Probab. Theory Relat. Fields (2014) 160:655–678
DOI 10.1007/s00440-013-0537-5

Large deviations in the reinforced random walk model on trees

Yu Zhang

Received: 14 October 2012 / Revised: 9 July 2013 / Published online: 30 October 2013

Abstract In this paper, we consider the linearly reinforced and the once-reinforced random walk models in the transient phase on trees. We show the large deviations for the upper tails for both models. We also show the exponential decay for the lower tail in the once-reinforced random walk model. However, the lower tail is in polynomial decay for the linearly reinforced random walk model.

Keywords Reinforced random walks on trees · Large deviation

Mathematics Subject Classification 60K37 · 60J15

1 Introduction

Let $\mathbf{T}$ be an infinite tree with vertex set $\mathbf{V}$. Each $v \in \mathbf{V}$ has $b+1$ neighbors except a vertex, called the *root*, which has b neighbors for $b \geq 2$. We denote the root by $\mathbf{0}$. For any two vertices $u, v \in \mathbf{V}$, let $e = [u, v]$ be the edge with vertices u and v. We denote by $\mathbf{E}$ the edge set. Consider a Markov chain $\mathbf{X} = \{X_i, \omega(e, i)\}$, which starts at $X_0 = \mathbf{0}$ with $\omega(e, 0) = 1$ for all $e \in \mathbf{E}$, where $\omega(e, 0)$ is called the *initial weight*. For $i \geq 1$ and $e \in \mathbf{E}$, let $X_i \in \mathbf{V}$ and let $\omega(e, i) \geq 1$ be the i-th *weight*. The transition from X_i to the nearest neighbor X_{i+1} is randomly selected with probabilities proportional to weights $\omega(e, i)$ of incident edges e to X_i.

After X_i has changed to X_{i+1}, the weights are updated by the following rule:

$$w(e, i+1) = \begin{cases} 1 + k(c-1) & \text{for } [X_i, X_{i+1}] = e \text{ and } e \text{ had been traversed } k \text{ times,} \\ w(e, i) & \text{otherwise} \end{cases}$$

Y. Zhang (✉)
Department of Mathematics, University of Colorado, Colorado Springs, USA
e-mail: yzhang3@uccs.edu

for fixed $c > 1$. With this weight change, the model is called a *linearly reinforced random walk*. Note that if $c = 1$, then it is a simple random walk.

The linearly reinforced random walk model was first studied by Coppersmith and Diaconis in 1986 (see [4]) for finite graphs on the $\mathbf{Z}^d$ lattice. They asked whether the walks are recurrent or transient. For $d = 1$, the walks are recurrent for all $c \geq 1$ (see [3] and [10]). For $d \geq 1$, Sabot and Tarres [9] showed that the walks are also recurrent for a large c. The other cases on the $\mathbf{Z}^d$ lattice still remain open. Pemantle [8] studied this model on trees and showed that there exists $c_0 = c_0(b) \geq 4.29$ such that when $1 < c < c_0$, then the walks are transient and when $c > c_0$, then the walks are recurrent. Furthermore, Collevecchio [2] and Aidekon [1] investigated the behavior of $h(X_n)$ on the transient phase, where $h(x)$ denotes by the number of edges from the root to x for $x \in \mathbf{T}$. They focused on $c = 2$ and showed that the law of large numbers holds for $h(X_n)$ with a positive speed for any $b \geq 2$. More precisely, if $c = 2$, then there exists $0 < T = T(b) < b/(b+2)$ such that

$$\lim_{n\to\infty} \frac{h(X_n)}{n} = T \text{ a.s..} \tag{1.1}$$

By the dominated convergence theorem,

$$\lim_{n\to\infty} \mathbf{E}\frac{h(X_n)}{n} = T. \tag{1.2}$$

By a simple computation, the probability that the walks repeatedly move between an edge connected to the root is larger than n^{-C} for some $C = C(b) > 0$. Therefore,

$$n^{-C} \leq \mathbf{P}(h(X_n) \leq 1), \tag{1.3}$$

so the lower tail of $h(X_n)$ has the following behavior:

$$n^{-C} \leq \mathbf{P}(h(X_n) \leq n(T - \epsilon)) \tag{1.4}$$

for all $\epsilon < T$ and for all large n. In this paper, C and C_i are positive constants depending on c, b, ϵ, N, M, and δ, but not on n, m, and k. They also change from appearance to appearance. From (1.4), unlike a simple random walk on a tree, we have

$$\lim_{n\to\infty} \frac{-1}{n^\eta} \log \mathbf{P}(h(X_n) \leq n(T - \epsilon)) = 0 \tag{1.5}$$

for all $\epsilon < T$ and for all $\eta > 0$.

We may ask what the behavior of the upper tail is. Unlike the lower tail, we show that the upper tail has a standard large deviation behavior for large b.

Theorem 1 *For the linearly reinforced random walk model with $c = 2$ and $b \geq 70$, and for $\epsilon > 0$, there exists a positive number $\alpha = \alpha(b, \epsilon)$ such that*

$$\lim \frac{-1}{n} \log \mathbf{P}(h(X_n) \geq (T + \epsilon)n) = \alpha.$$

Remark 1 The proof of Theorem 1 depends on a few Collevecchio's estimates (see Lemma 2.1 as follows). Since his estimates need a requirement that $b \geq 70$, Theorem 1 also needs this restriction. We conjecture that Theorem 1 holds for all $b \geq 2$. Durrett et al. [5] also investigated a similar reinforced random walk $\{Y_k, w(e, i)\}$, except that the weight changes by

$$w(e, i+1) = \begin{cases} c & \text{for } [Y_i, Y_{i+1}] = e, \\ w(e, i) & \text{otherwise} \end{cases} \tag{1.6}$$

for fixed $c > 1$. This random walk model is called a *once-reinforced random walk*. For the once-reinforced random walk model, Durrett et al. [5] showed that for any $c > 1$, the walks are always transient. In addition, they also showed the law of large numbers for $h(Y_n)$. More precisely, they showed that there exists $0 < S = S(c) < b/(b+c)$ such that

$$\lim_{n\to\infty} \frac{h(Y_n)}{n} = S \text{ a.s..} \tag{1.7}$$

We also investigate the large deviations for $h(Y_n)$. We have the following theorem, similar to the linearly reinforced random walk model.

Theorem 2 *For the once-reinforced random walk model with $c > 1$ and for $\epsilon > 0$, there exists a finite positive number $\beta = \beta(c, b, \epsilon)$ such that*

$$\lim \frac{-1}{n} \log \mathbf{P}(h(Y_n) \geq (S+\epsilon)n) = \beta.$$

Remark 2 It is difficult to compute the precise rate functions α and β. But we may obtain some properties such as the continuity in ϵ for them.

We may ask what the lower tail deviation for $h(Y_n)$ is. Unlike in the linearly reinforced random walk model, the lower tail is still exponentially decaying.

Theorem 3 *For the once-reinforced random walk model with $c > 1$ and $0 < \epsilon < S$,*

$$0 < \liminf \frac{-1}{n} \log \mathbf{P}(h(Y_n) \leq (S-\epsilon)n) \leq \limsup \frac{-1}{n} \log \mathbf{P}(h(Y_n) \leq (S-\epsilon)n) < \infty.$$

Remark 3 Durrett et al. [5] also showed that (1.7) holds for a finitely many times reinforced random walk. We can also adopt the same proof of Theorems 2 and 3 to show that the same arguments hold for a finitely many times reinforced random walk. In fact, our proofs in Theorems 2 and 3 depend on Durrett, Kesten, and Limic's Lemmas 7 and 8 (2002). These proofs in their lemmas can be extended for the finitely many times reinforced random walk model.

Remark 4 We believe that the limit exists in Theorem 3, but we are unable to show it.

