

## BOUNDS ON THE SPEED AND ON REGENERATION TIMES FOR CERTAIN PROCESSES ON REGULAR TREES

BY ANDREA COLLEVECCHIO<sup>1</sup> AND TOM SCHMITZ<sup>2</sup>

*Università Ca' Foscari and Max Planck Institute for Mathematics in the Sciences*

We develop a technique that provides a lower bound on the speed of transient random walk in a random environment on regular trees. A refinement of this technique yields upper bounds on the first regeneration level and regeneration time. In particular, a lower and upper bound on the covariance in the annealed invariance principle follows. We emphasize the fact that our methods are general and also apply in the case of once-reinforced random walk. Durrett, Kesten and Limic [*Probab. Theory Related Fields.* **122** (2002) 567–592] prove an upper bound of the form  $b/(b + \delta)$  for the speed on the  $b$ -ary tree, where  $\delta$  is the reinforcement parameter. For  $\delta > 1$  we provide a lower bound of the form  $\gamma^2 b/(b + \delta)$ , where  $\gamma$  is the survival probability of an associated branching process.

**1. Introduction.** Random processes with long memory have gained considerable attention in the recent past. Two emblematic examples of such processes are random walks in a random environment and reinforced processes. Although considerable progress has been achieved, there are many basic questions that remain open. We refer to the overviews by Sznitman [24] and Zeitouni [26, 27] for random walk in a random environment on  $\mathbb{Z}^d$ , and by Pemantle [20] for reinforced processes on  $\mathbb{Z}^d$  and on trees.

For the latter topic, we also refer the reader to Davis [9]. In this article we look at certain transient processes on regular trees, more precisely at random walk in a random environment and at once-reinforced random walk. An important question is to obtain an explicit expression for the speed (if at all it exists), or at least to get good estimates. This is in general a hard question, even for Markov chains as the biased random walk on a general tree, that is, a graph without cycles. For this model there is in general no explicit expression for the speed, and often only an upper bound is at hand. It is, in general, hard to find a lower bound, and we refer to

---

Received May 2010.

<sup>1</sup>Supported by the DFG-Forschergruppe 718 “Analysis and Stochastics in Complex Physical Systems” and by Italian PRIN 2007 Grant 2007TKLTSR “Computational markets design and agent-based models of trading behavior.”

<sup>2</sup>Supported by a postdoctoral research grant from the Max Planck Institute Mathematics in the Sciences.

*MSC2010 subject classifications.* 60K37, 60K99.

*Key words and phrases.* Random walk in a random environment, once edge-reinforced random walk, lower bound on the speed, regeneration times, regular trees.

Chen [4] for several examples. We also point out random walks on general graphs (Virág [25]) where basically no lower bound on the speed is available.

For random walk in a random environment, the speed is explicitly known only in one-dimensional models. On  $\mathbb{Z}^d$ ,  $d \geq 2$ , not much is known about the speed, and even worse, if  $d \geq 3$ , it is still open if a law of large numbers with constant speed holds (see [24, 26, 27]). On regular trees, however, a law of large numbers holds (see [14]) and transience implies that the speed is positive. This follows from Theorem 1.1 in Aidékon [1] that treats the more general setting of Galton–Watson trees. In Proposition 1.1 in [3], Aidékon introduced a simple lower bound for the speed on Galton–Watson trees. One of our goals is to find a different bound on the speed for random walks in a random environment on regular trees. Our approach looks more general, and we apply it to another class of processes with long memory: once edge-reinforced random walk. Moreover, we provide examples where our bound works better than the one introduced by Aidékon (see the example following Theorem 2.12). For an analysis of the recurrent regime of random walks in random environment on regular trees, we refer the reader to Hu and Shi [15, 16].

Once edge-reinforced random walk on regular trees is transient and has positive speed (see Theorems 1 and 2 in Durrett, Kesten and Limic [13]). They propose an upper bound on the speed, but no lower bound that is always positive is at hand. With similar techniques to those in the setting of random walk in random environment, we derive a lower bound.

In order to provide a lower bound on the speed, it is instrumental to find a lower bound for the escape probability from the root, as well as an upper bound for the expected number of returns to the root. Both of these bounds are obtained with the help of an auxiliary branching process that already appeared in Collecchio [6]. In particular the escape probability is bounded from below by the survival probability of the branching process (see Propositions 2.6 and 2.15). For once-reinforced random walk, the branching process can be constructed in such a way that its survival probability is always positive, whereas for random walk in random environment we need additional assumptions.

By a refinement of our methods, we are, moreover, able to derive a *common* explicit upper bound on all the moments of a first regeneration time  $\tau_1$ . These bounds are general and hold for random walk in a random environment as well as for once edge-reinforced random walk (see Theorem 3.7). In words, this first regeneration time is the first time the height of the walk reaches a new maximum and from then on never backtracks below this maximum. Regeneration times enjoy a widespread use in different settings, and we refer for instance to Lyons, Pemantle and Peres [18] for biased random walk on a Galton–Watson tree, to Durrett, Kesten and Limic [13] for once-reinforced random walk on a regular tree and to Sznitman [24] for random walk in a random environment on  $\mathbb{Z}^d$ .

The main step is to derive an explicit upper exponential tail on the first regeneration level  $\ell_1$ , defined as  $\ell_1 = |X_{\tau_1}|$ , where  $|\cdot|$  denotes the height of a vertex (see Theorem 3.5). We inspire ourselves from Collecchio [7], where a similar

technique was introduced, although in the setting of the vertex-reinforced jump process. Let us mention that a detailed analysis of the tail behavior of the first regeneration time is presented in Propositions 2.1 and 2.2 in Aidékon [2], revealing an exponential and a subexponential regime on regular trees. We emphasize that we obtain *explicit* upper bounds on all moments of the first regeneration time under certain assumptions, in contrast to [2], where only the finiteness of the moments follows. In particular, these bounds on the first regeneration level, respectively, regeneration time, imply a lower and an upper bound on the covariance of the Brownian motion that appears as the limiting object in an annealed invariance principle (see Theorem 3.8 and Proposition 3.9).

This article is organized as follows. In Section 2, we provide a lower bound on the speed for random walk in a random environment and for once edge-reinforced random walk, and in Section 3 we derive moment bounds on the first regeneration time that are completely general and hold for random walk in a random environment and for once edge-reinforced random walk.

**2. On the speed.** Let us start by introducing some notation. Consider the  $b$ -ary regular tree  $\mathcal{G}_b$  with root  $\rho$ . We assume that the root  $\rho$  has a parent  $\overleftarrow{\rho}$ . Hence each vertex in the tree is connected to  $b + 1$  vertices, except for  $\overleftarrow{\rho}$ , that is only connected to  $\rho$ . For any vertex  $v$ , denote by  $|v|$  its distance to the root, that is, the number of edges on the unique self-avoiding path connecting  $v$  and  $\rho$ . Level  $i$  is the set of vertices  $v$  such that  $|v| = i$ , with the exception that  $|\overleftarrow{\rho}| = -1$ . For  $v \neq \overleftarrow{\rho}$ , define  $\overleftarrow{v}$ , called the parent of  $v$ , to be the unique vertex at level  $|v| - 1$  connected to  $v$ . We say that  $v$  is a child of  $\overleftarrow{v}$ . We say that a vertex  $v_0$  is a descendant of the vertex  $v$  if the latter lies on the unique self-avoiding path connecting  $v_0$  to  $\rho$ , and  $v_0 \neq v$ . In this case,  $v$  is said to be an ancestor of  $v_0$ . For any vertex  $\mu$ , let  $\Lambda_\mu$  be the subtree of  $\mathcal{G}_b$  consisting of  $\mu$ , its descendants and the edges connecting them, that is, the  $b$ -ary subtree rooted at  $\mu$ . Let  $\overleftarrow{\Lambda}_\mu$  be the smallest subtree of  $\mathcal{G}_b$  containing  $\Lambda_\mu$  and the vertex  $\overleftarrow{\mu}$ .

