\} = \{\rho_j < \sigma_{i+1}\}$, we have for any $M \geq 1$

$$\begin{aligned}
&P\{h(X_.) \text{ does not reach } h(v) + m^2 \text{ during } [\sigma_i, \sigma_{i+1}] | \mathcal{F}_{\sigma_i}\} \\
&\quad \leq P\{\nu < M | \mathcal{F}_{\sigma_i}\} \\
&\qquad + E\Big\{ I[\nu \geq M] I\big[h(X_.) \text{ does not reach } h(v) + m^2 \text{ during } [\sigma_i, \rho_M)\big] \\
&\qquad \times P\{h(X_.) \text{ does not reach } h(v) + m^2 \text{ during } [\rho_M, \sigma_{i+1}) | \mathcal{F}_{\rho_M}\} \Big| \mathcal{F}_{\sigma_i} \Big\} \\
&\quad \leq P\{\nu < M | \mathcal{F}_{\sigma_i}\} + [1 - K_1 m^{-2c/b}] E\Big\{ I[\nu \geq M - 1] \\
&\qquad \times P\{h(X_.) \text{ does not reach } h(v) + m^2 \text{ during } [\sigma_i, \rho_{M-1} \wedge \sigma_{i+1}) | \mathcal{F}_{\rho_{M-1}}\} \Big| \mathcal{F}_{\sigma_i} \Big\} \\
&\quad \leq P\{\nu < M | \mathcal{F}_{\sigma_i}\} + [1 - K_1 m^{-2c/b}]^2 E\Big\{ I[\nu \geq M - 2] \\
&\qquad \times P\{h(X_.) \text{ does not reach } h(v) + m^2 \text{ during } [\sigma_i, \rho_{M-2} \wedge \sigma_{i+1}) | \mathcal{F}_{\rho_{M-2}}\} \Big| \mathcal{F}_{\sigma_i} \Big\} \\
&\quad \cdots \leq P\{\nu < M | \mathcal{F}_{\sigma_i}\} + [1 - K_1 m^{-2c/b}]^{M-1}.
\end{aligned} \tag{2.35}$$

To complete the proof we bound

$$P\{\nu < M|\mathcal{F}_{\sigma_i}\} = P\{\rho_M \geq \sigma_{i+1}|\mathcal{F}_{\sigma_i}\} \tag{2.36}$$

and choose M appropriately. To obtain such a bound we note that on the event $A(m, v)$ all vertices on the self-avoiding path from v_0 to v and all vertices u in $T(v)$ with $h(u) \leq h(v) + m^{1/4}$ have been visited by time σ_m. Thus, at time $\sigma_i \geq \sigma_m$, all edges incident to a $u \in T(v)$ with $h(u) \leq h(v) + m^{1/4} - 1$ have been reinforced. Consequently, on $A(v, m) \cap \{X_{\sigma_i+1}$ is a child of $X_{\sigma_i}\}$, $X_.$ behaves like simple random walk on T during $[\sigma_i + 1, \rho_1 \wedge \sigma_{i+1}]$. If we write $\{\widehat{X}_n\}$ for simple random walk on T, then this shows that on $A(v, m)$

$$\begin{aligned}
&P\{\rho_1 \geq \sigma_{i+1}|\mathcal{F}_{\sigma_i}\}\\
&\quad = P\{X_{\sigma_i+1} \text{ is the parent of } v|\mathcal{F}_{\sigma_i}\} + P\{X_{\sigma_i+1} \text{ is a child of } X_{\sigma_i}|\mathcal{F}_{\sigma_i}\}\\
&\qquad \times P\{\widehat{X}_. \text{ returns to } v_0 \text{ before } h(\widehat{X}_.) \text{ reaches } m^{1/4}|\widehat{X}_0 \text{ is a child of } v_0\}\\
&\quad = \frac{c}{(b+1)c}\\
&\qquad + \frac{bc}{(b+1)c} P\{\widehat{X}_. \text{ returns to } v_0 \text{ before } h(\widehat{X}_.) \text{ reaches } m^{1/4}|\widehat{X}_0 \text{ is a child of } v_0\}\\
&\quad \leq \frac{1}{b+1} + \frac{b}{b+1} P\{\widehat{X}_. \text{ ever returns to } v_0|\widehat{X}_0 = v_0\}\\
&\quad = \frac{1}{b+1} + \frac{b}{b+1}\frac{1}{b} = \frac{2}{b+1}.
\end{aligned}$$

In the one but last equality we used that $h(\widehat{X}_.)$ is a nearest neighbor random walk on $\{0, 1, 2, \ldots\}$ with

$$\begin{aligned}
&P\{h(\widehat{X}_{n+1}) = h(\widehat{X}_n) + 1|\widehat{X}_0, \ldots, \widehat{X}_n\}\\
&\qquad\qquad = 1 - P\{h(\widehat{X}_{n+1}) = h(\widehat{X}_n) - 1|\widehat{X}_0, \ldots, \widehat{X}_n\} = \frac{b}{b+1}
\end{aligned}$$

and Feller (1968), formula XIV.2.4 or XIII.4.5. For similar reasons, one has for $j \geq 1$ on $A(m, v) \cap \{\rho_j < \sigma_{i+1}\}$

$$P\{\rho_{j+1} \geq \sigma_{i+1}|\mathcal{F}_{\rho_j}\} \leq P\{\widehat{X}_. \text{ ever reaches } v_0|h(\widehat{X}_0) = \lfloor m^{1/4}\rfloor\} \leq b^{-\lfloor m^{1/4}\rfloor}.$$

It follows that

$$P\{\nu < M|\mathcal{F}_{\sigma_i}\} \leq \frac{2}{b+1} + \frac{b-1}{b+1}(M-2)b^{-\lfloor m^{1/4}\rfloor}.$$

We now take

$$M = b^{\frac{1}{2}\lfloor m^{1/4}\rfloor}$$

so that

$$P\{\nu < M|\mathcal{F}_{\sigma_i}\} \leq \frac{5/2}{b+1} \tag{2.37}$$

for $m \geq$ some constant $K_5 = K_5(b)$. When we substitute the last estimate in (2.35) we see that for $K_4 = K_4(b, c)$ large enough and $i \geq m \geq K_4$, on the event $A(v, m)$, it holds that

$$\begin{aligned} P\{h(X_{.}) \quad &\text{does not reach } h(v) + m^2 \text{ during } [\sigma_i, \sigma_{i+1}] | \mathcal{F}_{\sigma_i}\} \\ &\leq \frac{b + 1/2}{b+1} + [1 - K_1 m^{-2c/b}]^{M-1} \leq \frac{b+3/4}{b+1}. \end{aligned}$$