2 Preliminaries

In this section, we focus on the linearly reinforced random walk model with $c = 2$. We define a *hitting time* sequence $\{t_i\}$ as follows.

$$t_k = \min\{j \geq 0 : h(X_j) = k\}.$$

Note that walks are transient, so $h(X_j) \to \infty$ as $j \to \infty$. Thus, t_k is finite and

$$0 = t_0 < t_1 < t_2 < \cdots < t_k < \cdots < \infty. \tag{2.1}$$

With this definition, for each $k \geq 1$,

$$h(X_{t_k}) - h(X_{t_{k-1}}) = 1. \tag{2.2}$$

We also define a *leaving time* sequence $\{\rho_i\}$ as follows.

$$\rho_i = \max\{j \geq 0 : h(X_j) = i\}.$$

Since the walk $\mathbf{X}$ is transient,

$$\rho_0 < \rho_1 < \cdots < \rho_k < \cdots < \infty. \tag{2.3}$$

However, unlike the simple random walk model, $\{t_j - t_{j-1}\}$ are not independent increments. So we need to look for independence from these times. To achieve this target, we call t_i a *cut time* if

$$\rho_i - t_i = 0. \tag{2.4}$$

Since the walks $\mathbf{X}$ is transient, we may select these cut times and list all of them in increasing order as

$$\tau_1 < \cdots < \tau_k < \cdots < \infty. \tag{2.5}$$

With these cutting times, we consider difference

$$H_k = h(X_{\tau_{k+1}}) - h(X_{\tau_k}) \quad \text{for } k = 1, 2, \ldots. \tag{2.6}$$

By this definition, it can be shown that for $k = 1, 2, \ldots,$

$$(\tau_{k+1} - \tau_k, H_k) \text{ is an i.i.d. sequence.} \tag{2.7}$$

In fact (see page 97 in [2]), to verify (2.7), it is enough to realize that X_{τ_k}, $k \geq 1$, are regenerative points for the process $\mathbf{X}$. These points split the process $\mathbf{X}$ into *i.i.d.* pieces, which are $\{X_m, \tau_k \leq m < \tau_{k+1}\}$, $k \geq 1$.

Level $k \geq 1$ is the set of vertices v such that $h(v) = k$. Level k is a *cut* level if the walk visits it only once. We also call X_k, the only vertex to be visited, the *cut vertex*. It follows from the cut time definition that X_{τ_k} is a cut vertex for $k \geq 1$. We want to remark that τ_1 may or may not be equal zero. If $\tau_1 = 0$, the root is a cut vertex. For convenience, we just call $\tau_0 = 0$ whether the root is a cut vertex or not. In addition, let

$$H_0 = h(X_{\tau_1}) - h(X_{\tau_0}) = h(X_{\tau_1}). \tag{2.8}$$

With these definitions, Collevecchio [2] proved the following lemma.

Lemma 2.1 *For* $c = 2$ *and* $b \geq 70$,

$$\mathbf{P}(H_k \geq k) \leq 0.115^k \quad \textit{for } k \geq 0. \tag{2.9}$$

Furthermore, for $p_0 = 1002/1001$,

$$\mathbf{E}\tau_1^{p_0} < \infty. \tag{2.10}$$

With Lemma 2.1, we can see that $h(X_{\tau_{k+1}}) - h(X_{\tau_k})$ is large with a small probability. Also, $\tau_{k+1} - \tau_k$ is large with a small probability. However, to show a large deviation result, we need a much shorter tail requirement. Therefore, we need to truncate both $H_k = h(X_{\tau_{k+1}}) - h(X_{\tau_k})$ and $\tau_{k+1} - \tau_k$. We call τ_k N-*short* for $k \geq 1$ if

$$H_k = h(X_{\tau_{k+1}}) - h(X_{\tau_k}) \leq N; \tag{2.11}$$

otherwise, we call it N-*long*. Since we only focus on the transient phase, we have

$$\tau_k(N) < \infty.$$

We list all N-short cut times as

$$\tau_1(N) < \tau_1(N) < \cdots < \infty. \tag{2.12}$$

For convenience, we also call $\tau_0(N) = 0$ whether the root is a cut vertex or not. We know that $\tau_k(N) = \tau_i$ for some i.We denote it by $\tau'_k(N) = \tau_{i+1}$. In particular, let $\tau'_0(N) = 0$. For $N > 0$, let

$$I_n = \max\{i : \tau_i(N) \leq n\}$$

and

$$h_n(N) = \sum_{i=0}^{I_n} \left(h\left(X_{\tau'_i(N)}\right) - h\left(X_{\tau_i(N)}\right)\right).$$

If $I_n = 0$,

$$h_n = 0. \tag{2.13}$$

Now we state standard tail estimates for an i.i.d. sequence. The proof can be followed directly from Markov's inequality.

Lemma 2.2 *Let* $Z_1, \dots Z_k, \dots$ *be an i.i.d. sequence with* $\mathbf{E}Z_1 = 0$ *and* $\mathbf{E}\exp(\theta Z_1) < \infty$ *for some* $\theta > 0$, *and let*

$$S_m = Z_1 + Z_2 + \cdots + Z_m.$$

For any $\epsilon > 0$, $i \le n$ *and* $j \ge n$, *there exist* $C_i = C_i(\epsilon)$ *for* $i = 1, 2$ *such that*

$$\mathbf{P}(S_i \ge n\epsilon) \le C_1 \exp(-C_2 n),$$

and

$$\mathbf{P}(S_j \le -\epsilon n) \le C_1 \exp(-C_2 n).$$

Now we show that $h_n(N)/n$ and $h(X_n)/n$ are not very different if N is large.

Lemma 2.3 *For* $\epsilon > 0$, $c = 2$, *and* $b \ge 70$, *there exist* $N = N(\epsilon)$ *and* $C_i = C_i(\epsilon, N)$ *for* $i = 1, 2$ *such that*

$$\mathbf{P}(h(X_n) \ge h_n(N) + n\epsilon) \le C_1 \exp(-C_2 n).$$

Proof If

$$h(X_n) - h_n(N) \ge \epsilon n, \tag{2.14}$$

we may suppose that there are only $k \ge 1$ many N-long cut time pairs $\{\tau_{i_j}, \tau_{i_j+1}\}$ for $j = 1, \dots, k$ such that

$$\tau_{i_1} < \tau_{i_1+1} < \tau_{i_2} < \tau_{i_2+1} < \cdots < \tau_{i_j} < \tau_{i_j+1} \cdots < \tau_{i_{k-1}} < \tau_{i_{k-1}+1} < \tau_{i_k} \le n \le \tau_{i_k+1}$$

with $i_1 \ge 1$ and with

$$\sum_{j=1}^{k} H_{i_j} = \sum_{j=1}^{k} h(X_{\tau_{i_j+1}}) - h(X_{\tau_{i_j}}) \ge \epsilon n/2, \tag{2.15}$$

where

$$H_{i_j} = h(X_{\tau_{i_j+1}}) - h(X_{\tau_{i_j}}) > N \quad \text{for } j = 1, 2, \dots, k \le n/N, \tag{2.16}$$

or

$$h(X_{\tau_1}) \ge \epsilon n/2. \tag{2.17}$$

For the second case in (2.17), by Lemma 2.1, there exist $C_i = C_i(\epsilon)$ for $i = 1, 2$ such that

$$\mathbf{P}(h(X_{\tau_1}) \geq \epsilon n/2) = \mathbf{P}(H_0 \geq \epsilon n/2) \leq C_1 \exp(-C_2 n). \tag{2.18}$$

We focus on the first case in (2.15). By (2.7) and Lemma 2.1, $\{H_1, H_2, \ldots\}$ is an i.i.d sequence with

$$\mathbf{P}(H_i \geq m) \leq 0.115^m \quad \text{for } i \geq 1. \tag{2.19}$$

Thus, if (2.15) holds, by (2.15) and (2.16), it implies that there exist k many H_is in $\{H_1, \ldots, H_n\}$ for $1 \leq k \leq \lceil n/N \rceil$ such that $H_i > N$ and their sum is large than $\epsilon n/2$.