2.1. *Random walk in random environment.* Let us define the random environment. To each vertex  $v$ , different from  $\overleftarrow{\rho}$ , we assign a  $b$ -dimensional random vector with positive entries

$$\mathbf{A}_v \stackrel{\text{def}}{=} (A_v^{(1)}, A_v^{(2)}, \dots, A_v^{(b)}).$$

We assume that these vectors are i.i.d. under the measure  $\mathbb{P}$ . Moreover, following Lyons and Pemantle [17], we assume that the coordinates are identically distributed. The random environment  $\omega$  is defined by  $\omega(\overleftarrow{\rho}, \rho) = 1$  and for any vertex  $v \neq \overleftarrow{\rho}$ ,

$$(2.1) \quad \omega(v, \overrightarrow{v}^{(i)}) = \frac{A_v^{(i)}}{1 + \sum_j A_v^{(j)}}; \quad \omega(v, \overleftarrow{v}) = \frac{1}{1 + \sum_j A_v^{(j)}}.$$

For a vertex  $v$  we define the Markov chain  $\{X_n, n \geq 0\}$  started at  $v$  by

$$\begin{aligned} \mathbf{P}_{v,\omega}(X_0 = v) &= 1, \\ \mathbf{P}_{v,\omega}(X_{n+1} = \mu_1 | X_n = \mu_0) &= \omega(\mu_0, \mu_1), \end{aligned}$$

for any pair of neighbors  $\mu_0, \mu_1$ . We introduce further the annealed measure as the semi-direct product  $\mathbf{P}_v = \mathbb{P} \times \mathbf{P}_{v,\omega}$ . We write  $\mathbf{P}_\omega$  and  $\mathbf{P}$  for  $\mathbf{P}_{\rho,\omega}$ , respectively,  $\mathbf{P}_\rho$ . We also write  $A$  and  $\mathbf{A} = (A^{(1)}, \dots, A^{(b)})$  for a generic copy of  $A_v^{(i)}$ ,  $1 \leq i \leq b$ , respectively, for a generic copy of  $\mathbf{A}_v = (A_v^{(1)}, \dots, A_v^{(b)})$ . We introduce the hitting times of a vertex  $v$ , respectively, of a level  $i$

$$(2.2) \quad T(v) \stackrel{\text{def}}{=} \inf\{k \geq 0 : X_k = v\} \quad \text{and} \quad T_i \stackrel{\text{def}}{=} \inf\{k \geq 0 : |X_k| = i\}.$$

We further introduce the respective return times

$$(2.3) \quad D \stackrel{\text{def}}{=} \inf\{n \geq 1 : X_n = \overleftarrow{X}_0\}, \quad D(v) \stackrel{\text{def}}{=} \inf\{n \geq 1 : X_{n-1} = v, X_n = \overleftarrow{v}\},$$

and the annealed return probability

$$(2.4) \quad \beta \stackrel{\text{def}}{=} \mathbf{P}(D < \infty).$$

For any graph  $G$ , denote with  $\text{Vert}(G)$  the set of vertices of  $G$ . To each ordered pair of neighbors  $v, \mu \in \text{Vert}(\mathcal{G}_b)$  assign a collection of independent exponentials  $h_k(v, \mu), k \geq 0$ , each with mean one. We assume that all these collections are independent. Using these exponentials, we now provide a construction of random walk in random environment on an arbitrary subtree (see [22] for a similar construction for reinforced processes).

DEFINITION 2.1 (Extension  $\mathbf{Y}^{\mathcal{C}}$ ). Fix a subtree  $\mathcal{C}$  of  $\mathcal{G}_b$ . The extension  $\mathbf{Y}^{\mathcal{C}}$  of  $\mathbf{X}$  on the subtree  $\mathcal{C}$  is defined as follows. Fix a starting point  $\eta$  in  $\mathcal{C}$ , that is,  $Y_0^{\mathcal{C}} = \eta$ . We define  $\mathbf{Y}^{\mathcal{C}}$  iteratively in the following way. Let  $s_1(v)$  be the first time  $\mathbf{Y}^{\mathcal{C}}$  reaches some vertex  $v$ . Define  $N_v^{\mathcal{C}}$  to be the set of neighbors of  $v$  in  $\mathcal{C}$ . The first jump after  $s_1(v)$  is toward the neighbor  $\mu \in N_v^{\mathcal{C}}$  for which the following minimum,

$$(2.5) \quad \min_{\eta \in N_v^{\mathcal{C}}} \frac{h_1(v, \eta)}{\omega(v, \eta)},$$

is a.s. attained. The probability that the minimum in (2.5) is achieved by the vertex  $\mu$  is  $\omega(v, \mu) / \sum_{\eta \in N_v^{\mathcal{C}}} \omega(v, \eta)$ . We define  $s_k(v), k \geq 2$ , inductively via

$$s_k \stackrel{\text{def}}{=} \inf\{n > s_{k-1} : Y_n^{\mathcal{C}} = v\}$$

and

$$j_k(v, \mu) \stackrel{\text{def}}{=} 1 + \text{number of times } \mathbf{Y}^{\mathcal{C}} \text{ jumped from } v \text{ to its neighbor } \mu \text{ by time } s_k.$$

The first jump after  $s_k$  is toward the neighbor  $\mu$  for which the following minimum,

$$(2.6) \quad \min_{\eta \in N_v^C} \frac{h_{j_k}(v, \eta)}{\omega(v, \eta)},$$

is a.s. attained. Due to the memoryless property, the probability that the minimum in (2.6) is achieved by the vertex  $\mu$  is still  $\omega(v, \mu) / \sum_{\eta \in N_v^C} \omega(v, \eta)$ .

With a slight abuse of notation, we denote the quenched and annealed law of the extension  $\mathbf{Y}^C$  again by  $\mathbf{P}_{\cdot, \omega}$ , respectively,  $\mathbf{P}$ .

REMARK 2.2. The extension processes will play a crucial role in our proofs. They are coupled to the original process  $\mathbf{X}$  in the following sense. The behavior of  $\mathbf{Y}^C$  is determined by the same exponentials used to generate the jumps of  $\mathbf{X}$ . The process  $\mathbf{Y}^C$  and the process  $\mathbf{X}$ , only observed when it visits  $C$ , coincide up to the random time the latter process leaves forever this subgraph. Suppose  $C$  is finite. The utility of the extension process stands, as we will see later, in the fact that the process  $\mathbf{Y}^C$  visits, a.s., infinitely many times every vertex of this subgraph, while  $\mathbf{X}$ , which we assumed to be transient, does not. Fix a vertex  $v \in \text{Vert}(C)$ . The behavior of  $Y^C$  when it jumps from  $v$  for the  $i$ th time, with  $i \geq 1$ , would coincide with that of  $\mathbf{X}$ , if the latter reaches  $v$  at least  $i$  times. For a rigorous definition of restriction process see [5] or [10]. Extension processes were used in [7] to prove the strong law of large numbers for vertex jump-reinforced processes.

A child  $v^{(j)}$  of  $v$  is called a *first child* if it is a.s. the minimizer of

$$(2.7) \quad \min_{1 \leq i \leq b} \frac{h_1(v, v^{(i)})}{\omega(v, v^{(i)})} \quad \text{a.s.}$$

Let us now turn to the lower bound on the speed. Lyons and Pemantle [17] (see also Menshikov and Petritis [19]) established the following recurrence-transience dichotomy:

$$(2.8) \quad \mathbf{X} \text{ is transient} \quad \text{if} \quad \inf_{0 \leq t \leq 1} \mathbb{E}[A^t] > \frac{1}{b}, \text{ and recurrent otherwise.}$$

Our standing assumption is that the walk is transient. Gross [14] proves a strong law of large numbers

$$(2.9) \quad v \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{|X_n|}{n} \geq 0, \quad \mathbf{P}\text{-a.s.}$$

The natural question to ask now is in which cases  $v$  is positive. This question was answered recently in Aidékon [1] in the more general setting of Galton–Watson trees. In particular, if  $A$  is bounded, it turns out that  $v$  is always positive (see Theorem 1.1 in [1]). We will now derive a lower bound on the speed  $v$ . This should be compared with the bound provided in Proposition 1.1 of [3], that is,

$$(2.10) \quad v \geq \frac{1 - \mathbb{E}[1/(\sum_i A^{(i)})]}{1 + \mathbb{E}[1/(\sum_i A^{(i)})]},$$

which holds also for RWRE defined on Galton–Watson trees, with no leaves, under additional assumptions. Our approach is quite different from the one used by Aidékon and can be applied also to other processes, such as once reinforced random walks (see Section 2.2). Moreover, the lower bound we propose seems to work better than the one given in (2.10) when there is a strong dependence between the  $A^{(i)}$ 's (see the example following Theorem 2.12). For  $n \geq 1$ , we define

$$(2.11) \quad L(v, n) \stackrel{\text{def}}{=} \sum_{j=0}^n \mathbb{1}_{\{X_j=v\}} \quad \text{and} \quad L(v) \stackrel{\text{def}}{=} \sum_{j=0}^{\infty} \mathbb{1}_{\{X_j=v\}},$$

the number of visits to  $v$  by time  $n$ , respectively, the total number of visits. Here is the main result of this subsection.