Of course this gives for $m \geq K_4$, on the event $A(v, m)$, that

$$P\{h(X_{.}) \text{ does not reach } h(v) + m^2 \text{ during } [\sigma_m, \sigma_{2m}] | \mathcal{F}_{\sigma_m}\} \leq \left[\frac{b+3/4}{b+1}\right]^m. \tag{2.38}$$

Finally, the left hand side of (2.33) is at least

$$P\{A(v, m) | \mathcal{F}_{\sigma_0}\} \left[1 - \left[\frac{b+3/4}{b+1}\right]^m\right],$$

which together with (2.32) proves (2.33) (possibly after a suitable increase of K_4). □

We are finally ready for the required lower bound of $\gamma(L, q)$.

Lemma 6. *For some constants $K_i = K_i(b, c) < \infty$ and $L \geq K_6$,*

$$\gamma(L, \lfloor L/2 \rfloor^2) \geq 1 - \exp[-K_7 L^{1/2}]. \tag{2.39}$$

In addition, for all vertices v with $h = h(v) \geq 1$,

$$\beta(h(v), v) \leq K_8 \exp[-K_9 h^{1/4}]. \tag{2.40}$$

Proof. We apply (2.33) with $m = \lfloor L/2 \rfloor,\ v = v_{0,b+1}$ and $\sigma_0 = t_p + 1$, where t_p and s_p are as in the lines preceding (2.20). Then $s_{p+L} \geq t_{p+L} + 1 \geq \sigma_L \geq \sigma_{2m}$, with σ_j as in (2.25). This gives

$$\begin{aligned} &P\{H(z, L) \geq \lfloor L/2 \rfloor^2 | \mathcal{F}_z\} \\ &\geq P\{X_{.} \text{ visits some descendant } u \text{ of } v_{0,b+1} \text{ with} \\ &\qquad\qquad h(u) - h(v_{0,b+1}) = h(u) - 1 \geq m^2 \text{ during } [\sigma_0, \sigma_{2m}] | \mathcal{F}_{\sigma_0}\} \\ &\geq 1 - 2\exp[-K_3 m^{1/2}] \geq 1 - \exp[-K_7 L^{1/2}]. \end{aligned}$$

This is the required (2.39).

To obtain (2.40) we substitute (2.39) into (2.22). This, together with obvious monotonicity of γ, shows that

$$\begin{aligned} \beta(h(v), v) \quad &\leq \sup_n \text{ ess } \sup_\omega P\{L(n, j, k) \leq 4h^{1/2} \text{ for all } 1 \leq j \leq h-1, 2 \leq k \leq b | \mathcal{F}_n\} \\ &\qquad + [1 - \gamma(4h^{1/2}, h)] \\ &\leq \sup_n \text{ ess } \sup_\omega P\{L(n, j, k) \leq 4h^{1/2} \text{ for all } 1 \leq j \leq h-1, 2 \leq k \leq b | \mathcal{F}_n\} \\ &\qquad + \exp[-K_7 h^{1/4}]. \end{aligned}$$

Next let

$$\widehat{L}(n, j) = \text{number of } t \in (n, \tau(n, 0)) \text{ for which } X_t = v_j.$$

If $X_n = v, h(v) = h$ and $\widehat{L}(n, j) \leq 8(b + 2c)h^{1/2}$ for all $1 \leq j \leq h - 1$, then the imbedded random walk on the path r described just before Lemma 3 moves from $v = v_h$ to v_0 with no more than $\sum_{j=1}^{h-1} \widehat{L}(n, j) \leq 8(b+2c)h^{3/2}$ visits to $[1, h-1]$. The probability for this event is bounded by the probability that a symmetric simple random walk $\{S_n\}$ starting at 0 reaches h in no more than $8(b+2c)h^{3/2}$ steps from vertices in $[1, h - 1]$. We next deduce from standard random walk estimates that this probability is at most $\exp[-K_{10}h^{1/2}]$ for a suitable constant K_{10} and $h \geq$ some h_0. In fact, let $0 < n_1 < n_2 < \ldots < n_\lambda$ be all the indices n for which $S_n = 0$ before S_n first reaches h, at time τ say. Then $S_\tau - S_{n_\lambda} = h - 0 = h$, while $\tau - n_\lambda - 1$ cannot exceed the total time spent in $[1, h-1]$ before τ. Thus we must have $\tau - n_\lambda \leq 8(b + 2c)h^{3/2} + 1$. Consequently,

$$\begin{aligned}
&P\Big\{\sum_{j=1}^{h-1} \widehat{L}(n, j) \leq 8(b + 2c)h^{3/2}\Big\} \\
&\leq P\{\lambda > h^2\} \\
&\quad + \sum_{i=1}^{h^2} P\{S_\cdot \text{ reaches } h \text{ before time } n_i + 8(b + 2c)h^{3/2} + 1 | S_0, \ldots, S_{n_i}, n_i < \tau\} \\
&\leq \Big[1 - \frac{1}{h}\Big]^{h^2} + 4h^2 \max_{n \leq 8(b+2c)h^{3/2}+1} P\{S_n \geq h/4\} \leq \exp[-K_{10}h^{1/2}]
\end{aligned}$$

for large h (see Billingsley (1986), Theorem 22.5).