For a fixed k, it costs at most $\binom{n}{k}$ to fix the subsequence of these H_is from $\{H_1, \ldots, H_n\}$. We denote by $H_{i_1}, \ldots, H_{i_k}$ these fixed random variables. Since $\{H_i\}$ is an i.i.d sequence, the joint distribution of $H_{i_1}, \ldots, H_{i_k}$ is always the same for different i_js. With these observations,

$$\mathbf{P}\left(h(X_n) \geq h_n(N) + n\epsilon/2, \text{(2.15) holds}\right) \leq \sum_{k=1}^{\lceil n/N \rceil} \binom{n}{k} \mathbf{P}(H_{i_1} + \cdots + H_{i_k} \geq n\epsilon/2). \tag{2.20}$$

By (2.19), we know that

$$EH_i = EH_1 < \infty \quad \text{for each } i \geq 1.$$

Since $k \leq n/N + 1$, we may take $N = N(\epsilon)$ large such that for each $k \leq n$ and fixed $i_1, \ldots, i_k$

$$\mathbf{P}(H_{i_1} + \cdots + H_{i_k} \geq n\epsilon/2) \leq \mathbf{P}([H_{i_1} - EH_{i_1}] + \cdots + [H_{i_k} - EH_{i_k}] \geq n\epsilon/4) \tag{2.21}$$

Note that $\{H_{i_j} - EH_{i_j}\}$ is an i.i.d sequence with a zero-mean and an exponential tail for $j = 1, \ldots, k$, so by Lemma 2.2,

$$\mathbf{P}([H_{i_1} - EH_{i_1}] + \cdots + [H_{i_k} - EH_{i_1}] \geq n\epsilon/4) \leq C_3 \exp(-C_4 n). \tag{2.22}$$

By a standard entropy bound, as given in Corollary 2.6.2 of Engel [6], for $k \leq n/N$,

$$\binom{n}{k} \leq \exp(n \log N/N). \tag{2.23}$$

By (2.19)–(2.22), if we take N large, then there exist $C_i = C_i(\epsilon, N)$ for $i = 5, 6$ such that

$$\mathbf{P}\left(h_n(X_n) \geq h_n(N) + n\epsilon, \text{(2.15) holds}\right) \leq C_5 n \exp(-C_6 n). \tag{2.24}$$

So Lemma 2.3 holds by (2.18) and (2.24). □

We also need to control the time difference such that $\tau'_k(N) - \tau_k(N)$ cannot be large. We call $\tau_k(N)$ *M-tight* for $k \geq 1$ if

$$\tau'_k(N) - \tau_k(N) \leq M.$$

We list all M-tight N-short cut times as

$$\tau_1(N, M), \tau_2(N, M), \ldots, \tau_k(N, M), \ldots.$$

Suppose that $\tau_k(N, M) < \infty$. We know that $\tau_k(N, M) = \tau_i$ for some i. We denote $\tau'_k(N, M) = \tau_{i+1}$. For convenience, we also call $\tau_0(N, M) = 0$ and $\tau'_0(N, M) = 0$ whether the root is a cut vertex or not. Let

$$J_n = \max\{i : \tau_i(N, M) \leq n\}$$

and

$$h_n(N, M) = \sum_{i=0}^{J_n} \left(h\left(X_{\tau'_i(N,M)}\right) - h\left(X_{\tau_i(N,M)}\right)\right). \tag{2.25}$$

If $J_n = 0$, then

$$h_n(N, M) = 0. \tag{2.26}$$

The following lemma shows that $h_n(N, M)/n$ and $h_n(N)/n$ are not far away.

Lemma 2.4 *For $\epsilon > 0$, for N, and for each n, there exists $M = M(\epsilon, N)$ such that*

$$h_n(N) \leq h_n(N, M) + n\epsilon.$$

Proof If $h_n(N) > h_n(N, M) + n\epsilon$, we know that there are at least $\epsilon n/2N$ many $\{\tau_i(N)\}$ such that

$$\tau'_i(N) - \tau_i(N) > M. \tag{2.27}$$

If we take $M \geq 3N\epsilon^{-1}$, then

$$n \geq \sum_{i=1}^{I_n} \left(\tau'_i(N) - \tau_i(N)\right) > M\epsilon n/2N > n. \tag{2.28}$$

The contradiction shows that

$$h_n(N) \leq h_n(N, M) + n\epsilon.$$

So Lemma 2.4 follows. □

Let $\mathcal{E}(\epsilon)$ be the event that $h(X_n) \geq n(T - \epsilon)$. By Lemmas 2.3 and 2.4,

$$\lim_{n\to\infty} \mathbf{P}(h_n(N, M) \leq Tn/2, \mathcal{E}_n(\epsilon)) = 0.$$

Note that $\mathbf{P}(\mathcal{E}_n(\epsilon))$ is near one for large n, so there are at least $Tn/2M$ many $\tau_i(N, M)$s with $\tau_i(N, M) \leq n$ that also have a probability near one for large n. Hence, $\tau_k(N, M) = \infty$ cannot have a positive probability for each k. Therefore,

$$\tau_1(N, M) < \tau_2(N, M) < \cdots \tau_k(N, M) < \cdots < \infty. \tag{2.29}$$

By (2.29), we know that $\tau_k(N, M) = \tau_i$ for some i and

$$\tau_k'(N, M) - \tau_k(N, M) = \tau_{i+1} - \tau_i.$$

Therefore, by the same proof of (2.7), for $k \geq 1$

$$\left\{\left(\tau_k'(N, M) - \tau_k(N, M), h\left(X_{\tau_k'(N,M)}\right) - h\left(X_{\tau_k(N,M)}\right)\right)\right\} \text{ is an i.i.d. sequence.} \tag{2.30}$$

3 Large deviations for $h_n(N, M)$

By Lemma 2.1, we let

$$\mathbf{E}(\tau_2 - \tau_1) = A \geq 1 \quad \text{and} \quad \mathbf{E}\left(\tau_1'(N, M) - \tau_1(N, M)\right) = A(N, M)$$

and

$$\mathbf{E}(h(X_{\tau_2}) - h(X_{\tau_1})) = B \geq 1 \quad \text{and} \quad \mathbf{E}\left(h\left(X_{\tau_1'(N,M)}\right) - h\left(X_{\tau_1(N,M)}\right)\right) = B(N, M).$$

We set

$$T_n = \sum_{k=1}^{n} (\tau_{k+1} - \tau_k) \quad \text{and} \quad T_n(N, M) = \sum_{k=1}^{n} \left(\tau_k'(N, M) - \tau_k(N, M)\right)$$

and

$$H_n = \sum_{k=1}^{n} \left(h\left(X_{\tau_{k+1}}\right) - h\left(X_{\tau_k}\right)\right) \quad \text{and} \quad H_n(N, M) = \sum_{k=1}^{n} \left(h\left(X_{\tau_k'(N,M)}\right) - h\left(X_{\tau_k(N,M)}\right)\right).$$

By the law of large numbers,

$$\lim_{n\to\infty}\frac{T_n}{n}=A \quad \text{and} \quad \lim_{n\to\infty}\frac{T_n(N,M)}{n}=A(N,M) \tag{3.1}$$

and

$$\lim_{n\to\infty}\frac{H_n}{n}=B \quad \text{and} \quad \lim_{n\to\infty}\frac{H_n(N,M)}{n}=B(N,M). \tag{3.2}$$

If $\tau_i \le n \le \tau_{i+1}$ for $i \ge 1$, then

$$h(X_{\tau_i}) \le h(X_n) \le h(X_{\tau_{i+1}}). \tag{3.3}$$

Thus,

$$\frac{h(X_{\tau_i})}{\tau_{i+1}} \le \frac{h(X_n)}{n} \le \frac{h(X_{\tau_{i+1}})}{\tau_i}. \tag{3.4}$$

By (3.1) and (3.2),

$$\lim_{i\to\infty}\frac{h(X_{\tau_i})}{\tau_{i+1}} = \lim_{i\to\infty}\frac{h(X_{\tau_{i+1}})}{\tau_i} = \frac{B}{A}. \tag{3.5}$$

So by (1.1), (3.4), and (3.5),

$$\frac{B}{A}=T. \tag{3.6}$$

Regarding $B(N,M)$ and $A(N,M)$, we have the following lemma.