PROPOSITION 2.3. *Recall  $\beta$  in (2.4). Under transience, it holds that*

$$(2.12) \quad v \geq \frac{1 - \beta}{\mathbf{E}[L(\rho)]} > 0, \quad \mathbf{P}\text{-a.s.}$$

Before proving Proposition 2.3, we provide first a lemma. Let

$$(2.13) \quad \Pi_k = \sum_{v: |v|=k} \mathbb{1}_{\{T(v) < \infty\}}$$

be the number of vertices visited at level  $k$ . We have

LEMMA 2.4. *Assume transience, that is,  $\beta < 1$ . Then  $\Pi_k$  is stochastically dominated by a geometric random variable with parameter  $1 - \beta$ .*

PROOF. One vertex at level  $k$  is visited for sure. Call this vertex  $\sigma_1$ . Notice that, after  $T(\sigma_1)$ , a necessary condition to visit a further vertex at level  $k$  is that the walk returns to the parent of  $\sigma_1$ . To obtain an upper bound for  $\Pi_k$ , we adopt the following strategy. If the walk returns to the parent of  $\sigma_1$ , we consider the extension  $Y^{(\sigma_1)}$  of  $\mathbf{X}$  to the subtree obtained by cutting the subtree  $\Lambda_{\sigma_1}$ . This ensures that the second visit at level  $k$  will be at a new vertex  $\sigma_2$ , different from  $\sigma_1$ . We repeat this procedure iteratively, and it clearly yields an upper bound on the number of vertices  $\sigma_i$  visited at level  $k$ . Each time a new vertex  $\sigma_i$  is visited, there is a chance of escape to infinity with annealed probability  $1 - \beta > 0$ , because of stationarity. Since all subtrees  $\Lambda_{\sigma_i}$  are disjoint, the trials of escape are independent. It follows that  $\Pi_k$  is dominated by a geometric with parameter  $1 - \beta$ . This ends the proof.  $\square$

PROOF OF PROPOSITION 2.3. Notice that

$$(2.14) \quad \lim_{n \rightarrow \infty} \frac{T_n}{n} = 1/v, \quad \mathbf{P}\text{-a.s.}$$

Label the vertices at level  $k$  by  $v_{k,1}, v_{k,2}, \dots, v_{k,b^k}$ . We have that for  $n \geq 1$ ,

$$(2.15) \quad \mathbf{E}[T_n] \leq 1 + \mathbf{E}[L(\overleftarrow{\rho})] + \sum_{k=0}^{n-1} \sum_{j=1}^{b^k} \mathbf{E}[L(v_{k,j})\mathbb{1}_{\{T(v_{k,j}) < \infty\}}].$$

Fix a vertex  $v$ , and define  $\tilde{L}(v)$  to be the total time spent in the vertex  $v$  by the extension of  $\mathbf{X}$  to  $\overleftarrow{\Lambda}_v$  started at  $v$ . Then  $L(v) \leq \tilde{L}(v)$ , and the law of  $\tilde{L}(v)$  under  $\mathbf{P}_v$  is equal to the law of  $L(\rho)$  under  $\mathbf{P}$ . Moreover, the random variables  $\tilde{L}(v)$  and  $\mathbb{1}_{\{T(v) < \infty\}}$  are independent under the annealed measure. We use independence, and then stationarity, and obtain that the sum on the right-hand side of (2.15) is smaller than

$$(2.16) \quad \sum_{k=0}^{n-1} \sum_{j=1}^{b^k} \mathbf{E}[\tilde{L}(v_{k,j})\mathbf{P}[T(v_{k,j}) < \infty]] = \mathbf{E}[L(\rho)] \sum_{k=0}^{n-1} \mathbf{E}[\Pi_k] \leq \mathbf{E}[L(\rho)] \frac{n}{1 - \beta},$$

where in the last step we used Lemma 2.4. Using (2.15) and (2.16), and by Fatou’s lemma, we obtain that  $\mathbf{P}$ -a.s.,

$$(2.17) \quad \lim_{n \rightarrow \infty} T_n/n \leq \liminf_{n \rightarrow \infty} \mathbf{E}[T_n/n] \leq \mathbf{E}[L(\rho)](1 - \beta)^{-1}.$$

The claim of the proposition follows now from (2.14).  $\square$

Our main task is now to derive upper bounds on  $\beta$  and on the expectation of  $L(\rho)$ .

2.1.1. *Estimates on the return probability  $\beta$ .* In the last section, we provided a lower bound in terms of the annealed return probability  $\beta$ . In this section, we will derive an upper bound on  $\beta$  in terms of the extinction probability  $\alpha$  of a certain branching process, in the spirit of Collevocchio [6]. This allows us to obtain an explicit lower bound on the speed.

Let us start by constructing the branching process.

DEFINITION 2.5 (Color scheme). Fix an integer  $\psi \geq 1$ , and denote with  $\mathbf{Y}(v, \mu)$  the extension of  $\mathbf{X}$  to the unique ray connecting the vertices  $\mu$  and  $v$ . We introduce the following color scheme. A vertex  $v$  at level  $\psi$  is colored if and only if the  $\mathbf{Y}(\overleftarrow{\rho}, v)$ , started at  $\rho$ , hits  $v$  before  $\overleftarrow{\rho}$ . A vertex  $v$  at level  $k\psi, k \geq 2$ , is colored if and only if:

- its ancestor at level  $(k - 1)\psi$ , say  $\mu$ , is colored, and
- $Y(\overleftarrow{\mu}, v)$ , started at  $\mu$ , hits  $v$  before  $\overleftarrow{\mu}$ .

All the other vertices are uncolored, and only vertices that are at a level  $k\psi, k \geq 1$ , can be colored.

Under the *annealed* measure, the number of colored vertices form a homogeneous branching process, since the offspring is each time determined by disjoint parts of the environment. We denote this branching process with  $\mathbf{Z}_\psi$ . We formulate the following:

PROPOSITION 2.6. *Denote with  $\alpha_\psi$  the extinction probability of  $\mathbf{Z}_\psi$ . Then  $\beta \leq \alpha_\psi$ . If, moreover,*

$$(2.18) \quad \sup_{n \geq 1} b^n \mathbb{E} \left[ \left( \sum_{r=1}^{n+1} \prod_{j=1}^{r-1} A_j^{-1} \right)^{-1} \right] > 1,$$

*then there is an integer  $\psi \geq 1$  such that  $\alpha_\psi < 1$ .*

REMARK 2.7. The condition in (2.18) is quite natural. In fact the expectation in the left-hand side measures the drift of the process on fixed paths joining the root to a vertex at level  $n$ . This quantity is multiplied by the number of possible paths and becomes a measure of the drift. If this quantity is large enough, we can estimate the annealed probability of never visiting the parent of the root, when the process is started at the root.

PROOF OF PROPOSITION 2.6. Let us show that  $\beta \leq \alpha_\psi$  in the case  $\alpha_\psi < 1$  (otherwise there is nothing to prove). Assume that  $\mathbf{Z}_\psi$  survives. Choose vertices  $\mu$  and  $\nu$  as in Definition 2.5. By Remark 2.2, the processes  $Y(\tilde{\mu}, \nu)$  and  $\mathbf{X}$  coincide, from the time the latter hits  $\mu$  until its last visit to the path connecting  $\tilde{\mu}$  to  $\nu$ . It follows that, if  $\mathbf{X}$  hits  $\nu$  before  $\mu$ , then so does  $Y(\tilde{\mu}, \nu)$ . If the branching process survives, then there exists at least one colored vertex at each level  $k\psi, k \geq 1$ . It means that each level  $k\psi, k \geq 1$  is hit before visiting to the parent of the root. Hence  $\{Z_\psi \text{ survives}\} \subseteq \{D = \infty\}$ , and  $\beta \leq \alpha_\psi$  follows. We choose a vertex  $\mu$ , and then a vertex  $\nu$  at level  $|\mu| + \psi$ . Then the extension  $\mathbf{Y}(\tilde{\mu}, \nu)$ , started at  $\mu$ , hits  $\nu$  before  $\tilde{\mu}$  with (annealed) probability

$$(2.19) \quad \mathbb{E} \left[ \left( \sum_{r=1}^{\psi+1} \prod_{j=1}^{r-1} A_j^{-1} \right)^{-1} \right],$$

where  $A_j, 1 \leq j \leq \psi$ , is an enumeration of the variables  $A$  along the ray connecting  $\mu$  to  $\nu$ . To prove this, we just need to use the fact that under the quenched probability the process is a birth-death process and then apply a standard result, which can be found in [12], page 296. Hence, by virtue of (2.18), we can choose  $\psi$  that

$$(2.20) \quad b^\psi \mathbb{E} \left[ \left( \sum_{r=1}^{\psi+1} \prod_{j=1}^{r-1} A_j^{-1} \right)^{-1} \right] > 1.$$



Notice that the left-hand side of the last display is the expected offspring of the branching process  $Z_\psi$ , so that we can choose  $\psi$  s.t.  $Z_\psi$  is supercritical. This finishes the proof of the proposition.  $\square$

PROPOSITION 2.8. *If  $\mathbb{E}[\log(bA)] > 0$ , then  $\alpha_\psi < 1$ .*

PROOF. We prove that  $\mathbb{E}[\log(bA)] > 0$  implies

$$(2.21) \quad \liminf_{n \rightarrow \infty} \mathbb{E} \left[ b^n \left( \sum_{r=1}^{n+1} \prod_{j=1}^{r-1} A_j^{-1} \right)^{-1} \right] = \infty,$$

which implies (2.18). To prove (2.21), in virtue of Fatou’s lemma, it is enough to prove that

$$(2.22) \quad \limsup_{n \rightarrow \infty} b^{-n} \sum_{r=1}^{n+1} \prod_{j=1}^{r-1} A_j^{-1} = 0, \quad \mathbb{P}\text{-a.s.}$$

To see this, rewrite the left-hand side of (2.22) as

$$(2.23) \quad \sum_{r=1}^{n+1} b^{r-n-1} \prod_{j=1}^{r-1} (bA_j)^{-1}.$$

The strong law of large numbers shows that for large  $r$ ,

$$(2.24) \quad \prod_{j=1}^r (bA_j)^{-1} \leq \exp\left(-\frac{1}{2}r\mathbb{E}[\log(bA)]\right), \quad \mathbb{P}\text{-a.s.}$$

Using this, it is standard to show that the expression in (2.23) tends to zero. The claim (2.22) now follows, and the proof is finished.  $\square$

DEFINITION 2.9. We denote with  $\mathbf{p} := \{p_k, k \in \{0, 1, \dots, b^\psi\}\}$  the offspring distribution of the branching process  $Z_\psi$ . The mean offspring is

$$(2.25) \quad m_\psi \stackrel{\text{def}}{=} \sum_{k=0}^{b^\psi} kp_k.$$

Proposition 2.8 implies that if  $\mathbb{E}[A^{-1}] \leq b$  and  $\mathbb{P}(A^{-1} = b) < 1$ , then there is  $\psi \geq 1$  such that  $m_\psi > 1$ .