Finally, if $\widehat{L}(n, j) \geq 8(b + 2c)h^{1/2}$, but all $L(n, j, k) \leq 4h^{1/2}, 2 \leq k \leq b$, then in the first $8(b + 2c)h^{1/2}$ visits after time n to v_j, the walk $X_\cdot$ moves fewer than $4(b - 1)h^{1/2}$ times to one of the children $v_{j,2}, \ldots, v_{j,b}$ in the next step. Since at any visit by $X_\cdot$ to v_j the probability for $X_\cdot$ to move to one of these children in the next step is at least $(b - 1)/(b - 1 + 2c) \geq (b - 1)/(b + 2c)$,

$$\begin{aligned}
P\{\widehat{L}(n, j) \geq 8(b + 2c)h^{1/2} \quad &\text{but } L(n, j, k) \leq 4h^{1/2}, 2 \leq k \leq b | \mathcal{F}_n\} \\
&\leq \exp[-K_{11}h^{1/2}],\ 1 \leq j \leq h - 1,
\end{aligned}$$

(by standard exponential bounds for the binomial distribution; see Billingsley (1986), pp. 148, 149). From these estimates we obtain that

$$\beta(h(v), v) \leq \exp[-K_7h^{1/4}] + \exp[-K_{10}h^{1/2}] + h \exp[-K_{11}h^{1/2}]$$

for $h \geq$ some h_1. By adjusting K_8 this implies (2.40). □

Proof of Theorem 1. It now only takes a few lines to complete the proof of this theorem. Indeed, let $X_0 = 0,\ \tau(0) = 0$ and

$$\tau(j + 1) = \inf\{n > \tau(j) : h(X_n) = j + 1\}. \tag{2.41}$$

These are the hitting times of the various heights. $\tau(j)$ is the same as the $\tau(0, j)$ of Lemma 2. Analogously to the proof of (2.17),

$$
\begin{aligned}
&P\{h(X_.) \text{ reaches } h \text{ before } X_. \text{ returns to } v_0 | X_0 = v_0\} \\
&\geq \prod_{j=0}^{h-1} [\text{ess inf } P\{\tau(\tau(j), j+1) < \tau(\tau(j), 0) | \tau(j) < \text{first return time by } X_. \text{ to } v_0\}] \\
&\geq \prod_{j=0}^{\infty} [1 - \sup_{h(v)=j} \beta(j, v)] > 0, \qquad (2.42)
\end{aligned}
$$

by virtue of (2.40). Since this holds uniformly in h, $\{X_n\}$ cannot be recurrent.

3. Proof of Theorem 2

We call $t \geq 1$ a $cut - time$ if $h(X_n) < h(X_t) < h(X_m)$ for all $0 \leq n < t$ and for all $m > t$. It is easy to see from the fact that $\{X_n\}$ is transient that almost surely there exist infinitely many cut-times (see (2.6) and its proof). It is also well known that if the process starts with all weights

$$w(e, 0) = 1, \qquad (3.1)$$

then the cut-times are regeneration points. Specifically, if $0 < \mathcal{T}_1 < \mathcal{T}_2 < \dots$ are the successive cut-times, then the "excursions" $E_i := \{h(X_m) - h(X_{\mathcal{T}_i}) : \mathcal{T}_i \leq m \leq \mathcal{T}_{i+1}\}, i \geq 1$, are i.i.d. In particular, the vectors $\big(\mathcal{T}_{i+1} - \mathcal{T}_i, h(X_{\mathcal{T}_{i+1}}) - h(X_{\mathcal{T}_i})\big)$, $i \geq 1$, are i.i.d. In fact, these cut-times are splitting times for the Markov chain $\{X_n, w(e, n)\}_{n \geq 0}$ and the regenerative property can be obtained from Theorem 1 in Jacobsen (1974). Jacobsen ascribes this theorem to D. Williams; see also Kesten (1977) and Lyons, Pemantle and Peres (1996), Section 3, for regeneration points in circumstances very similar to the present one. It follows from the strong law of large numbers that

$$\lim_{i \to \infty} \frac{1}{i} \mathcal{T}_i = \lim_{i \to \infty} \frac{1}{i} \sum_{j=1}^{i-1} [\mathcal{T}_{j+1} - \mathcal{T}_j] = E\{\mathcal{T}_2 - \mathcal{T}_1\} \leq \infty \text{ a.s.}, \qquad (3.2)$$

and

$$\lim_{i \to \infty} \frac{1}{i} h(X_{\mathcal{T}_i}) = E\{h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})\} \leq \infty \text{ a.s.} \qquad (3.3)$$

In the next two lemmas we shall show that the expectations in the right hand sides of (3.2) and (3.3) are finite. Since $\mathcal{T}_2 - \mathcal{T}_1$ and $h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})$ are both at least 1, these expectations are also strictly positive. If $\mathcal{T}_i \leq n < \mathcal{T}_{i+1}$, then by the definition of the cut-times $\mathcal{T}_i$ we also have

$$h(X_{\mathcal{T}_i}) \leq h(X_n) < h(X_{\mathcal{T}_{i+1}}).$$

Therefore,

$$\limsup_{n\to\infty}\frac{1}{n}h(X_n)\le\lim_{i\to\infty}\frac{1}{\mathcal{T}_i}h(X_{\mathcal{T}_{i+1}})$$
$$=\lim_{i\to\infty}\frac{1}{i+1}h(X_{\mathcal{T}_{i+1}})\Big[\lim_{i\to\infty}\frac{1}{i}\mathcal{T}_i\Big]^{-1}$$
$$=\frac{E\{h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})\}}{E\{\mathcal{T}_2-\mathcal{T}_1\}}<\infty \text{ a.s..}$$

Together with a similar estimate for the liminf this will establish that

$$0<\lim_n\frac{1}{n}h(X_n)=\frac{E\{h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})\}}{E\{\mathcal{T}_2-\mathcal{T}_1\}}<\infty \text{ a.s.,}$$

and thus complete the proof of Theorem 2 with

$$S=\frac{E\{h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})\}}{E\{\mathcal{T}_2-\mathcal{T}_1\}}. \tag{3.4}$$

We note that for Theorem 2 we only need to know that $h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})$ and $\mathcal{T}_2-\mathcal{T}_1$ have a first moment. For Theorem 3 we also need second moments and for this reason we prove in Lemmas 7 and 8 that these two quantities have all moments. There exist much simpler proofs to establish the existence of the first moments only.

Lemma 7. *Under the conditions of Theorem 1, $h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})$ has all moments finite.*

Proof. It will be useful to introduce cut-levels, as spatial analogues of the cut-times. Let $\tau(j)$ be as in (2.41). We call ℓ a *cut-level* if $\tau(\ell)$ is a cut-time, or equivalently if

$$h(X_n)>h(X_{\tau(\ell)})=\ell \text{ for all } n>\tau(\ell). \tag{3.5}$$

(The inequalities

$$h(X_n)<h(X_{\tau(\ell)})=\ell \text{ for all } n<\tau(\ell)$$

are true by definition of $\tau(\ell)$.)