Lemma 3.1 *For $c=2$ and $b\ge 70$,*

$$\lim_{N,M\to\infty} A(N,M)=A \quad \textit{and} \quad \lim_{N,M\to\infty} B(N,M)=B \quad \textit{and} \quad \lim_{N,M\to\infty}\frac{B(N,M)}{A(N,M)}=T.$$

Proof By (2.5) and the definitions of $\tau_1(N)$ and $\tau_1(N,M)$, for each sample point ω, there exist large N and M such that

$$\tau_1(N,M)(\omega)=\tau_1(\omega),$$

where $\tau_1(N,M)(\omega)$ and $\tau_1(\omega)$ are $\tau_1(N,M)$ and τ_1 with ω. It also follows from the definition of $\tau_1'(N,M)$ that for the above N and M,

$$\tau_1'(N,M)(\omega)=\tau_2(\omega).$$

Thus, for each ω

$$\lim_{N,M\to\infty}\tau_1'(N,M)(\omega)-\tau_1(N,M)(\omega)=\tau_2(\omega)-\tau_1(\omega). \tag{3.7}$$

By the dominated convergence theorem,

$$\lim_{N,M\to\infty} A(N,M) = \lim_{N,M\to\infty} \mathbf{E}(\tau_1'(N,M) - \tau_1(N,M)) = \mathbf{E}(\tau_2 - \tau_1) = A. \quad (3.8)$$

Similarly,

$$\lim_{N,M\to\infty} B(N,M) = \lim_{N,M\to\infty} \mathbf{E}\left(h\big(X_{\tau_1'(N,M)}\big) - h\big(X_{\tau_1(N,M)}\big)\right) = B. \quad (3.9)$$

Therefore, Lemma 3.1 follows from (3.8), (3.9), and (3.6). □

Now we show that $h_n(N,M)$ has an exponential upper tail.

Lemma 3.2 *If $c = 2$ and $b \geq 70$, then for $\epsilon > 0$, there exist $N_0 = N_0(\epsilon)$ and $M_0 = M_0(\epsilon)$ such that for all $N \geq N_0$ and $M \geq M_0$*

$$\mathbf{P}(h_n(N,M) \geq n(T+\epsilon)) \leq C_1 \exp(-C_2 n), \quad (3.10)$$

where $C_i = C_i(\epsilon, N, M)$ for $i = 1, 2$ are constants.

Proof Recall that

$$J_n = \max\{i : \tau_i(N,M) \leq n\}.$$

So

$$\begin{aligned}
&\mathbf{P}(h_n(N,M) \geq n(T+B\epsilon)) \\
&\quad = \mathbf{P}\left(\sum_{i=1}^{J_n}\left(h\big(X_{\tau_i'(N,M)}\big) - h\big(X_{\tau_i(N,M)}\big)\right) \geq n(T+B\epsilon)\right) \\
&\quad \leq \mathbf{P}\left(\sum_{i=1}^{J_n}\left(h\big(X_{\tau_i'(N,M)}\big) - h\big(X_{\tau_i(N,M)}\big)\right) \geq n(T+B\epsilon),\, J_n \leq n\left(\frac{T}{B(N,M)} + \epsilon/2\right)\right) \\
&\qquad + \mathbf{P}\left(J_n > n\left(\frac{T}{B(N,M)} + \epsilon/2\right)\right) \\
&\quad \leq \mathbf{P}\left(\sum_{i=1}^{n(T/B(N,M)+\epsilon/2)}\left(h\big(X_{\tau_i'(N,M)}\big) - h\big(X_{\tau_i(N,M)}\big)\right) \geq n(T+B\epsilon)\right) \\
&\qquad + \mathbf{P}\left(J_n > n\left(\frac{T}{B(N,M)}\right) + \epsilon/2)\right) \\
&\quad = I + II. \qquad (3.11)
\end{aligned}$$

Here without loss of generality, we assume that $n(T/B(N,M) + \epsilon/2)$ is an integer, otherwise we can use $\lceil n(T/B(N,M)) + \epsilon/2)\rceil$ to replace $n(T/B(N,M)) + \epsilon/2)$.

We will estimate I and II separately. For I, note that by Lemma 3.2, there exist $N_0 = N_0(\epsilon)$ and $M_0 = M_0(\epsilon)$ such that for all $N \geq N_0$ and $M \geq M_0$

$$\mathbf{E}\left(\sum_{i=1}^{n(T/B(N,M)+\epsilon/2)}\left(h\left(X_{\tau_i'(N,M)}\right) - h\left(X_{\tau_i(N,M)}\right)\right)\right)$$
$$\leq nT(1 + B(N, M)\epsilon/2) \leq nT(1 + 2B\epsilon/3).$$

Note also that by (2.30),

$$\left\{h\left(X_{\tau_i'(N,M)}\right) - h\left(X_{\tau_i(N,M)}\right)\right\} \text{ is a uniformly bounded i.i.d. sequence,}$$

so by Lemma 2.2, there exist $C_i = C_i(\epsilon, N, M)$ for $i = 3, 4$ such that

$$\mathbf{P}\left(\sum_{i=1}^{n(T/B(N,M)+\epsilon/2)}\left(h\left(X_{\tau_i'(N,M)}\right) - h\left(X_{\tau_i(N,M)}\right)\right) \geq n(T + B\epsilon)\right) \leq C_3 \exp(-C_4 n). \tag{3.12}$$

Now we estimate II. By Lemma 3.1, there exist $N_0 = N_0(\epsilon, b)$ and $M_0 = M_0(\epsilon, b)$ such that for all $N \geq N_0$ and $M \geq M_0$

$$\mathbf{P}\left(J_n > n\left(\frac{T}{B(N, M)}\right) + \epsilon/2\right) = \mathbf{P}\left(J_n > n\left(A^{-1}(N, M) + \epsilon/3\right)\right). \tag{3.13}$$

Here without loss of generality, we also assume that $n(A^{-1}(N, M)+\epsilon/3)$ is an integer, otherwise we can use $\lceil n(A^{-1}(N, M) + \epsilon/3)\rceil$ to replace $n(A^{-1}(N, M) + \epsilon/3)$. Note that

$$\left\{J_n \geq n(A^{-1}(N, M) + \epsilon/3)\right\} \subset \left\{\sum_{i=1}^{n(A^{-1}(N,M)+\epsilon/3)} (\tau_i'(N, M) - \tau_i(N, M)) \leq n\right\}. \tag{3.14}$$

Note also that

$$\mathbf{E} \sum_{i=1}^{n(A^{-1}(N,M)+\epsilon/3)} \left(\tau_i'(N, M) - \tau_i(N, M)\right) = n(1 + \epsilon A(N, M)/3),$$

and, by (2.30), $\{\tau_i'(N, M) - \tau_i(N, M)\}$ is a uniformly bounded i.i.d. sequence, so by (3.13), and (3.14), and Lemma 2.2, there exist $C_i = C_i(\epsilon, b, N, M)$ for $i = 5, 6$ such that

$$\begin{aligned}
&\mathbf{P}\left(J_n > n\left(\frac{T}{B(N,M)} + \epsilon/2\right)\right)\\
&\le \mathbf{P}\left(J_n > n(A^{-1}(N,M) + \epsilon/3)\right)\\
&\le \mathbf{P}\left(\sum_{i=1}^{n(A^{-1}(N,M)+\epsilon/3)} (\tau_i'(N,M) - \tau_i(N,M)) \le n\right)\\
&\le C_5 \exp(-C_6 n).
\end{aligned} \tag{3.15}$$

For all large N and M, we substitute (3.12) and (3.15) in (3.11) to have

$$\mathbf{P}(h_n(N,M) \ge n(T+\epsilon)) \le I + II \le C_7 \exp(-C_8 n) \tag{3.16}$$

for $C_i = C_i(\epsilon, N, M)$ for $i = 7, 8$. Therefore, we have an exponential tail estimate for $h_n(N, M)$. So Lemma 3.2 follows. □

Let

$$L_n = \max\{i : \tau_i \le n\}$$

and

$$h_n = \sum_{i=1}^{L_n} \left(h(X_{\tau_i}) - h(X_{\tau_{i-1}})\right) \text{ if } L_n \ge 1 \quad \text{and} \quad h_n = 0 \text{ if } L_n = 0. \tag{3.17}$$

Recall that ρ_i is the leaving time defined in (2.3). We show the following subadditive argument for h_n.