2.1.2. *An explicit upper bound on the expectation of  $L(\rho)$ .* Our standing assumption in the remaining subsections is that

$$(2.26) \quad \text{we can find } \psi \geq 1 \text{ such that } \alpha_\psi < 1,$$

where we recall  $\alpha_\psi$  in Proposition 2.6. For  $p \geq 1, n \geq 1$ , we introduce the function

$$(2.27) \quad \theta(p, n) \stackrel{\text{def}}{=} \begin{cases} c_p b \mathbb{E} \left[ \left( 1 + \frac{1}{\sum_{i=1}^b A^{(i)}} \right)^p \right] n^{p-1} \frac{\mathbb{E}[A^{-p}]^n - 1}{\mathbb{E}[A^{-p}] - 1}, & \text{if } n \geq 2, \\ c_p \mathbb{E} \left[ \left( 1 + \frac{1}{\sum_{i=1}^b A^{(i)}} \right)^p \right], & \text{if } n = 1, \end{cases}$$

where the right-hand side is infinite if  $\mathbb{E}[A^{-p}] = \infty$ , and the constants  $c_p$  are introduced in Lemma A.1 in the Appendix. We have the following:

PROPOSITION 2.10. *If  $\mathbb{E}[A^{-p-\varepsilon}] < \infty$  for some  $p \geq 1$  and some  $\varepsilon > 0$ , then for all  $n \geq 1, \theta(p + \varepsilon, n) < \infty$ , and*

$$(2.28) \quad \begin{aligned} \mathbf{E}[L(\rho)^p] &\leq \theta(p + \varepsilon, 1)^{1/q} \\ &+ \sum_{n=2}^{\infty} \theta(p + \varepsilon, n)^{1/q} \alpha_\psi^{b^{n-2}/q'} \\ &\times \left( \sum_{i=1}^b (-1)^{i-1} \binom{b}{i} \alpha_\psi^{b^{n-2}(i-1)} \right)^{1/q'}, \end{aligned}$$

where  $q = 1 + \varepsilon/p$ , and  $q' = 1 + p/\varepsilon$  is the dual of  $q$ .

Before proving Proposition 2.10, we formulate an auxiliary result. We first introduce some notation. Fix  $n \geq 2$ . Choose  $b$  distinct vertices  $v_i, 1 \leq i \leq b$ , at level  $n$ , with different ancestors at level one. More precisely, we choose  $v_i$  with ancestor  $\vec{\rho}_i$  at level one, and call this set of vertices  $\mathcal{A}_n$ . We label the vertices on the ray connecting  $\vec{\rho}_i$  to  $v_i$  by  $\sigma_j^{(i)}, 1 \leq j \leq n$ , with  $\sigma_1^{(i)} = \vec{\rho}_i$  and  $\sigma_n^{(i)} = v_i$ . Denote with  $\Gamma_n$  the subtree composed by the root  $\rho$ , its parent  $\vec{\rho}$ , the vertices  $\sigma_j^{(i)}, 1 \leq j \leq n, 1 \leq i \leq b$ , and the edges connecting them. For  $n = 1, \Gamma_1$  is simply the subtree composed by the root and its children, with the edges connecting them and  $\mathcal{A}_1$  is the set of children of the root. We denote with  $\mathbf{Y}$  the extension of  $\mathbf{X}$  to  $\Gamma_n$ , and we introduce  $\tilde{T}_{\mathcal{A}_n} = \inf\{n \geq 0 : Y_n \in \mathcal{A}_n\}$ , and  $\tilde{T}(\rho) \stackrel{\text{def}}{=} \inf\{n \geq 1 : Y_n = \rho\}$ . We further define

$$\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n}) \stackrel{\text{def}}{=} \sum_{i=0}^{\infty} \mathbb{1}_{\{Y_i = \rho, i < \tilde{T}_{\mathcal{A}_n}\}}.$$

Recall  $\theta(p, n)$  in (2.27). We have the following:

PROPOSITION 2.11. *If  $\mathbb{E}[A^{-p}] < \infty$  for some  $p \geq 1$ , then*

$$(2.29) \quad \mathbf{E}[\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})^p] \leq \theta(p, n) < \infty.$$

PROOF. Fix  $n \geq 2$ . To escape from the root, the walk  $\mathbf{Y}$  has to jump to one of the children of the root, and then hit the set  $\mathcal{A}_n$  before returning to the root. Hence

$$(2.30) \quad q_\omega \stackrel{\text{def}}{=} \mathbf{P}_\omega(\tilde{T}_{\mathcal{A}_n} < \tilde{T}(\rho)) = \sum_{i=1}^b \omega(\rho, \rho^{(i)}) p_{i,\omega},$$

where

$$p_{i,\omega} = \left( \sum_{j=1}^n \prod_{k=1}^{j-1} \frac{\omega(\sigma_k^{(i)}, \sigma_{k-1}^{(i)})}{\omega(\sigma_k^{(i)}, \sigma_{k+1}^{(i)})} \right)^{-1}.$$

It follows that under the quenched measure,  $\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})$  is a geometric variable with parameter  $q_\omega$ . Hence, with the help of Lemma A.1 in the Appendix, we find that

$$(2.31) \quad \mathbf{E}[\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})^p] \leq c_p \mathbf{E}[q_\omega^{-p}].$$

It follows from (2.30), and by independence, that

$$(2.32) \quad \begin{aligned} \mathbf{E}[q_\omega^{-p}] &\leq \mathbf{E}\left[\left(\min_i p_{i,\omega}\right)^{-p} (1 - \omega(\rho, \overleftarrow{\rho}))^{-p}\right] \\ &= \mathbf{E}\left[\left(\min_i p_{i,\omega}\right)^{-p}\right] \mathbf{E}[(1 - \omega(\rho, \overleftarrow{\rho}))^{-p}]. \end{aligned}$$

We use that

$$(2.33) \quad \mathbf{E}\left[\left(\min_i p_{i,\omega}\right)^{-p}\right] = \mathbf{E}\left[\max_i p_{i,\omega}^{-p}\right] \leq \mathbf{E}[\sum_i p_{i,\omega}^{-p}] = b \mathbf{E}[p_{1,\omega}^{-p}],$$

and we find by (2.30), by Jensen’s inequality and by independence that

$$(2.34) \quad \mathbf{E}[p_{1,\omega}^{-p}] \leq n^{p-1} \sum_{j=1}^n \mathbf{E}[A^{-p}]^{j-1} = n^{p-1} \frac{\mathbf{E}[A^{-p}]^n - 1}{\mathbf{E}[A^{-p}] - 1}.$$

Now observe that

$$(2.35) \quad \mathbf{E}[(1 - \omega(\rho, \overleftarrow{\rho}))^{-p}] = \mathbf{E}\left[\left(1 + \frac{1}{\sum_i A^{(i)}}\right)^p\right],$$

and by collecting the results from (2.31) to (2.35), the claim of the proposition follows for  $n \geq 2$ . For  $n = 1$ , a similar (and simpler) argument shows the claim. This finishes the proof of the proposition.  $\square$

PROOF OF PROPOSITION 2.10. In the course of this proof, we denote with  $\mathbf{Y}^{(v)}$  the extension of  $\mathbf{X}$  to  $\overleftarrow{\Lambda}_v$ , and let

$$(2.36) \quad D^{(v)} \stackrel{\text{def}}{=} \inf\{n \geq 1 : Y_n^{(v)} = \overleftarrow{\rho}\} \quad \text{and} \quad C(v) \stackrel{\text{def}}{=} \{D^{(v)} = \infty\}.$$

Suppose that  $|v| \geq 1$  and  $C(v)$  holds. Then if the process visits  $v$  it will never return to  $\overleftarrow{v}$ , and in particular it will not increase the local time spent at the root  $\rho$ . Define

$$(2.37) \quad \begin{aligned} d = \inf\{k \geq 1 : \text{there are } b \text{ distinct vertices } v_1, \dots, v_b \\ \text{at level } k \text{ with different ancestors at level } 1 \text{ s.t.} \\ C(v_i) \text{ holds for all } 1 \leq i \leq b\}. \end{aligned}$$