We shall bound $h(X_{\mathcal{T}_2})-h(X_{\mathcal{T}_1})$ by a sum of a random number of "almost" i.i.d. random variables. Essentially the same decomposition can be found in Piau (1998). Another closely related argument (corresponding to the situation without reinforcement) appears in Lemma 5.1 of Dembo, Peres and Zeitouni (1996). To define the appropriate random variables note that if $\mathcal{T}_1=t_1$ and $h(X_{\mathcal{T}_1})=h_1$, then $h(X_n)>h_1$ for all $n>t_1$ and t_1 is the first hitting time by $h(X_.)$ of the level h_1. In this case, the smallest possible value for the next cut-level, $h(X_{\mathcal{T}_2})$, is h_1+1. Moreover, h_1+1 is indeed the next cut-level if and only if $h(X_t)>h_1+1$ for all $t>t_1+1$. If this is not the case, then none of the levels reached between time t_1+1 and the next return by $h(X_t)$ to the level h_1+1 can be a cut-level. This suggests that the following random variables will be useful:

$$\widehat{s}_1=\inf\{t>t_1+1:h(X_t)=h_1+1\}\quad(=\infty \text{ if no such } t \text{ exists}),$$

$$\lambda_1 = \sup\{h(X_t) : t_1 + 1 \le t \le \widehat{s}_1\}.$$

As observed, if $\widehat{s}_1 = \infty$, then $\mathcal{T}_2 = \mathcal{T}_1 + 1$ and $h(X_{\mathcal{T}_2}) = h_1 + 1$. On the other hand, if $\widehat{s}_1 < \infty$, then none of the values in $[h_1 + 1, \lambda_1]$ are cut-levels and hence $h(X_{\mathcal{T}_2}) \ge \lambda_1 + 1$. On the event $\widehat{s}_1 < \infty$ we therefore define also

$$s_1 = \inf\{t > t_1 : h(X_t) = \lambda_1 + 1\} = \inf\{t > \widehat{s}_1 : h(X_t) = \lambda_1 + 1\},$$

and we take $s_1 = \infty$ on $\{\widehat{s}_1 = \infty\}$. Note that $s_1 < \infty$ a.e. on $\{\widehat{s}_1 < \infty\}$. (This is true both when the X-process is transient and when it is recurrent.) If $\widehat{s}_1, \dots, \widehat{s}_k, \lambda_1, \dots, \lambda_k, s_1, \dots, s_k$ have already been defined, and $s_k < \infty$, then we define similarly

$$\widehat{s}_{k+1} = \inf\{t > s_k : h(X_t) = h(X_{s_k}) = \lambda_k + 1\} \quad (= \infty \text{ if no such } t \text{ exists}),$$

$$\lambda_{k+1} = \sup\{h(X_t) : s_k + 1 \le t \le \widehat{s}_{k+1}\}$$

and

$$s_{k+1} = \inf\{t > s_k : h(X_t) = \lambda_{k+1} + 1\} = \inf\{t > \widehat{s}_{k+1} : h(X_t) = \lambda_{k+1} + 1\}$$

(see Figure 1). Note that if $s_{k+1} < \infty$, then

$$s_k < \widehat{s}_{k+1} < s_{k+1},$$

and that the process reaches a new level at the times s_k. Indeed,

$$s_k = \tau(s_{k-1}, \lambda_k + 1) \tag{3.6}$$

in the notation of (2.15). The times s_k are the potential cut-times. However, if $\widehat{s}_{k+1} < \infty$, then s_k turns out not to be a cut-time. Finally, we define

$$\kappa = \inf\{k : \widehat{s}_k = \infty\}.$$

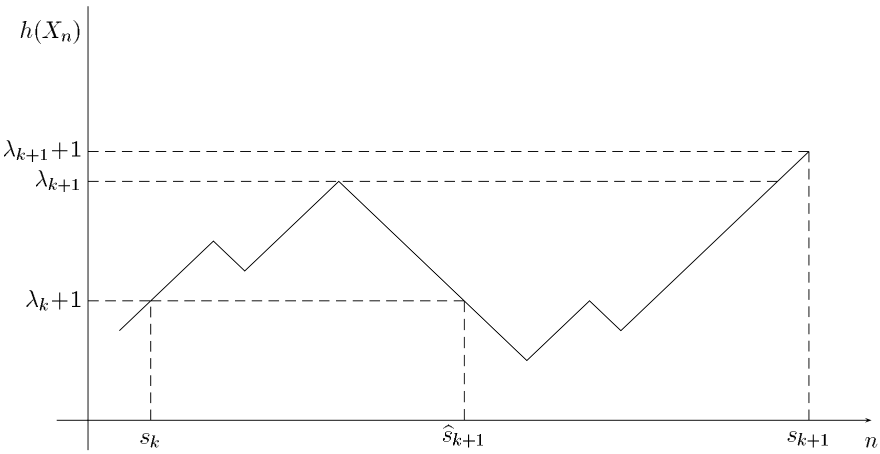

Fig. 1. Part of the sample path of $h(X_n)$ with $s_k, \widehat{s}_{k+1}$ and s_{k+1} marked. In this figure, s_k turns out not to be a cut-time.

Then $\widehat{s}_{\kappa-1} < \infty$ and $s_{\kappa-1} < \infty$ a.s., while $\widehat{s}_\kappa = \infty$. Thus $h(X_t) > \lambda_{\kappa-1} + 1$ for all $t > s_{\kappa-1} + 1$ and $\mathcal{T}_2 = s_{\kappa-1} + 1, h(X_{\mathcal{T}_2}) = \lambda_{\kappa-1} + 1$. Thus we have (with $\lambda_0 = h_1$)

$$h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1}) = 1 + \sum_{i=1}^{\kappa-1} [\lambda_i - \lambda_{i-1}]. \tag{3.7}$$

This is the desired decomposition for $h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})$.