Lemma 3.3 *For $c = 2$, $b \ge 2$, $N > 0$, and for each pair of positive integers n and m,*

$$\begin{aligned}
&\mathbf{P}(h_n \ge nC, \rho_0 \le N)\mathbf{P}(h_m \ge mC, \rho_0 \le N)\\
&\quad \le 2^N(b+1)n\mathbf{P}(h_{n+m+1}\\
&\quad \ge (n+m)C + 1, \rho_0 \le N),
\end{aligned}$$

for any $C > 0$.

Proof By the definition in (3.17), there exists $0 \le k \le n$ such that

$$\tau_k \le n \le \tau_{k+1}.$$

So

$$h_n = h(X_{\tau_k}) \le h(X_n) \le h(X_{\tau_{k+1}}). \tag{3.18}$$

For $i \geq nC$, we denote by $\mathcal{F}(x, i, N, nC)$ the event that walks $\{X_1, X_2, \ldots, X_i\}$ have

$$h(X_j) < nC \quad \text{for} j < i \quad \text{and} \quad h(X_i) = x \text{ with } h(x) \geq nC. \tag{3.19}$$

In addition, the number of walks $\{X_1, X_2, \ldots, X_i\}$ visiting the root is no more than N.

Note that on $\{h_n \geq nC, \rho_0 \leq N\}$, walks eventually move to some vertex x at some time i with $h(x) \geq nC$, and walks $\{X_1, X_2, \ldots, X_i\}$ visit the root no more than N times. So we may control $\{h_n \geq nC, \rho_0 \leq N\}$ by a finite step walks $\{X_1, X_2, \ldots, X_i\}$ in order to work on a further coupling process. More precisely,

$$\mathbf{P}(h_n \geq nC, \rho_0 \leq N) \leq \sum_{i \leq n} \sum_{x} \mathbf{P}\left(\mathcal{F}(x, i, N, nC)\right). \tag{3.20}$$

There are $b + 1$ many vertices adjacent to x. We just select one of them and denote it by z with $h(z) = h(x) + 1$. Let e_z be the edge with the vertices x and z. On $\mathcal{F}(x, i, N, nC)$, we require that the next move X_{i+1} will be from x to z. Thus, $X_{i+1} = z$. We denote this subevent by $\mathcal{G}(x, z, i, N, nC) \subset \mathcal{F}(x, i, N, nC)$. We have

$$\sum_{i \leq n} \sum_{x} \mathbf{P}\left(\mathcal{F}(x, i, N, nC)\right) \leq (b + 1) \sum_{i \leq n} \sum_{x} \mathbf{P}\left(\mathcal{G}(x, z, i, N, nC)\right). \tag{3.21}$$

Now we focus on $\{h_m \geq Cm, \rho_0 \leq N\}$. Let $\mathbf{T}_z$ be the subtree with the root at z and vertices in $\{v : h(v) \geq h(z)\}$. We define $\{X_n^i(z)\}$ to be the linearly reinforced random walks starting from z in subtree $\mathbf{T}_z$ for $n \geq i + 1$ with

$$X_{i+1}^i(z) = z \quad \text{and} \quad w(e_z, i + 1) = 2.$$

Note that walks $\{X_n^i(z)\}$ stay inside $\mathbf{T}_z$, so

$$w(e_z, n) = 2 \quad \text{for } n \geq i + 1. \tag{3.22}$$

We can define τ_k^i, ρ_0^i and $h_m^i(z)$ for $\{X_n^i(z)\}$ similar to the definitions of τ_k, ρ_0 and h_m for $\{X_n\}$.

On $w(e_z, i + 1) = 2$, we consider a probability difference between $\mathbf{P}(h_m \geq Cm, \rho_0 \leq N)$ and $\mathbf{P}(h_m^i(z) \geq Cm, \rho_0^i \leq N)$. Note that there are only b edges from the root, but there are $b + 1$ edges from vertex z with $w(e_z, n) = 2$, so the two probabilities are not the same. We claim that

$$\mathbf{P}(h_m \geq mC, \rho_0 \leq N) \leq 2^N \mathbf{P}(h_m^i(z) \geq Cm, \rho_0^i \leq N \mid w(e_z, i + 1) = 2). \tag{3.23}$$

To show (3.23), we consider a fixed path $(u_0 = \mathbf{0}, u_1, u_2, \ldots)$ in $\mathbf{T}$ with $\{X_1 = u_1, X_2 = u_2, \ldots\} \in \{h_m \geq Cm, \rho_0 \leq N\}$. Note that $[u_j, u_{j+1}]$ is an edge in $\mathbf{E}$. If we remove $\mathbf{T}$ from the root to z, it will be $\mathbf{T}_z$. So path $(\mathbf{0}, u_1, u_2, \ldots)$ in $\mathbf{T}$ will be a new path $(u_0(z) = z, u_1(z), u_2(z), \ldots)$ in $\mathbf{T}_z$ after removing. Thus, if

$$\{X_0 = \mathbf{0}, X_1 = u_1, X_2 = u_1, \ldots\} \in \{h_m \geq Cm, \rho_0 \leq N\},$$

then

$$\{X^i_{i+1} = z, X^i_{i+2}(z) = u_1(z), \ldots\} \in \{h^i_m(z) \geq Cm, \rho^i_0 \leq N\}.$$

On the other hand, given a fixed paths $\{\mathbf{0}, u_1, \ldots, u_j, \ldots\}$, it follows from the definition of $\{z, u_1(z), \ldots, u_j(z), \ldots\}$ that

$$w\left([u_j, u_{j+1}], k\right) = w\left([u_j(z), u_{j+1}(z)], i+1+k\right) \tag{3.24}$$

for any positive integers j and k. We may focus on a finite part $\{\mathbf{0}, u_1, \ldots u_l\}$ from $\{\mathbf{0}, u_1, \ldots\}$. Now if we can show that for all large l, and for each path $\{\mathbf{0}, u_1, u_2, \ldots, u_l\}$,

$$\begin{aligned}&\mathbf{P}(X_1 = u_1, X_2 = u_2, \ldots, X_l = u_l)\\ &\quad \leq 2^N \mathbf{P}\Big(X^i_{i+2}(z) = u_1(z), X^i_{i+3}(z) = u_2(z), \ldots, X^i_{i+2+l}(z)\\ &\quad = u_l(z) \mid w(e_z, i+1) = 2\Big),\end{aligned} \tag{3.25}$$

then (3.23) will be followed by the summation of all possible paths $\{\mathbf{0}, u_1, u_2, \ldots u_l\}$ for both sides in (3.25) and by letting $l \to \infty$. Therefore, to show (3.23), we need to show (3.25).

Note that

$$\mathbf{P}(X_1 = u_1, X_2 = u_2, \ldots, X_l = u_l) = \prod_{j=1}^{l} \mathbf{P}(X_j = u_j \mid X_{j-1} = u_{j-1}, \ldots, X_1 = u_1) \tag{3.26}$$

and

$$\begin{aligned}&\mathbf{P}(X^i_{i+2} = u_1(z), X^i_{i+3}(z) = u_2(z), \ldots, X^i_{i+2+l}(z) = u_l(z))\\ &\quad = \prod_{j=1}^{l} \mathbf{P}(X^i_{i+1+j}(z) = u_j(z) \mid X^i_{i+j}(z) = u_{j-1}(z), \ldots, X^i_{i+2}(z)\\ &\quad = u_1(z), w(e_z, i+1) = 2).\end{aligned} \tag{3.27}$$