Notice that  $d < \infty$ , a.s. On  $\{d = n\}$ , we choose  $b$  distinct vertices  $v_1, \dots, v_b$  at level  $n$  with different ancestors at level 1 s.t.  $C(v_i)$  holds for all  $1 \leq i \leq b$ , and in the notation used in Proposition 2.11, we denote this set of vertices with  $\mathcal{A}_n$ . Notice that

$$(2.38) \quad L(\rho)\mathbb{1}_{\{d=n\}} \leq \tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})\mathbb{1}_{\{d=n\}}.$$

With the help of (2.38), we infer that for  $q, q'$  as in the proposition

$$(2.39) \quad \begin{aligned} \mathbf{E}[L(\rho)^p] &\leq \sum_{n=1}^{\infty} \mathbf{E}[\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})^p, d = n] \\ &\leq \sum_{n=1}^{\infty} \mathbf{E}[\tilde{L}(\rho, \tilde{T}_{\mathcal{A}_n})^{pq}]^{1/q} \mathbf{P}[d = n]^{1/q'}, \end{aligned}$$

where in the last inequality we used Hölder’s inequality. Let us now estimate  $\mathbf{P}(d = n)$ . The events  $C(v)|_{|v|=n}$  are determined by disjoint parts of the environment, and are thus independent and identically distributed under the annealed measure. Fix  $n \geq 2$ . At level  $n - 1$ , there are  $b$  families of  $b^{n-2}$  vertices, each that have different ancestors at level one. If  $\{d = n\}$  holds, then the event  $C(\cdot)^c$  holds for all  $b^{n-2}$  vertices in at least one of these families of vertices at level  $n - 1$ . With  $\mathbf{P}(C(\cdot)) = 1 - \beta$ , it follows that

$$\mathbf{P}(d = n) \leq 1 - (1 - \mathbf{P}(C^c)^{b^{n-2}})^b = 1 - (1 - \beta^{b^{n-2}})^b,$$

and with Proposition 2.6, it follows that

$$(2.40) \quad \mathbf{P}(d = n) \leq 1 - (1 - \alpha_\psi^{b^{n-2}})^b = \alpha_\psi^{b^{n-2}} \sum_{i=1}^b (-1)^{i-1} \binom{b}{i} \alpha_\psi^{b^{n-2}(i-1)}.$$

Together with the trivial bound  $\mathbf{P}(d = 1) \leq 1$ , this finishes the proof of the proposition.  $\square$

2.1.3. *An explicit lower bound on the speed and an example.* Recall  $\alpha_\psi$  in Proposition 2.6. The Propositions 2.3, 2.6 and 2.10 (applied with  $p = \varepsilon = 1$ ) imply the following:

**THEOREM 2.12.** *Assume (2.26), and that  $\mathbb{E}[A^{-2}] < \infty$ . Then it holds  $\mathbf{P}$ -a.s. that*

$$\begin{aligned}
 v &\geq \frac{1 - \alpha_\psi}{\mathbb{E}[L(\rho)]} \\
 &\geq \frac{1 - \alpha_\psi}{\theta(2, 1)^{1/2} + \sum_{n=2}^\infty \theta(2, n)^{1/2} \alpha_\psi^{b^{n-2}/2} \left(\sum_{i=1}^b (-1)^{i-1} \binom{b}{i} \alpha_\psi^{b^{n-2}(i-1)}\right)^{1/2}} \\
 &> 0.
 \end{aligned}$$

*An example.* Let us now provide an explicit example on the regular binary tree (i.e.,  $b = 2$ ). We choose  $A^{(1)} = A^{(2)}$ , and we write  $A$  for a copy of  $A_1$ , respectively,  $A_2$ . We consider the following cases:

$$\mathbb{P}[A = a] = a, \quad \mathbb{P}[A = 1 - a] = 1 - a,$$

where  $a \in (0.5, 1)$ . We have

$$\mathbb{E}[\log(2A)] = a \log(2a) + (1 - a) \log(2(1 - a)).$$

The right-hand side equals 0 for  $a = 1/2$ , and its derivative is  $\log(a/(1 - a))$  which is positive for  $a \in (0.5, 1)$ . Hence

$$\mathbb{E}[\log(2A)] > 0 \quad \text{for } a \in (0.5, 1).$$

In virtue of Proposition 2.8 we have  $\alpha_\psi < 1$ , for some  $\psi \in \mathbb{N}$ , implying that the lower bound given in Theorem 2.12 is positive. In Table 1 we summarize the results obtained for the cases  $a = 0.8, 0.9$ . In each of these cases, we simulated 50,000 RWRE on the binary tree and estimated  $\alpha_\psi$ , which we plugged in the lower bound given in Theorem 2.12.

On the other hand,  $\mathbb{E}[1/(A^{(1)} + A^{(2)})] = \mathbb{E}[1/(2A^{(1)})] = 1$ . Hence the lower bound for the speed, given in Proposition 1.1 in the Introduction of [3],

$$\frac{1 - \mathbb{E}[1/(A^{(1)} + A^{(2)})]}{1 + \mathbb{E}[1/(A^{(1)} + A^{(2)})]},$$

equals 0 and gives no information in this case.

TABLE 1  
*Lower bound for the speed of some RWRE*

$a$	$\psi$	$\alpha_\psi$	$v \geq$
0.8	5	0.8826681	0.00009
0.9	5	0.7354767	0.00016

2.2. *Once edge-reinforced random walk.* Durrett, Kesten and Limic [13] prove transience and provide a law of large numbers with positive speed for once edge-reinforced random walk on a regular tree. However their methods do not give a lower bound for the speed that is always positive. Collevecchio [6] proves transience for this process defined on supercritical Galton–Watson trees. The same was proved, independently and with different methods by Dai [8]. In this section, we provide a lower bound on the speed by using a refinement of the methods from [6].

Let us first define the process. Fix  $\delta > 0$ , and denote with  $\{v, \mu\}$  the edge connecting the neighboring vertices  $v$  and  $\mu$ . Once  $\delta$ -edge-reinforced random walk [ORRW( $\delta$ ) or simply ORRW]  $\mathbf{X} = \{X_k, k \geq 0\}$  is a discrete-time process on the regular  $b$ -ary tree  $\mathcal{G}_b$ , and is defined as follows. Each edge has initial weight one, that is,  $W(\{v, \mu\}, 0) = 1$ , with the exception of the edge  $\{\overleftarrow{\rho}, \rho\}$ , which has weight  $\delta$ , that is,  $W(\{\overleftarrow{\rho}, \rho\}, 0) = \delta$ . This exception helps to simplify our exposition. This initial weight configuration is called *initially fair*. For  $n \geq 1$ , we update the weight  $W$  of the edges according to the following rule:

$$(2.41) \quad W(\{v, \mu\}, n) = \begin{cases} \delta, & \text{if } \{X_{k-1}, X_k\} = \{v, \mu\} \text{ for some } 1 \leq k \leq n, \\ 1, & \text{otherwise.} \end{cases}$$

ORRW starts from  $\rho$ , that is,  $X_0 = \rho$ , and we define inductively  $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ , and the transition probabilities

$$(2.42) \quad \mathbf{P}(X_{n+1} = \mu | \mathcal{F}_n) = \frac{W(\{X_n, \mu\}, n)}{\sum_{v: v \sim X_n} W(\{X_n, v\}, n)},$$

if  $\mu$  is a neighbor of  $X_n$ , and zero otherwise. The canonical law of this process is denoted with  $\mathbf{P}$ . Later on, we will also use the following initial weights, where not only the edge  $\{\overleftarrow{\rho}, \rho\}$  has weight  $\delta$ , but a connected collection of edges containing the edge  $\{\overleftarrow{\rho}, \rho\}$ , that is, if some edge has weight  $\delta$ , then each edge on the path connecting this edge to the root has weight  $\delta$ . We denote with  $\mathbb{W}$  the set of such initial weight configurations. Of course,  $\mathbb{W}$  contains the initially fair weights that we denote from now on with  $w_0$ . For  $w \in \mathbb{W}$  let  $w(\{v, \mu\})$  be the weight that  $w$  assigns to the edge  $\{v, \mu\}$ . For any weight configuration  $w \in \mathbb{W}$ , define  $W_w(\{v, \mu\}, 0) = w(\{v, \mu\})$ , and for  $n \geq 1$ ,

$$W_w(\{v, \mu\}, n) = \begin{cases} \delta, & \text{if } \{X_{k-1}, X_k\} = \{v, \mu\} \text{ for some } 1 \leq k \leq n, \\ w(\{v, \mu\}), & \text{otherwise.} \end{cases}$$

The transition probabilities are defined similarly as in (2.42), with  $W(\cdot, n)$  replaced by  $W_w(\cdot, n)$ . The canonical law of ORRW started at  $\rho$ , and in the initial weight configuration  $w \in \mathbb{W}$  is denoted with  $\mathbf{P}_w$  (clearly  $\mathbf{P} = \mathbf{P}_{w_0}$ ). Recall the exponential random variables  $h_k(\cdot, \cdot), k \geq 1$ , with mean one, used in Definition 2.1, and fix a subtree  $\mathcal{C}$  of  $\mathcal{G}_b$ .