We remind the reader that we already assumed that $\mathcal{T}_1 = t_1, h(X_{t_1}) = h_1$. We shall therefore be interested in conditional expectations given the event

$$\begin{aligned} C = C(h_1, t_1) := \{&h(X_n) < h_1 \text{ for } n < t_1, h(X_{t_1}) \\ &= h_1, \text{ and } h(X_t) > h_1 \text{ for } t > t_1\}. \end{aligned}$$

It is convenient to write this event as $C'_{i-1}(m) \cap C''_{i-1}(m)$, where $m \geq 1$ is an integer, and (with $s_0 = t_1 + 1$ and $\tau(n, j)$ as in (2.15))

$$\begin{aligned} C'_{i-1}(m) &:= \{h(X_n) < h_1 \text{ for } n < t_1, h(X_{t_1}) = h_1 \text{ and } h(X_t) \\ &\qquad > h_1 \text{ for } t_1 < t \leq \tau(t_1, h(X_{s_{i-1}} + m - 1))\}, \\ C''_{i-1}(m) &:= \{h(X_t) > h_1 \text{ for } t > \tau(t_1, h(X_{s_{i-1}} + m - 1))\}. \end{aligned}$$

$C'_{i-1}(m) \in \mathcal{F}_{\tau(t_1, h(X_{s_{i-1}} + m - 1))}$, whereas $C''_{i-1}(m)$ depends on the future after time $\tau(t_1, h(X_{s_{i-1}} + m - 1))$. An immediate consequence of (3.7) is that for $p > 1$

$$\begin{aligned} &\Big(E\{\big[h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})\big]^p | C\}\Big)^{1/p} \\ &\leq 1 + \sum_{i=1}^{\infty} \Big(E\{I[\kappa > i][\lambda_i - \lambda_{i-1}]^p | C\}\Big)^{1/p} \text{ (by Minkowski's inequality)} \\ &\leq 1 + \sum_{i=1}^{\infty} \Big(P\{\kappa > i - 1 | C\}\Big)^{1/p} \Big(E\{I[\kappa > i][\lambda_i - \lambda_{i-1}]^p | C, \kappa > i - 1\}\Big)^{1/p}. \end{aligned} \tag{3.8}$$

Next we estimate the tail of the conditional distribution of $I[\kappa > i] \cdot [\lambda_i - \lambda_{i-1}]^p$ given the event $D_{i-1} := C \cap \{\kappa > i - 1\} = C \cap \{\widehat{s}_{i-1} < \infty\}$. Now we have for any integer $m \geq 2$,

$$\begin{aligned} &P\{I[\kappa > i] \cdot [\lambda_i - \lambda_{i-1}] \geq m | D_{i-1}\} \\ &\quad = P\{h(X_t) \text{ returns to } h(X_{s_{i-1}}) \\ &\qquad \text{after reaching } h(X_{s_{i-1}}) + m - 1 | D_{i-1}\} \\ &\quad = 1 - P\{h(X_t) \text{ never returns to } h(X_{s_{i-1}}) \\ &\qquad \text{after reaching } h(X_{s_{i-1}}) + m - 1 | \kappa > i - 1, C'_{i-1}(m), C''_{i-1}(m)\} \\ &\quad \leq 1 - P\{C''_{i-1}(m),\ h(X_t) > h(X_{s_{i-1}}) \text{ for all} \\ &\qquad t \geq \tau\big(t_1, h(X_{s_{i-1}})\big) + m - 1 | \kappa > i - 1, C'_{i-1}(m)\}. \end{aligned} \tag{3.9}$$

We now apply the argument in (2.42) to the tree $T(X_{s_{i-1}})$. In (3.9) we have to find a lower bound for the probability that we do not return to the root of this tree once we reached height $m-1$ above this root. Essentially as in (2.42) we have for large m

$$\begin{aligned}
&P\{C''_{i-1}(m),\ h(X_t) > h(X_{s_{i-1}}) \text{ for all } t \geq \tau\big(t_1, h(X_{s_{i-1}}) + m - 1\big)| \\
&\qquad \kappa > i-1, C'_{i-1}(m)\} \\
&\quad \geq \prod_{j=m-1}^{\infty} [1 - \sup_{h(v)=j} \beta(h(v), v)] \\
&\quad \geq \prod_{j=m-1}^{\infty} \Big[1 - K_8 \exp[-K_9 j^{1/4}]\Big] \text{ (see (2.40))} \\
&\quad \geq \exp\Big[-2K_8 \sum_{j=m-1}^{\infty} \exp[-K_9 j^{1/4}]\Big] \\
&\quad \geq \exp\Big[-2K_8 \int_{m-2}^{\infty} \exp[-K_9 x^{1/4}]dx\Big] \\
&\quad \geq \exp\Big[-K_{12} m^{3/4} \exp[-K_9 m^{1/4}]\Big] \\
&\quad \geq 1 - K_{12} m^{3/4} \exp[-K_9 m^{1/4}]. \qquad (3.10)
\end{aligned}$$

Substitution of this estimate into (3.9) yields

$$P\{I[\kappa > i] \cdot [\lambda_i - \lambda_{i-1}] \geq m | D_{i-1}\} \leq K_{12} m^{3/4} \exp[-K_9 m^{1/4}].$$

In turn, this implies

$$\begin{aligned}
&E\{I[\kappa > i] \cdot [\lambda_i - \lambda_{i-1}]^p \big| D_{i-1}\} \\
&\qquad \leq K_{13} \sum_{m=1}^{\infty} m^{p-1} P\{I[\kappa > i] \cdot [\lambda_i - \lambda_{i-1}] \geq m | D_{i-1}\} \leq K_{14},
\end{aligned}$$

for a suitable constant $K_{14} = K_{14}(p) < \infty$. It follows that the right hand side of (3.8) is bounded by

$$1 + K_{14}^{1/p} \sum_{i=1}^{\infty} \Big(P\{\kappa > i-1 | C\}\Big)^{1/p}. \qquad (3.11)$$

Finally we show that the sum here is finite. To see this we show that κ is stochastically smaller than a geometric random variable. Indeed, if we take into account that $\{s_{i-1} < \infty\}$ and $\{\widehat{s}_{i-1} < \infty\}$ differ only by a set of probability 0, and the fact that $h(X_{s_{i-1}}) = \lambda_{i-1} + 1$ is reached for the first time at time s_{i-1}, so that $\tau\big(t_1, h(X_{s_{i-1}})\big)$ we obtain

$$\begin{aligned}
&P\{\kappa = i | \kappa > i-1, C\} \\
&\quad = P\{\kappa = i | \widehat{s}_{i-1} < \infty, C\} \\
&\quad = P\{h(X_t) > h(X_{s_{i-1}}) \text{ for all } t > s_{i-1} | s_{i-1} < \infty, C'_{i-1}(1), C''_{i-1}(1)\} \\
&\quad \geq P\{C''_{i-1}(1),\ h(X_t) > h(X_{s_{i-1}}) \text{ for all } t > s_{i-1} | s_{i-1} < \infty, C'_{i-1}(1)\} \\
&\quad \geq \frac{b}{b+c} P\{X_n \text{ never returns to } v_0 | A_0\} > 0 \text{ (by Lemma 1).} \qquad (3.12)
\end{aligned}$$