If $u_{j-1} = \mathbf{0}$, then

$$\mathbf{P}(X_j = u_j \mid X_{j-1} = u_{j-1}, \ldots, X_1 = u_1) = \frac{w([u_{j-1}, u_j], j-1)}{\sum_e w(e, j)}, \tag{3.28}$$

where the sum in (3.28) takes over all possible edges adjacent to the root with vertices in $\mathbf{T}$. On the other hand, if $u_{j-1} = \mathbf{0}$, we know that $u_{j-1}(z) = z$, then by (3.22),

$$\begin{aligned}
&\mathbf{P}(X^i_{i+1+j}(z) = u_j(z) \mid X^i_{i+j}(z) = u_{j-1}(z), \ldots, X^i_{i+2}(z) = u_1(z), w(e_z, i+1) = 2) \\
&= \frac{w([u_{j-1}(z), u_j(z)], i+j)}{\sum_e w(e, i+j) + w(e_z, i+j)} = \frac{w([u_{j-1}(z), u_j(z)], i+j)}{\sum_e w(e, i+j) + 2},
\end{aligned} \tag{3.29}$$

where the sum in (3.29) takes all edges adjacent to z with vertices in $\mathbf{T}_z$ (not including e_z). We check the numerators in the right sides of (3.28) and (3.29). If $X_1, \ldots X_{j-1}$ never visit u_j, then both $w([u_{j-1}, u_j], j-1] = 1$ and $w([u_{j-1}(z), u_j(z)], i+j) = 1$. Otherwise, by (3.24) the two numerators are also the same. Similarly, the two sums in the denominators in the right sides of (3.28) and (3.29) are the same. Therefore, if $u_{j-1} = \mathbf{0}$, note that $\sum_e w(e, j) \geq 2$ for all j, so

$$\begin{aligned}
&2\mathbf{P}(X^i_{i+1+j}(z) = u_j(z) \mid X^i_{i+j}(z) = u_{j-1}(z), \ldots, X^i_{i+2}(z) = u_1(z), w(e_z, i+1) = 2) \\
&\geq \mathbf{P}(X_j = u_j \mid X_{j-1} = u_{j-1}, \ldots, X_1 = u_1).
\end{aligned} \tag{3.30}$$

If $u_{j-1} \neq \mathbf{0}$, we do not need to consider the extra term $w(e_z, i+j)$ in the denominator of the second right side of (3.29). So by the same argument of (3.30), if $u_{j-1} \neq \mathbf{0}$,

$$\begin{aligned}
&\mathbf{P}(X^i_{i+1+j}(z) = u_j(z) \mid X^i_{i+j}(z) = u_{j-1}(z), \ldots, X^i_{i+2}(z) = u_1(z), w(e_z, i+1) = 2) \\
&= \mathbf{P}(X_j = u_j \mid X_{j-1} = u_{j-1}, \ldots, X_1 = u_1)
\end{aligned} \tag{3.31}$$

Since we restrict $\rho_0 \leq N$ and $\rho^i_0 \leq N$, walks $\{X_1, X_2, \ldots\}$ visit the root no more than N times. On the other hand, walks $\{X^i_{i+2}(z), X^i_{i+3}(z), \ldots\}$ also visit z no more than N times. This indicates that there are at most N vertices u_js with $u_j = \mathbf{0}$ for $1 \leq j \leq l$ for the above path $\{\mathbf{0}, u_1, \ldots, u_l\}$. Thus, (3.25) follows from (3.26)–(3.31). So does (3.23).

With (3.23), we will show Lemma 3.3. Note that $\{h^i_m(z) \geq mC, \rho^i_0 \leq N\}$ only depends on the weight configurations of the edges with vertices inside $\mathbf{T}_z$, and weight $w(e_z, i+1)$, and the time interval $[i+2, \infty)$. In contrast, on $\mathcal{G}(x, z, i, N, nC)$, the last move of walks $\{X_1, \ldots, X_i, X_{i+1}\}$ is from x to z, but the other moves use the edges with the vertices inside $\{y : h(y) \leq h(z) - 1\}$. So by (3.23),

$$\begin{aligned}
&\mathbf{P}(h_m \geq Cm, \rho_0 \leq N) \\
&\quad \leq 2^N \mathbf{P}\left(h^i_m(z) \geq Cm, \rho^i_0 \leq N \mid w(e_z, i+1) = 2\right) \\
&\quad \leq 2^N \mathbf{P}\left(h^i_m(z) \geq Cm, \rho^i_0 \leq N \mid \mathcal{G}(x, z, i, N, nC)\right).
\end{aligned} \tag{3.32}$$

By (3.21) and (3.32),

$$\begin{aligned}
&\mathbf{P}(h_n \geq nC, \rho_0 \leq N)\mathbf{P}(h_m \geq mC, \rho_0 \leq N) \\
&\quad \leq \sum_{i \leq n} \sum_x 2^N (b+1) \mathbf{P}\left(\mathcal{G}(x, z, i, N, nC), h^i_m(z) \geq mC, \rho^i_0 \leq N\right).
\end{aligned} \tag{3.33}$$

If $i \le n$, then

$$h^i_m(z) \le h^i_{m+n-i}(z). \tag{3.34}$$

By (3.33) and (3.34),

$$\begin{aligned}
&\mathbf{P}(h_n \ge nC, \rho_0 \le N)\mathbf{P}(h_m \ge mC, \rho_0 \le N) \\
&\quad \le \sum_{i \le n} \sum_x 2^N (b+1)\mathbf{P}\left(\mathcal{G}(x, z, i, N, nC), h^i_m(z) \ge mC, \rho^i_0 \le N\right) \\
&\quad \le \sum_{i \le n} 2^N (b+1)\mathbf{P}\left(\bigcup_x \left\{\mathcal{G}(x, z, i, N, nC), h^i_{m+n-i}(z) \ge mC\right\}\right).
\end{aligned} \tag{3.35}$$

Note that for each x and i,

$$\left\{\mathcal{G}(x, z, i, N, nC), h^i_{m+n-i}(z) \ge mC\right\}$$

implies that the walks first move to x at time i with $h(x) \ge nC$ and the number of walks $\{X_1, \ldots, X_i\}$ back to the root is not more than N. After that, the walks continue to move from x to z. After this move, the walks move inside subtree $\mathbf{T}_z$. So i is a cut time and X_i is a cut vertex with $h(X_i) \ge nC$. Therefore, together with $h^i_{n+m-i}(z) \ge mC$, $\left\{\mathcal{G}(x, z, i, N, nC), h^i_{m+n-i}(z) \ge mC\right\}$ implies that $\{h_{n+m+1} \ge (n+m)C + 1, \rho_0 \le N\}$ occurs. In other words,

$$\left\{\mathcal{G}(x, z, i, N, nC), h^i_{m+n-i}(z) \ge mC\right\} \subset \{h_{n+m+1} \ge (n+m)C + 1, \rho_0 \le N\}. \tag{3.36}$$

Therefore,

$$\bigcup_x \left\{\mathcal{G}(x, z, i, N, nC), h^i_{m+n-i}(z) \ge mC\right\} \subset \{h_{n+m+1} \ge (n+m)C + 1, \rho_0 \le N\}. \tag{3.37}$$

Finally, by (3.35) and (3.37),

$$\begin{aligned}
&\mathbf{P}(h_n \ge nC, \rho_0 \le N)\mathbf{P}(h_m \ge mC, \rho_0 \le N) \\
&\quad \le 2^N (b+1) n \mathbf{P}(h_{n+m+1} \ge (n+m)C + 1, \rho_0 \le N).
\end{aligned} \tag{3.38}$$

Therefore, Lemma 3.3 follows from (3.38). □

We let

$$a_n = -\log \mathbf{P}(h_n \ge (T + \epsilon)n, \rho_0 \le N). \tag{3.39}$$

We may take ϵ small such that $T+\epsilon<1$. By Lemma 3.3, for any n and m

$$a_{n+m+1} \le a_n + a_m + \log n + N\log 2 + \log(b+1). \tag{3.40}$$

By (3.40) and a standard subadditive argument (see (II.6) in Grimmett [7]), we have the following lemma.