DEFINITION 2.13 (Extension  $\mathbf{Y}^{\mathcal{C}}$  on the subtree  $\mathcal{C}$ ). The extension  $\mathbf{Y}^{\mathcal{C}}$  of  $\mathbf{X}$  on the subtree  $\mathcal{C}$  is defined as follows. Fix a starting point  $\eta$  in  $\mathcal{C}$ , that is,  $Y_0^{\mathcal{C}} = \eta$  and

an initial weight configuration  $w \in \mathbb{W}$ . We define  $\mathbf{Y}^C$  iteratively in the following way. Let  $s_1(v)$  be the first time  $\mathbf{Y}^C$  reaches some vertex  $v$ . Define  $N_v^C$  to be the set of neighbors of  $v$  in  $\mathcal{C}$ . The first jump after  $s_1(v)$  is toward the neighbor  $\mu \in N_v^C$  for which the following minimum,

$$(2.43) \quad \min_{\mu \in N_v^C} \frac{h_1(v, \mu)}{W_w(\{v, \mu\}, s_1(v))},$$

is a.s. attained. We define  $s_k(v), k \geq 2$ , inductively via

$$s_k(v) \stackrel{\text{def}}{=} \inf\{n > s_{k-1} : Y_n^C = v\}$$

and

$$j_k(v, \mu) \stackrel{\text{def}}{=} 1 + \text{number of times } \mathbf{Y}^C \text{ jumped from } v \text{ to its neighbor } \mu \text{ by time } s_k.$$

The first jump after  $s_k(v)$  is toward the neighbor  $\mu$  for which the following minimum,

$$(2.44) \quad \min_{\mu \in N_v^C} \frac{h_{j_k}(v, \mu)}{W_w(\{v, \mu\}, s_1(v))},$$

is a.s. attained. For any vertex  $v$ , enumerate its children with  $v^{(j)}$ , with  $j \in \{1, 2, \dots, b\}$ . A child  $v^{(j)}$  of  $v$  is called a *first child* if it is a.s. the minimizer of

$$(2.45) \quad \min_{1 \leq i \leq b} \frac{h_1(v, v^{(i)})}{W_w(\{v, v^{(i)}\}, 1)} \quad \text{a.s.}$$

The comments in Remark 2.2 also apply here. We now introduce a similar color scheme as in Definition 2.5.

DEFINITION 2.14. Fix an integer  $\psi \geq 1$ , and denote with  $\mathbf{Y}(\overleftarrow{\mu}, v)$ , for a descendant  $v$  of  $\mu$ , the extension of ORRW on the ray connecting  $\overleftarrow{\mu}$  to  $v$ , started at  $\mu$ , in the following initial weight configuration. The edge  $\{\overleftarrow{\mu}, \mu\}$  has weight  $\delta$ , and all the other edges in the path connecting  $\mu$  to  $v$  have initial weight 1. A vertex  $v$  at level  $\psi$  is colored if and only if  $\mathbf{Y}(\overleftarrow{\rho}, v)$  hits  $v$  before  $\overleftarrow{\rho}$ . A vertex  $v$  at level  $k\psi, k \geq 2$ , is colored if and only if:

- its ancestor at level  $(k - 1)\psi$ , say  $\mu$ , is colored, and
- $Y(\overleftarrow{\mu}, v)$  hits  $v$  before  $\overleftarrow{\mu}$ .

All the other vertices are uncolored, and only vertices that are at a level  $k\psi, k \geq 1$ , can be colored.

This color scheme constitutes again a homogeneous branching process, with extinction probability  $\alpha_\psi$ . Notice that for every  $b \geq 2$ , and every  $\delta > 0$ , we can always find an integer  $\psi \geq 1$  such that

$$(2.46) \quad b^\psi \prod_{j=1}^\psi \frac{j}{j + \delta} > 1.$$

We define  $D$  in the same way as in (2.3), and also  $\beta_w = \mathbf{P}_w(D = \infty)$ , and we write  $\beta = \beta_{w_0}$ . Recall  $\mathbb{W}$  below (2.42). We have the following:

**PROPOSITION 2.15.** *If  $\psi$  is such that (2.46) holds, then  $\alpha_\psi < 1$ . If  $\delta > 1$ , then for every  $w \in \mathbb{W}$ , it holds that  $\beta_w \leq \alpha_\psi$ .*

**PROOF.** The probability that  $\mathbf{Y}(\overleftarrow{\rho}, \nu)$ , started at  $\rho$ , in the initially fair weight configuration  $w_0$ , hits level  $\psi$  before it hits  $\overleftarrow{\rho}$  is equal to (see Lemma 1 in [6])

$$(2.47) \quad \prod_{j=1}^\psi \frac{j}{j + \delta}.$$

Hence the mean of the offspring distribution of the colored process is equal to  $b^\psi \prod_{j=1}^\psi \frac{j}{j + \delta}$ , which is larger than one by our choice of  $\psi$ . This shows that  $\alpha_\psi < 1$ . Now choose an initial weight configuration  $w \in \mathbb{W}$ . If  $\delta > 1$ , we can couple the extension  $\mathbf{Y}(\overleftarrow{\rho}, \nu)$ , started at  $\rho$ , in the initially fair weight configuration  $w_0$ , to the extension  $\tilde{\mathbf{Y}}(\overleftarrow{\rho}, \nu)$ , started at  $\rho$ , in the weight configuration  $w \in \mathbb{W}$ , in such a way that  $|\tilde{\mathbf{Y}}| \geq |\mathbf{Y}|$ . To do this, we choose a family of independent variables  $(E_n^\uparrow, E_n^\downarrow)_{n \geq 1}$ , with i.i.d. exponential entries with mean 1. At each time point  $n$ , the vector  $(E_n^\uparrow, E_n^\downarrow)$  is attached *both* to the positions  $Y_n$  and  $\tilde{Y}_n$ , with  $E_n^\uparrow$  attached to the edge connecting  $Y_n$  and  $\tilde{Y}_n$  to the vertex  $\nu$  at level  $|Y_n| + 1$ , respectively,  $\tilde{\nu}$  at level  $|\tilde{Y}_n| + 1$ , and  $E_n^\downarrow$  attached to the edge connecting  $Y_n$  and  $\tilde{Y}_n$  to the vertex  $\mu$  at level  $|Y_n| - 1$ , respectively,  $\tilde{\mu}$  at level  $|\tilde{Y}_n| - 1$ . The jump of  $\mathbf{Y}$  at time  $n + 1$  is to the vertex  $\nu$  or  $\mu$  for which the minimum

$$(2.48) \quad \min \left\{ \frac{E_n^\uparrow}{W_{w_0}(\{Y_n, \nu\}, n)}, \frac{E_n^\downarrow}{W_{w_0}(\{Y_n, \mu\}, n)} \right\}$$

is a.s. attained, and similarly for  $\tilde{\mathbf{Y}}$ , where we replace the weights  $W_{w_0}$  by  $W_w$ , and the vertices  $\nu, \mu$  by  $\tilde{\nu}, \tilde{\mu}$ . Notice that in this way the extensions  $\mathbf{Y}$  and  $\tilde{\mathbf{Y}}$  have the same distribution as in Definition 2.13. Let

$$r = \inf\{n \geq 1 : |Y_n| \neq |\tilde{Y}_n|\}$$

be the first splitting time, and for ease of notation, let  $e_0, e_1$  be the two edges incident to  $Y_{r-1} = \tilde{Y}_{r-1}$ , where  $e_1$  connects  $Y_{r-1}$  to its child on the path, and  $e_0$  connects  $Y_{r-1}$  to its parent  $\tilde{Y}_{r-1}$ . Clearly  $W_w(e_0, r - 1) = W_{w_0}(e_0, r - 1) = \delta$ ,



since the edge  $e_0$  is crossed by both processes. Also, by construction,  $W_w(e, r - 1) \geq W_{w_0}(e, r - 1)$  for any edge  $e$  lying on the path connecting  $\overleftarrow{\rho}$  to  $v$ . If we would have  $W_w(e_1, r - 1) = W_{w_0}(e_1, r - 1)$ , then, by the construction of the coupling in (2.48),  $Y_r = \tilde{Y}_r$ , a contradiction. Hence  $W_w(e_1, r - 1) = \delta$  and  $W_{w_0}(e_1, r - 1) = 1$ . It follows again from (2.48) that the only way  $\mathbf{Y}$  and  $\tilde{\mathbf{Y}}$  can split is that  $|Y_r| = |\tilde{Y}_r| + 2$ . Define

$$s = \inf\{n > r : |Y_n| = |\tilde{Y}_n|\}.$$

For any edge  $e$  lying on the path connecting  $\overleftarrow{\rho}$  to  $v$ , we have that  $W_w(e, s) \geq W_{w_0}(e, s)$ , and we can reiterate the previous argument to prove that  $|\tilde{\mathbf{Y}}| \geq |\mathbf{Y}|$ . Consider the coloring process, defined in the same way as above (2.46), but on the weight configuration  $w$ . It follows that, if the coloring process associated to  $\mathbf{Y}$  survives, then as  $|\tilde{\mathbf{Y}}| \geq |\mathbf{Y}|$ , the coloring process associated to  $\tilde{\mathbf{Y}}$  survives. But on this last event,  $D = \infty$ . Hence  $\beta_w = \mathbf{P}_w(D < \infty) \leq \alpha_\psi$ .  $\square$