Thus, if we denote the last member of (3.12) by δ, then

$$P\{\kappa > i|C\} \le (1-\delta)P\{\kappa > i-1|C\} \le \ldots \le (1-\delta)^i.$$

Together with (3.8) and (3.11) this proves the lemma. □

Lemma 8. *Under the conditions of Theorem 1, $\mathcal{T}_2 - \mathcal{T}_1$ has all moments finite.*

Proof. We shall reduce this Lemma to the preceding one. We first want a lower bound for

$$P\{h(X_.) \text{ reaches } h(v)+q \text{ before it reaches } h(v)-q|\mathcal{F}_n\}$$

on the set $\{X_n = v\}$. As before, let $r = (v_0, \ldots, v_h)$ be the self-avoiding path from the root to v. Until $h(X_.)$ reaches $h(v)-q$, $X_.$ has to stay inside $T(v_{h-q})$. This time we apply the argument for (2.42) to the tree $T(v_{h-q})$ which is rooted at v_{h-q}. We then obtain (with the help of (2.40) for the last step)

$$\begin{aligned}
&P\{h(X_.) \text{ reaches } h(v)+q \text{ before it reaches } h(v)-q|\mathcal{F}_n\}\\
&\quad\ge \inf_m \operatorname*{ess\,inf}_{\omega: h(X_m(\omega))=q} P\{h(X_.) \text{ reaches } 2q \text{ in } T(v_{h-q}) \text{before it reaches } v_{h-q}|\mathcal{F}_m\}\\
&\quad\ge \inf_m \prod_{j=q}^{2q-1} [\operatorname{ess\,inf} P\{\tau(\tau(m,j),j+1) < \tau(\tau(m,j),0)|\tau(m,j) < \tau(m,0)\}]\\
&\quad\ge \prod_{j=q}^{2q-1} [1 - \sup_{h(v)=j} \beta(h(v),v)]\\
&\quad\ge \prod_{j=q}^{2q-1} [1 - K_8 \exp[-K_9 j^{1/4}].
\end{aligned}$$

For the remainder of this proof we now fix q such that

$$P\{h(X_.) \text{ reaches } h(v)+q \text{ before it reaches } h(v)-q|\mathcal{F}_n\} \ge \frac{3}{4} \text{ a.s. on } \{X_n = v\} \tag{3.13}$$

for all n and v.

Next we consider the X-process at the times when $h(X_.)$ is a multiple of q. By (3.13) this is bounded below by a random walk with positive drift, and this will be crucial for our argument. Here are the details. Define $\nu(0) = 0$ and

$$\nu(j+1) = \inf\{n > \nu(j)\big||h(X_n) - h(X_{\nu(j)})| = q\}.$$

Also set

$$y_j = h(X_{\nu(j)}).$$

Then (3.13) implies

$$P\{y_{j+1} = y_j + q|\mathcal{F}_{\nu(j)}\} = 1 - P\{y_{j+1} = y_j - q|\mathcal{F}_{\nu(j)}\} \ge \frac{3}{4}. \tag{3.14}$$

Let

$$j^* = \inf\{j : \nu(j) \geq \mathcal{T}_1\}.$$

Thus $\nu(j^*)$ is the smallest $\nu(j)$ greater than or equal to $\mathcal{T}_1$. From the definition of the cut-times it follows that $h(X_n) > h(X_{\mathcal{T}_\ell})$ implies $n > \mathcal{T}_\ell$. We claim that this implies for any choice of the constants $K_{15}, K_{16} \in (0, \infty)$ and any $m \geq 1$ that

$$\begin{aligned} P\{\mathcal{T}_2 - \mathcal{T}_1 \geq m\} \leq\ & P\{h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1}) \geq K_{15}m\} \\ & + P\{y_{j^*+\lfloor K_{16}m\rfloor} < h(X_{\mathcal{T}_1}) + K_{15}m\} \\ & + P\{\nu(j^*) - \mathcal{T}_1 > m/2\} \\ & + P\{\nu(j^* + \lfloor K_{16}m\rfloor) - \nu(j^*) > m/2\}. \end{aligned} \tag{3.15}$$

To see this we note that if none of the four events on the right hand side of (3.15) occur, then

$$h(X_{\mathcal{T}_2}) < h(X_{\mathcal{T}_1}) + K_{15}m \leq y_{j^*+\lfloor K_{16}m\rfloor} = h\big(X_{\nu(j^*+\lfloor K_{16}m\rfloor)}\big),$$

and hence

$$\nu(j^* + \lfloor K_{16}m\rfloor) > \mathcal{T}_2.$$

But also

$$\nu(j^* + \lfloor K_{16}m\rfloor) \leq \mathcal{T}_1 + m,$$

so that

$$\mathcal{T}_2 < \mathcal{T}_1 + m.$$

Thus (3.15) holds. It therefore suffices for this lemma to prove that for any $p > 0$ and a suitable choice of K_{15}, K_{16}

$$\sum_{m=1}^{\infty} m^{p-1} P\{h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1}) \geq K_{15}m\} < \infty, \tag{3.16}$$

$$\sum_{m=1}^{\infty} m^{p-1} P\{y_{j^*+\lfloor K_{16}m\rfloor} < h(X_{\mathcal{T}_1}) + K_{15}m\} < \infty, \tag{3.17}$$

$$\sum_{m=1}^{\infty} m^{p-1} P\{\nu(j^*) - \mathcal{T}_1 > m/2\} < \infty \tag{3.18}$$

and

$$\sum_{m=1}^{\infty} m^{p-1} P\{\nu(j^* + \lfloor K_{16}m\rfloor) - \nu(j^*) > m/2\} < \infty. \tag{3.19}$$