Lemma 3.4 *For $c=2$ and any $N>0$ and $b\ge 2$, there exists $0\le \alpha(N)<\infty$ such that*

$$\begin{aligned}&\lim_{n\to\infty}\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n, \rho_0\le N)\\ &\quad=\inf_n\left\{\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n, \rho_0\le N)\right\}=\alpha(N).\end{aligned}$$

It follows from the definition and Lemma 3.4 that $\alpha(N)$ is a non-negative decreasing sequence in N. Thus, there exists a finite number $\alpha\ge 0$ such that

$$\lim_{N\to\infty}\alpha(N)=\alpha. \tag{3.41}$$

By (3.41) and Lemma 3.4, for each N,

$$\alpha\le\alpha(N)\le\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n, \rho_0\le N). \tag{3.42}$$

On the other hand, note that the walk is transient, so $\rho_0<\infty$. Thus, for any fixed n,

$$\lim_{N\to\infty}\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n, \rho_0\le N)=\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n) \tag{3.43}$$

By (3.42) and (3.43),

$$\alpha\le\liminf_n\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n) \tag{3.44}$$

Note that for each N,

$$\begin{aligned}&\limsup_n\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n)\\ &\quad\le\lim_{n\to\infty}\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n, \rho_0\le N)=\alpha(N).\end{aligned}$$

So for each $\delta>0$ we may take N large such that

$$\limsup_n\frac{-1}{n}\log \mathbf{P}(h_n\ge (T+\epsilon)n)\le\alpha(N)\le\alpha+\delta. \tag{3.45}$$

We summarize (3.44) and (3.45) as the following lemma.

Lemma 3.5 *For $c = 2$ and any $b \geq 2$, there exists a constant $\alpha \geq 0$ such that*

$$\lim_{n\to\infty} \frac{-1}{n} \log \mathbf{P}(h_n \geq (T+\epsilon)n) = \alpha.$$

4 Proof of Theorem 1

Note that for $\epsilon < 1 - T$, and for all large n,

$$\left(\frac{b}{b+1}\right)^n \leq \mathbf{P}(h(X_{i+1}) > h(X_i) \quad \text{for } 0 \leq i \leq n) \leq \mathbf{P}(h(X_n) \geq n(T+\epsilon)). \tag{4.1}$$

By (4.1),

$$\limsup_{n\to\infty} \frac{-1}{n} \log \mathbf{P}(h(X_n) \geq n(T+\epsilon)) < \infty. \tag{4.2}$$

Note also that

$$\begin{aligned}
&\mathbf{P}(h(X_n) \geq n(T+\epsilon)) \\
&\quad \leq \mathbf{P}(h(X_n) \geq n(T+\epsilon), h_n(N,M) \geq n(T+\epsilon/2)) \\
&\qquad + \mathbf{P}(h(X_n) - h_n(N,M) \geq n\epsilon/2).
\end{aligned} \tag{4.3}$$

By Lemmas 2.3 and 2.4, for $\epsilon > 0$, we select N and M such that

$$\mathbf{P}(h(X_n) \geq n(T+\epsilon)) \leq \mathbf{P}(h_n(N,M) \geq n(T+\epsilon/2)) + C_1 \exp(-C_2 n). \tag{4.4}$$

For N and M in (4.4), we may require that $N \geq N_0$ and $M \geq M_0$ for N_0 and M_0 in Lemma 3.2. By (4.4) and Lemma 3.2, there exist $C_i = C_i(\epsilon, N, M)$ for $i = 3, 4$ such that

$$\mathbf{P}(h(X_n) \geq n(T+\epsilon)) \leq C_3 \exp(-C_4 n). \tag{4.5}$$

By (4.5), for $\epsilon > 0$,

$$0 < \liminf_{n\to\infty} \frac{-1}{n} \log \mathbf{P}(h(X_n) \geq n(T+\epsilon)). \tag{4.6}$$

It remains for us to show the existence of the limit in Theorem 1. We use a similar proof in Lemma 3.3 to show it. Let $\mathcal{F}(x,k,n)$ be the event that $h(X_i) < n(T+\epsilon)$ for $i = 1, \ldots, k-1$, $h(X_k) \geq n(T+\epsilon)$ and $h(X_k) = x$ for $k \leq n$. Thus,

$$\mathbf{P}(h(X_n) \geq n(T+\epsilon)) \leq \sum_{k\leq n} \sum_{x\in\mathbf{T}} \mathbf{P}(\mathcal{F}(x,k,n)) \tag{4.7}$$

Note that $\mathcal{F}(x, k, n)$ depends on finite step walks $\{X_0, \ldots, X_k\}$. We need to couple the remaining walks $\{X_{k+1}, X_{k+2}, \ldots\}$ such that k is a cut time.Let $\mathcal{Q}(x, k)$ be the event that $X_k = x$ and $\{X_t\}$ will stay inside $\mathbf{T}_x$ but never returns to x for $t > k$. Since the walks are transient, we may let

$$\mathbf{P}(\mathcal{Q}(\mathbf{0}, 0)) = \nu > 0. \tag{4.8}$$

Let e_x denote the edge with vertices x and w for $h(w) = h(x) - 1$. We know that $\mathcal{Q}(x, k)$ depends on initial weight $w(e_x, k)$, and the weights in the edges with the vertices in $\mathbf{T}_x$, respectively.

Therefore, by the same discussion of (3.23) in Lemma 3.3,

$$2\mathbf{P}(\mathcal{Q}(x, k) \mid \mathcal{F}(x, k, n)) \geq \left(\frac{b+2}{b}\right) \mathbf{P}(\mathcal{Q}(x, k) \mid \mathcal{F}(x, k, n)) \geq \nu. \tag{4.9}$$

Thus, by (4.7) and (4.9),

$$\begin{aligned} &\mathbf{P}(h(X_n) \geq n(T + \epsilon)) \\ &\leq \sum_{x \in \mathbf{T}} \sum_{k \leq n} \mathbf{P}\left(\mathcal{F}(x, k, n)\right) \mathbf{P}(\mathcal{Q}(x, k) \mid \mathcal{F}(x, k, n)) \left(\frac{b+2}{b}\right) \nu^{-1} \\ &\leq 2\nu^{-1} \sum_{k \leq n} \mathbf{P}\left(\bigcup_{x \in \mathbf{T}} \mathcal{F}(x, k, n) \cap \mathcal{Q}(x, k) \right). \end{aligned} \tag{4.10}$$

If $\mathcal{F}(x, k, n) \cap \mathcal{Q}(x, k)$ occurs, it implies that the walks move to x at $k \leq n$ with $h(x) \geq n(T + \epsilon)$. After that, the walks continue to move inside $\mathbf{T}_x$ from x and never return to x. This implies that k is a cut time and X_k is a cut vertex with $h(X_k) \geq n(T + \epsilon)$. So for $0 \leq k \leq n$ and for each x,

$$\mathcal{F}(x, k, n) \cap \mathcal{Q}(x, k) \subseteq \{h_k \geq n(T + \epsilon)\}. \tag{4.11}$$

Thus,

$$\bigcup_{x \in \mathbf{T}} \mathcal{F}(x, k, n) \cap \mathcal{Q}(x, k) \subseteq \{h_k \geq n(T + \epsilon)\}. \tag{4.12}$$

Note that for $0 \leq k \leq n$,

$$h_k \leq h_n. \tag{4.13}$$

By (4.10)–(4.13),

$$\mathbf{P}(h(X_n) \geq n(T + \epsilon)) \leq 2\nu^{-1} n \mathbf{P}(h_n \geq n(T + \epsilon)). \tag{4.14}$$

On the other hand, we suppose that $h_n \geq n(T + \epsilon)$. Note that if $\tau_k \leq n \leq \tau_{k+1}$, then by (3.18),

$$h_n = h\left(X_{\tau_k}\right) \leq h(X_n). \tag{4.15}$$

By (4.15),

$$\mathbf{P}(h_n \geq n(T+\epsilon)) \leq \mathbf{P}(h(X_n) \geq n(T+\epsilon)). \tag{4.16}$$

Now we are ready to show Theorem 1.