The random variable  $L(\cdot)$  is defined in the same way as in (2.11). We have the following:

**PROPOSITION 2.16.** *If  $\delta > 1$ , under  $\mathbf{P}_{w_0}$ , the random variable  $L(\rho)$  is stochastically dominated by a geometric variable with parameter  $(1 - \alpha_\psi)b/(b + \delta)$ .*

**PROOF.** Recall that  $\mathbf{X}$  starts from  $\rho$  in the initially fair weight configuration  $w_0$ . With probability  $b/(b + \delta)$  the first jump will be toward one of the children of  $\rho$ . Then, started at this child of  $\rho$ , with probability  $1 - \beta$ , the process will never return to  $\rho$ . Whenever it returns to  $\rho$ , it starts on some random weight configuration  $w \in \mathbb{W}$ , depending on the past of the path. Under  $\mathbf{P}_w$ , the probability that ORRW jumps to one of the children of  $\rho$  is greater than  $b/(b + \delta)$ . To see this, recall that the edge  $\{\overleftarrow{\rho}, \rho\}$  has weight  $\delta$ , and we change all the weights on the edges connecting  $\rho$  to its children to one. Since  $\delta > 1$ , this decreases the probability to jump to level one, and we obtain the lower bound for this probability. Under  $\mathbf{P}_w$ , ORRW, started at a child  $v$  of  $\rho$ , has probability larger than  $1 - \beta_{\bar{w}}$  of never returning to  $\rho$ , where  $\bar{w}$  is the weight configuration induced by  $w$  on  $\overleftarrow{\Lambda}_v$ . With the help of Proposition 2.15, we find that, for any  $w \in \mathbb{W}$ , the escape probability from  $\rho$  is at least  $(1 - \alpha_\psi)b/(b + \delta)$ , and it follows that the number of returns to  $\rho$  is stochastically dominated by a geometric variable with parameter  $(1 - \alpha_\psi)b/(b + \delta)$ .  $\square$

We recall from [13] that a law of large numbers with positive speed holds, that is,  $\mathbf{P}$ -a.s.,  $v = \lim_{n \rightarrow \infty} |X_n|/n > 0$ . Further, it is shown that  $v \leq b/(b + \delta)$ , but no lower bound is available. We are now ready to provide a lower bound for the speed that is always positive.

TABLE 2  
Lower bound for the speed of some ORRW

$b$	$\delta$	$\psi$	$\alpha_\psi$	$v \geq$
2	2	9	0.7858936	0.02292078
2	2.1	10	0.7860250	0.02289265
2	2.5	10	0.8697370	0.008484225
7	7	6	0.6798824	0.05123764

THEOREM 2.17. *If  $\delta > 1$ , choose  $\psi \geq 1$  such that (2.46) holds. Then the speed  $v$  satisfies*

$$(2.49) \quad v \geq \frac{1 - \beta}{\mathbf{E}[L(\rho)]} \geq (1 - \alpha_\psi)^2 \frac{b}{b + \delta} > 0.$$

REMARK 2.18. Notice that in the case of  $\delta < b$  we can compare  $|\mathbf{X}|$  with a simple random walk on the nonnegative integers with drift equal to  $(b - \delta)/(b + \delta) > 0$ . It follows that for  $\delta < b$  we have  $v \geq (b - \delta)/(b + \delta)$ . In this case, we find that the lower bound in (2.49) is larger than  $(b - \delta)/(b + \delta)$  if and only if  $\alpha_\psi < 1 - \sqrt{1 - \delta/b}$ . The challenging case is  $\delta \geq b$ , which is covered by Theorem 2.17. In Table 2 we summarize our simulation results for ORRW on regular trees, with  $b = 2$  and  $b = 7$ , with  $\delta = 2, 2.1, 2.5, 7$ .

PROOF OF THEOREM 2.17. Define the random variable  $\Pi_k$  in the same way as in (2.13). Observe that the same result as Lemma 2.4 in the previous section holds, with exactly the same proof. By straightforward modifications, we further see that Proposition 2.3 holds in the setting of once edge-reinforced random walk. The first inequality follows. The second and third inequalities then follow directly from Propositions 2.16 and 2.15.  $\square$

Next we show monotonicity of the lower bound on the speed in (2.49).

PROPOSITION 2.19. *Choose  $\delta_2 > \delta_1 \geq 1$ . Then for every  $\psi \geq 1$ ,  $\alpha_\psi(\delta_1) \leq \alpha_\psi(\delta_2)$ , and in particular the lower bound in (2.49) is decreasing in  $\delta$  for  $\delta > 1$ .*

PROOF. Denote with  $\mathbf{Y}^{(1)}$  and  $\mathbf{Y}^{(2)}$  the extensions on rays  $[\overleftarrow{\rho}, \infty)$  corresponding to  $\text{ORRW}(\delta_1)$ , respectively,  $\text{ORRW}(\delta_2)$ , started at  $\rho$ , in the initially fair weight configuration  $w_0^{(\delta_1)}$ , respectively,  $w_0^{(\delta_2)}$ . Using the same coupling as in (2.48), we can show that  $|\mathbf{Y}^{(1)}| \geq |\mathbf{Y}^{(2)}|$ . To see this, call  $r$  to be the first time the two processes split, and let  $e_0$  and  $e_1$  be as in the proof of Proposition 2.15. Next we show that none of the processes traversed edge  $e_1$  by time  $r - 1$ . In fact, as the two processes coincide up to time  $r - 1$ , if one of them traversed  $e_1$ , also the

other did. On the other hand, both of them traversed  $e_0$  by time  $r - 1$ , in order to reach  $Y_{r-1}^{(1)} = Y_{r-1}^{(2)}$ . Hence  $\mathbf{P}(|Y_r^{(1)}| = |Y_{r-1}^{(1)}| + 1) = 1/2 = \mathbf{P}(|Y_r^{(2)}| = |Y_{r-1}^{(2)}| + 1)$ . By construction of the coupling, this would imply that  $Y_r^{(1)} = Y_r^{(2)}$ , which contradicts the definition of  $r$ . As none of the processes traversed edge  $e_1$  by time  $r - 1$ , while both traversed  $e_0$ , using the fact  $\delta_2 > \delta_1$  we infer that  $|Y_r^{(1)}| > |Y_r^{(2)}|$ . Denote with  $t$  the first time, after  $r$ , when the two processes meet, and let  $r_1$  be the first time after  $t$ , when the two processes split again. As  $|Y_k^{(1)}| \geq |Y_k^{(2)}|$  for all  $k \leq r_1 - 1$ , we have that there is no edge reinforced by  $Y_k^{(2)}$  which has not been reinforced by  $Y_k^{(1)}$ ,  $k \leq r_1 - 1$ . This, together with the fact that  $\delta_1 > \delta_2 > 1$ , and the construction of the coupling, implies that  $\mathbf{P}(|Y_{r_1}^{(1)}| = |Y_{r_1-1}^{(1)}| + 1) \geq \mathbf{P}(|Y_{r_1}^{(2)}| = |Y_{r_1-1}^{(2)}| + 1)$ . By construction of the coupling, we have that  $|Y_{r_1}^{(1)}| > |Y_{r_1}^{(2)}|$ . By reiterating this argument, we get  $|\mathbf{Y}^{(1)}| \geq |\mathbf{Y}^{(2)}|$ . This implies that for every  $\psi$ ,  $\alpha_\psi(\delta_1) \leq \alpha_\psi(\delta_2)$ , and it follows that the lower bound in (2.49) is decreasing in  $\delta$ .  $\square$

**3. Moment bounds on the first regeneration time.** In addition to providing an explicit lower bound on the speed, our methods can be extended to give an explicit upper bound on the tail of a certain regeneration level. We present a unified approach that applies both for random walk in a random environment and once edge-reinforced random walk. Hence, in what follows,  $\mathbf{X}$  denotes either one of these processes. We start by defining the regeneration times.

DEFINITION 3.1. We define the first regeneration level as follows:

$$\ell_1 \stackrel{\text{def}}{=} \inf\{k \geq 1 : D(X_{T_k}) = \infty\},$$

and iteratively

$$\ell_n \stackrel{\text{def}}{=} \inf\{k > \ell_{n-1} : D(X_{T_k}) = \infty\},$$

where  $D(\cdot)$  is defined in (2.3) and we use the convention  $\inf \emptyset = \infty$ . The regeneration times are defined as  $\tau_n = T_{\ell_n}$ ,  $n \geq 1$ , on the event  $\{\ell_n < \infty\}$ .