Clearly, (3.16) is equivalent to $E\{[h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})]^p\} < \infty$, and this holds by virtue of Lemma 7, no matter what the value of $K_{15} \in (0, \infty)$ is. (3.17) follows from the fact that

$$y_{j^*+\lfloor K_{16}m\rfloor} - h(X_{\mathcal{T}_1}) \geq \sum_{i=j^*}^{j^*+\lfloor K_{16}m\rfloor-1} [y_{i+1} - y_i]$$

and (3.14). Indeed (3.14) implies that $\{y_j/q\}$ stochastically dominates a nearest neighbor random walk which moves one step to the right (left) with probability 3/4 (respectively, 1/4) and

$$E\{y_{j+1} - y_j | \mathcal{F}_{\nu(j)}\} \geq \frac{q}{2}.$$

Standard exponential bounds for random walk (see Billingsley (1986), Theorem 9.3) then show that

$$P\left\{ \sum_{i=j^*}^{j^*+\lfloor K_{16}m \rfloor -1} [y_{i+1} - y_i] < K_{15}m \right\}$$

tends to zero exponentially in m as soon as

$$K_{16}\frac{q}{2} > K_{15}. \tag{3.20}$$

To prove that (3.18) and (3.19) hold we note that for any n, on $\{\nu(j) \leq n\}$ it holds that $P\{\nu(j+1) \leq n + q|\mathcal{F}_n\} \geq K_{17} > 0$. Therefore,

$$P\{\nu(j+1) - \nu(j) \geq m | \mathcal{F}_{\nu(j)}\} \leq K_{18}e^{-K_{19}m},$$
$$P\{\nu(j^*) - \mathcal{T}_1 \geq m | \mathcal{F}_{\mathcal{T}_1}\} \leq K_{18}e^{-K_{19}m} \tag{3.21}$$

and

$$E\{\nu(j+1) - \nu(j) | \mathcal{F}_{\nu(j)}\} \leq K_{20},$$

for suitable constants $0 < K_i < \infty$. (3.18) is now immediate from (3.21). Finally, (3.19) also holds by standard large deviation estimates as soon as

$$K_{16}K_{20} < \frac{1}{2},$$

because $\{\nu(j)\}$ is stochastically dominated by a random walk whose increments have exponential tails, and which have mean at most K_{20} (see Billingsley (1986), Theorem 9.3). Clearly we can choose first K_{16} and then K_{15} so as to satisfy this and (3.20). □

As we pointed out before, Theorem 2 follows directly from Lemmas 7 and 8.

Proof of Remark 2 As in (2.41), let

$$\tau(k) = \inf\{n : h(X_n) = k\}$$

and

$$\begin{aligned} N(k) &= \{\text{number of times } n < \tau(k+1) \text{ with } h(X_n) = k\} \\ &= \{\text{number of times } n \in [\tau(k), \tau(k+1)) \text{ with } h(X_n) = k\}. \end{aligned}$$

The second equality here holds because $h(X_n) = k$ does not occur before time $\tau(k)$, by definition. Clearly,

$$\tau(k) \geq \sum_{i=1}^{k-1} N(i)$$

and

$$S \le \liminf_{k\to\infty} \frac{h(X_{\tau(k)})}{\tau(k)} = \liminf_{k\to\infty} \frac{k}{\tau(k)}.$$

Thus, for any constant K, we have

$$\frac{1}{S} \ge \limsup_{k\to\infty} \frac{\tau(k)}{k} \ge \limsup_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k-1}\{N(i)\wedge K\}.$$

Taking expected values, using the fact that S is almost surely constant, and applying Fatou's lemma to $K - \limsup_{k\to\infty} k^{-1}\sum_{i=1}^{k-1}\{N(i)\wedge K\}$, gives

$$\frac{1}{S} = E\Big\{\frac{1}{S}\Big\} \ge \limsup_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k-1} E\{N(i)\wedge K\}. \tag{3.22}$$

Now, at any time $n < \tau(i+1)$, the X-process has not passed from height i to height $i+1$, and therefore the edges between any vertex x with $h(x) = i$ and its children have not yet been reinforced. On the other hand, if $X_n = x$ for some x with $h(x) = i$, then the edge between the parent of x and x has been traversed at least once, and this edge therefore has weight c at time n. It follows that conditionally on $\mathcal{F}_n$, on the event $\{n < \tau(i+1), X_n = x\}$ with $h(x) = i$ the probability for $h(X_{n+1}) = i-1$ equals $c/(b+c)$. If $h(X_{n+1}) = h(X_n) - 1 = i - 1$, then $h(X_{.})$ will eventually return to the level i, and then there is again a conditional probability of $c/(b+c)$ that the height will go down to $i-1$ in the next step. This implies that

$$P\{N(i) \ge j|\mathcal{F}_{\tau(i)}\} = \big[c/(b+c)\big]^{j-1},$$

and

$$E\{N(i)\wedge K\} \uparrow \frac{b+c}{b} \text{ as } K \uparrow \infty.$$

Together with (3.22) this implies (1.4).

4. Proof of Theorem 3

Since we have proven the existence of the appropriate moments in Lemmas 7 and 8, Theorem 3 follows from standard arguments for regenerative processes (see for instance S. Asmussen (1987), Section V.3 and also the proof of the central limit theorem for Markov chains in Chung (1967), Section I.16). We merely give an indication of the proof.

We define $\varphi(x)$ for any $x \ge 0$ as the unique index for which

$$\mathcal{T}_{\varphi(x)-1} < x \le \mathcal{T}_{\varphi(x)}.$$

The law of large numbers (3.2) shows that

$$\lim_{n\to\infty} \frac{1}{n}\sup_{t\le 1}\Big|\varphi(nt) - \frac{nt}{E\{\mathcal{T}_2 - \mathcal{T}_1\}}\Big| = 0 \text{ a.s.} \tag{3.23}$$

Furthermore, since $\mathcal{T}_2 - \mathcal{T}_1$ has a finite second moment,

$$\frac{1}{\sqrt{m}}\big|\mathcal{T}_{\varphi(m)} - \mathcal{T}_{\varphi(m-1)}\big| \to 0 \text{ a.s.}$$

Consequently also

$$\frac{1}{\sqrt{n}} \sup_{t\le 1} |\mathcal{T}_{\varphi(nt)} - nt| \to 0 \text{ a.s.} \tag{3.24}$$