Proof of Theorem 1 Together with (4.14), (4.16), and Lemma 3.5,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(h(X_n) \geq n(T+\epsilon)) = \alpha. \tag{4.17}$$

By (4.2) and (4.6),

$$0 < \alpha < \infty. \tag{4.18}$$

Therefore, Theorem 1 follows from (4.17) and (4.18). □

5 Proof of Theorem 2

Similarly, we define the same cut times τ_i that we defined for the linearly reinforced random walk. We have $\left(\tau_{k+1} - \tau_k, h(Y_{\tau_{k+1}}) - h(Y_{\tau_k})\right)$ as an i.i.d. sequence. We can also follow Durrett et al. [5] Lemmas 7 and 8 to show that there exist C_i for $i = 1, 2$ such that, for each $k \geq 1$,

$$\mathbf{P}(\tau_{k+1} - \tau_k \geq m) \leq C_1 \exp(-C_2 m) \tag{5.1}$$

and

$$\mathbf{P}(h(Y_{\tau_{k+1}}) - h(Y_{\tau_k}) \geq m) \leq C_1 \exp(-C_2 m). \tag{5.2}$$

By (5.1) and (5.2), similar to our approach the linearly reinforced random walk, we set

$$S_n = \sum_{k=1}^{n} (\tau_k - \tau_{k-1}) \quad \text{and} \quad K_n = \sum_{k=1}^{n} \left(h(Y_{\tau_k}) - h(Y_{\tau_{k-1}})\right). \tag{5.3}$$

By the law of large numbers,

$$\lim_{n\to\infty} \frac{S_n}{n} = A \quad \text{and} \quad \lim_{n\to\infty} \frac{K_n}{n} = B. \tag{5.4}$$

With these observations, Theorem 2 can follow from the exact proof of Theorem 1. In fact, we may not need to truncate τ_i to $\tau_i(N, M)$ as we did for Theorem 1, since we can use (5.1) and (5.2) directly. □

6 Proof of Theorem 3

Now we need to estimate $\mathbf{P}(h(Y_n) \le n(S-\epsilon))$. Let

$$L_n = \max\{i, \tau_i \le n\}$$

and let

$$h_n = \sum_{i=1}^{L_n} \left(h\left(Y_{\tau_i}\right) - h\left(Y_{\tau_{i-1}}\right)\right) \quad \text{if } L_n \ge 1 \quad \text{and} \quad h_n = 0 \text{ if } L_n = 0. \tag{6.1}$$

By (1.7), (5.3), and an argument similar to (3.6), we have

$$\frac{B}{A} = S. \tag{6.2}$$

Since $h_n \le h(Y_n)$, by (5.1)

$$\begin{aligned} &\mathbf{P}(h(Y_n) \le n(S-\epsilon B)) \\ &\quad \le \mathbf{P}(h_n \le n(S-\epsilon B)) \\ &\quad \le \mathbf{P}\left(\sum_{i=1}^{L_n} \left(h\left(Y_{\tau_i}\right) - h\left(Y_{\tau_{i-1}}\right)\right) \le n(S-\epsilon B\right) + \mathbf{P}(\tau_1 > n) \\ &\quad \le \mathbf{P}\left(\sum_{i=1}^{L_n} \left(h\left(Y_{\tau_i}\right) - h\left(Y_{\tau_{i-1}}\right)\right) \le n(S-\epsilon B\right) + C_1 \exp(-C_2 n). \end{aligned} \tag{6.3}$$

We split

$$\begin{aligned} &\mathbf{P}\left(\sum_{i=1}^{L_n} (h(Y_{\tau_i}) - h(Y_{\tau_{i-1}})) \le n(S-\epsilon B)\right) \\ &\le \mathbf{P}\left(\sum_{i=1}^{L_n} (h(Y_{\tau_i}) - h(Y_{\tau_{i-1}})) \le n(S-\epsilon B), L_n \ge n(SB^{-1}-\epsilon/2)\right) \\ &\quad + \mathbf{P}\left(L_n < n(SB^{-1}-\epsilon/2)\right) = I + II. \end{aligned}$$

We estimate I and II separately:

$$\begin{aligned} I &= \mathbf{P}\left(\sum_{i=1}^{L_n} (h(Y_{\tau_i}) - h(Y_{\tau_{i-1}})) \le n(S-\epsilon B), L_n \ge n(SB^{-1}-\epsilon/2)\right) \\ &\le \mathbf{P}\left(\sum_{i=1}^{n(SB^{-1}-\epsilon/2)} (h(Y_{\tau_i}) - h(Y_{\tau_{i-1}})) \le n(S-\epsilon B)\right). \end{aligned} \tag{6.4}$$

Note that

$$\mathbf{E}\left(\sum_{i=1}^{n(SB^{-1}-\epsilon/2)}(h(Y_{\tau_i})-h(Y_{\tau_{i-1}}))\right)=n(S-\epsilon B/2). \tag{6.5}$$

Note also that by (5.2), $\{h(Y_{\tau_i})-h(Y_{\tau_{i-1}})\}$ is an i.i.d. sequence with an exponential tail for $k\geq 2$, so by Lemma 2.2 there exist $C_i=C_i(\epsilon,B)$ for $i=3,4$ such that

$$I\leq C_3\exp(-C_4 n). \tag{6.6}$$

Also, by (6.2),

$$II=\mathbf{P}\left(L_n<n(SB^{-1}-\epsilon/2)\right)=\mathbf{P}\left(\sum_{i=1}^{n(SB^{-1}-\epsilon/2)}(\tau_i-\tau_{i-1})\geq n\right)$$
$$=\mathbf{P}\left(\sum_{i=1}^{n(A^{-1}-\epsilon/2)}(\tau_i-\tau_{i-1})\geq n\right).$$

Note that

$$\mathbf{E}\sum_{i=1}^{n(A^{-1}-\epsilon/2)}(\tau_i-\tau_{i-1})=n(1-\epsilon A/2). \tag{6.7}$$

Note also that by (5.1), $\{\tau_i-\tau_{i-1}\}$ is an i.i.d. sequence with an exponential tail for $k\geq 2$, so by Lemma 2.2, there exist $C_i=C_i(\epsilon,B)$ for $i=5,6$ such that

$$II\leq C_5\exp(-C_6 n). \tag{6.8}$$

Together with (6.3), (6.4), (6.6), and (6.8), there exist $C_i=C_i(c,\epsilon,B)$ for $i=7,8$ such that

$$\mathbf{P}(h(Y_n)\leq n(S-\epsilon))\leq C_7\exp(-C_8 n). \tag{6.9}$$

From (6.9),

$$0<\liminf\frac{-1}{n}\log\mathbf{P}(h(Y_n)\leq n(S-\epsilon)). \tag{6.10}$$

If the walks repeatedly move in the edge connecting the origin in n times, we have the probability C^n for a positive constant $C=C(b)$. Thus, for $\epsilon<S$ and for all large n,

$$C^n\leq\mathbf{P}(h(Y_n)\leq 1)\leq\mathbf{P}(h(Y_n)\leq n(S-\epsilon)). \tag{6.11}$$

So for $\epsilon < S$,

$$\limsup \frac{-1}{n} \log \mathbf{P}(h(Y_n) \le n(S-\epsilon)) < \infty. \tag{6.12}$$

Therefore, Theorem 3 follows from (6.10) and (6.12). □

Acknowledgments The author would like to thank a careful referee for his many valuable comments, which corrected a few mistakes and improved the quality of this paper. He would also like to thank Takei M. for pointing out a few typos and a few references.

References

1. Aidekon, E.: Transient random walks in random environment on a Galton Watson tree. Probab. Theory Relat. Fields **142**, 525–559 (2008)
2. Collevecchio, A.: Limit theorems for reinforced random walks on certain trees. Probab. Theory Relat. Fields **136**, 81–101 (2006)
3. Davis, B.: Reinforced random walks. Probab. Theory Relat. Fields **84**, 203–229 (1990)
4. Diaconis, P.: Recent progress on de Fietti's notions of exchangeability. Bayesian stat. 3, 115–125, (1988) Oxford Univ. Press.
5. Durrett, R., Kesten, H., Limic, V.: Once edge-reinforced random walk. Probab. Theory Relat. Fields **122**, 567–592 (2002)
6. Engel, E.: Sperner Theory. Cambridge Univ, Press, New York (1997)
7. Grimmett, G.: Percolation. Springer-Verlag, New York (1999)
8. Pemantle, R.: Phase transition in reinforced random walks and RWRE on trees. Ann. Probab. **16**, 1229–1241 (1988)
9. Sabot, C., Tarres, P.: Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model (2012) (arXiv 1111.3991v3)
10. Takeshima, M.: Behavior of 1-dimensional reinforced random walk. Osaka J. Math. **7**, 355–372 (2000)