In view of the fact that

$$h(X_{\mathcal{T}_{\varphi(m)-1}}) \le h(X_m) \le h(X_{\mathcal{T}_{\varphi(m)}}),$$

a similar argument shows that also

$$\frac{1}{\sqrt{n}} \sup_{t\le 1} |h(X_{\mathcal{T}_{\varphi(nt)}}) - h(X_{nt})| \to 0 \text{ a.s.} \tag{3.25}$$

(3.23)–(3.25) combined show that for any $\delta > 0$, $\varepsilon > 0$,

$$\limsup_{n\to\infty} P\left\{ \sup_{\substack{0\le t',t''\le 1\\ 0\le t'-t''\le\delta}} \frac{1}{\sqrt{n}}|h(X_{nt'}) - nt'S - h(X_{nt''}) + nt''S| > \varepsilon \right\}$$

$$\le \limsup_{n\to\infty} P\left\{ \sup_{\substack{0\le k',k''\le 2n/E\{\mathcal{T}_2-\mathcal{T}_1\}\\ 0\le k'-k''\le 2\delta n/E\{\mathcal{T}_2-\mathcal{T}_1\}}} \frac{1}{\sqrt{n}}\big|h(X_{\mathcal{T}_{k'}}) - \mathcal{T}_{k'}S - h(X_{\mathcal{T}_{k''}}) + \mathcal{T}_{k''}S\big| > \varepsilon \right\}. \tag{3.26}$$

Now define $\mathcal{T}_0 = 0$ and for $k \ge 0$

$$W_k = h(X_{\mathcal{T}_{k+1}}) - h(X_{\mathcal{T}_k}) - (\mathcal{T}_{k+1} - \mathcal{T}_k)S. \tag{3.27}$$

Then the right hand side of (3.26) can be written as

$$\limsup_{n\to\infty} P\left\{ \sup_{\substack{0\le k',k''\le 2n/E\{\mathcal{T}_2-\mathcal{T}_1\}\\ 0\le k'-k''\le 2\delta n/E\{\mathcal{T}_2-\mathcal{T}_1\}}} \frac{1}{\sqrt{n}}\Big| \sum_{i=k'}^{k''-1} W_i \Big| > \varepsilon \right\}. \tag{3.28}$$

The regenerative property of the cut-times says that $W_1, W_2, \ldots$ are i.i.d. A simple calculation and (3.4) show they have mean 0. Moreover, $EW_k^2 < \infty$ by virtue of Lemmas 7 and 8. Donsker's theorem (see Billingsley (1968), Section 2.10) tells us that the limit of (3.28) as $\delta \downarrow 0$ equals 0, and the same holds for the left hand side of (3.26). In turn this means that the family of functions $\{h(X_{nt}) - ntS\}_{0\le t\le 1}$ is tight in $C([0, 1], \mathbb{R})$. Thus to complete the proof of Theorem 3 it suffices to show that the finite dimensional distributions of $\{h(X_{nt}) - ntS\}_{0\le t\le 1}$ converge to those of $\{\sigma B(t)\}_{0\le t\le 1}$ with

$$\sigma^2 = \frac{1}{E\{\mathcal{T}_2 - \mathcal{T}_1\}} \text{Variance}(W_1), \tag{3.29}$$

and that $\sigma^2 > 0$.

The convergence of the finite dimensional distributions is immediate from the central limit theorem or Donsker's theorem and the fact that

$$\frac{1}{\sqrt{n}}\Big[h(X_{nt}) - ntS - \sum_{0\le i\le nt/E\{\mathcal{T}_2-\mathcal{T}_1\}} W_i\Big] \to 0 \text{ in probability} \tag{3.30}$$

(see (3.23)–(3.25) and use the fact that (3.26) tends to 0 as $\delta \downarrow 0$). Finally $\sigma^2 > 0$ follows easily from the fact that the variance of W_1 is strictly positive. Indeed, $\mathcal{T}_2 - \mathcal{T}_1$ still can take arbitrarily large values even when the difference of the two successive cut-levels, $h(X_{\mathcal{T}_2}) - h(X_{\mathcal{T}_1})$, is fixed; one merely has to let the X-process move back and forth over the same edge a large number of times.

Acknowledgements. The research of R.D. and H.K. was supported by the NSF through grants DMS 9877066 and DMS 9970943, respectively. The research of V.L. was supported by an NSF Postdoctoral Fellowship.

References

Asmussen, S.: *Applied Probability and Queues*, John Wiley & Sons 1987

Billingsley, P.: *Convergence of Probability Measures*, John Wiley & Sons 1968

Billingsley, P.: *Probability and Measure*, 2nd ed., John Wiley & Sons 1986

Breiman, L.: *Probability*, Addison-Wesley 1968

Chung, K.L.: *Markov Chains with Stationary Transition Probabilities*, 2nd ed., Springer-Verlag 1967

Coppersmith, D., Diaconis, P.: *Reinforced random walk*, Unpublished manuscript (1987)

Davis, B.: *Reinforced random walk*, Probab. Th. Rel. Fields **84**, 203–229 (1990)

Dembo, A., Peres, Y., Zeitouni, O.: *Tail estimates for one-dimensional random walk in random environment*, Comm. Math. Phys. **181**, 667–683 (1996)

Doyle, P.G., Snell, J.L.: *Random Walks and Electric Networks*, Carus Math. Monograph Nr 22 Math. Assoc. of America (1984)

Feller, W.: *Introduction to Probability Theory and its Applications*, vol I 3rd ed., John Wiley & Sons 1968

Jacobsen, M.: *Splitting times for Markov processes and a generalised Markov property for diffusions*, Z. Wahrsch. verw. Geb. **30**, 27–43 (1974)

Kesten, H.: *A renewal theorem for random walk in a random environment*, Proc. Symposia Pure Math. **31**, 67–77 (1977)

Lyons, R., Pemantle, R., Peres, Y.: *Biased random walks on Galton-Watson trees*, Probab. Th. Rel. Fields **106**, 249–264 (1996)

Lyons, R., Peres, Y.: *Probability on Trees and Networks*, available at http://php.indiana.edu/rdlyons/prbtree/prbtree.html (1997)

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

Piau, D.: *Théorème central limite fonctionnel pour une marche au hasard en environnement aléatoire*, Ann. Probab. **26**, 1016–1040 (1998)