In other words, a regeneration time occurs when the walk hits a level for the first time and then never backtracks to the previous level. Clearly, these are not stopping times. It is easy to see that under transience, it holds that for all  $n \geq 1$ ,  $\tau_n < \infty$ ,  $\mathbf{P}$ -a.s. It is also known that in the setting of random walks in random environment, the first regeneration level  $\ell_1$  has exponential moments under the conditioned measure  $\mathbf{P}(\cdot | D = \infty)$ . This is, for instance, proved in Lemma 4.2 in [11] for biased random walks on Galton–Watson trees, and can be directly adapted to our setting. For once edge-reinforced random walk with  $\delta > 1$ , we know that  $\ell_1$  has all moments finite under  $\mathbf{P}(\cdot | D = \infty)$ , see Lemma 7 in [13] (this statement is actually proved for certain cut levels, but notice that our regeneration level is smaller than the cut level in [13]).

We now present a unified approach that applies to both settings and that provides explicit estimates for the tail of  $\ell_1$  and for the moments of  $\tau_1$ .

3.1. *The tail of the first regeneration level.* We assume that we can choose  $\psi$  such that (2.26) is fulfilled for random walk in a random environment, respectively, once edge-reinforced random walk. Recall that for ORRW( $\delta$ ), this is always possible [see (2.46) and Proposition 2.15].

We will find *explicit* exponential tails on  $\ell_1$ . These tail estimates on  $\ell_1$  are obtained by refining the color scheme from Definitions 2.5, respectively, 2.14.

DEFINITION 3.2. Recall the definition of first child from (2.7) and (2.45). Recall also that for any pair of vertices  $\mu_1$  and  $\mu_2$ , we denote by  $\mathbf{Y}(\mu_1, \mu_2)$  the extension of the process  $\mathbf{X}$  on the unique shortest path connecting the two vertices. Let  $v$  be a vertex at level  $k\psi$ ,  $k \geq 1$ , and fix a positive integer  $\zeta$ . Let  $\Theta_v = \Theta_v(\zeta)$  be the set of vertices  $\mu$  in  $\Lambda_v$  which are first children and whose distance from  $v$  is a multiple of  $\zeta\psi$ . We proceed with the following labelling of vertices. A vertex  $\mu$ , descendent of  $v$  and at a level multiple of  $\zeta\psi$ , is *good* if:

- its ancestor in  $\Lambda_v$  which is at level multiple of  $(k - 1)\zeta\psi$ , say  $\mu_0$ , is good, and
- $Y(\overleftarrow{\mu}_0, \mu)$ , started at  $\mu_0$ , hits  $\mu$  before  $\overleftarrow{\mu}_0$ .

We label  $v$  as good. Let  $\Sigma_v$  be the set of vertices  $\mu$  in  $\Lambda_v$  such that:

- $\mu$  is good (in particular  $|\mu| - |v|$  is a multiple of  $\psi$ ), and
- all ancestors of  $\mu$  in  $\Lambda_v$  do not belong to  $\Theta_v$ .

Further define  $B(v) = \{\Sigma_v \text{ is infinite}\}$ , and  $B_0 = B(\rho)$ ,  $B_i = B(X_{T_{\psi\zeta i}})$ ,  $i \geq 1$ .

In other words,  $\Sigma_v$  is the set of colored vertices in  $\Lambda_v$  minus the colored vertices that are elements of subtrees generated by vertices  $\mu$  that are first children and  $|\mu| - |v| = k\zeta\psi$ ,  $k \geq 1$ . In a first step, we introduce an auxiliary branching process and use it to derive an explicit lower bound on the probability of  $B_0$  (see Lemma 3.3). In a second step, in Lemma 3.4, we then show that the events  $B_i$  are independent. In [7], Section 3, the counterpart of these lemma for vertex-reinforced jump processes are stated and proved in a similar way.

For any pair of distributions  $f_1$  and  $f_2$ , denote by  $f_1 \overline{*} f_2$  the distribution of  $\sum_{k=1}^V M_k$ , where:

- $V$  has distribution  $f_1$ , and
- $\{M_k, k \in \mathbb{N}\}$  is a sequence of i.i.d. random variables, independent of  $V$ , each with distribution  $f_2$ .

Recall the definition of  $p$  for random walk in random environment from Definition 2.9, and use the same definition for once-edge reinforced random walk. We set  $\mathbf{p}^{(1)} := \mathbf{p}$ , and define, by recursion,  $\mathbf{p}^{(j)} := \mathbf{p}^{(j-1)} \overline{*} \mathbf{p}$  for  $j \geq 2$ . The distribution

$\mathbf{p}^{(j)}$  describes the number of elements, at time  $j$ , in a population which evolves like a branching process generated by one ancestor and with offspring distribution  $\mathbf{p}$ . Let  $q_0 = p_0 + p_1$ , and for  $k \in \{1, \dots, b^\psi - 1\}$ , set  $q_k = p_{k+1}$ . Set  $\mathbf{q}$  to be the distribution which assigns to  $i \in \{0, \dots, b^\psi - 1\}$  probability  $q_i$ . For  $j \geq 2$ , let  $\mathbf{q}^{(j)} := \mathbf{p}^{(j-1)} \bar{*} \mathbf{q}$ . Denote by  $q_i^{(j)}$  the weight that the distribution  $\mathbf{q}^{(j)}$  assigns to  $i \in \{0, \dots, (b^\psi - 1)b^{(j-1)\psi}\}$ . The mean of  $\mathbf{q}^{(j)}$  is  $m_\psi^{j-1}(m_\psi - 1)$ . From now on,  $\zeta$  denotes the smallest positive integer such that

$$(3.1) \quad m_\psi^{\zeta-1}(m_\psi - 1) > 1.$$

(This is possible since we chose  $\psi$  such that  $m_\psi > 1$ .) Define  $\gamma$  to be the smallest positive solution of the equation

$$(3.2) \quad x = \sum_{k=0}^{\vartheta} x^k q_k^{(\zeta)} \quad \text{where } \vartheta = b^{(\zeta-1)\psi}(b^\psi - 1).$$

LEMMA 3.3. *Assume (3.1). We have that for  $i \geq 0$ ,  $\mathbf{P}(B_i) = \mathbf{P}(B_0) \geq 1 - \gamma > 0$ .*

PROOF. Fix  $i$  and notice that by stationarity,  $\mathbf{P}(B_i) = \mathbf{P}(B_0)$ . From the definition of  $\Sigma_\rho$ , it follows that the offspring distribution of colored vertices at level  $\zeta\psi$  in  $\Sigma_\rho$  is obtained as follows. The number of vertices at level  $(\zeta - 1)\psi$  has law  $\mathbf{p}^{(\zeta-1)}$ . Each vertex at level  $\zeta\psi$  has a number of colored offspring distributed as  $\mathbf{p} = \mathbf{p}^{(1)}$ . If from each of these offspring we delete the first child, the number of the remaining colored offspring is distributed as  $\mathbf{q}$ . Hence the offspring distribution modeling  $\Sigma_\nu$  is given by  $\mathbf{q}^{(\zeta)} = \mathbf{p}^{(\zeta-1)} \bar{*} \mathbf{q}$ . Then, from the basic theory of branching processes we know that the extinction probability equals the smallest positive solution of equation (3.2). In virtue of (3.1) we have that  $\gamma < 1$ .  $\square$

LEMMA 3.4. *The events  $B_i, i \geq 1$ , are independent under  $\mathbf{P}$ .*

PROOF. Choose integers  $0 < i_1 < i_2 < \dots < i_k$ . It is enough to prove that

$$(3.3) \quad \mathbf{P}\left(\bigcap_{j=1}^k B_{i_j}\right) = \prod_{j=1}^k \mathbf{P}(B_{i_j}).$$

We proceed by backward recursion. We use the notation introduced in Definition 2.1. The set  $B(\nu)$  belongs to the sigma-algebra generated by  $\{h_i(\eta, \mu) : \eta \in \text{Vert}(\Lambda_\nu), \mu \in \text{Vert}(\overleftarrow{\Lambda}_\nu) \text{ and } i \geq 1\}$ . Notice that each  $X_{T_i}, i \geq 1$ , is a first child. Hence the set  $\bigcap_{j=1}^{k-1} B_{i_j} \cap \{X_{T_{\psi\zeta i_k}} = \nu\}$  belongs to  $\{h_i(\eta, \mu) : \text{either } \eta \notin \text{Vert}(\Lambda_\nu) \text{ or } \mu \notin \text{Vert}(\overleftarrow{\Lambda}_\nu)\}$ . As the two events belong to disjoint collections of independent

exponential variables, they are independent. We have

$$\begin{aligned} \mathbf{P}\left(\bigcap_{j=1}^k B_{i_j}\right) &= \sum_{\nu} \mathbf{P}\left(B_{i_k} \cap \bigcap_{j=1}^{k-1} B_{i_j} \cap \{X_{T_{\psi\zeta i_k}} = \nu\}\right) \\ &= \sum_{\nu} \mathbf{P}(B(\nu)) \mathbf{P}\left(\bigcap_{j=1}^{k-1} B_{i_j} \cap \{X_{T_{\psi\zeta i_k}} = \nu\}\right). \end{aligned}$$

From stationarity, it follows that  $\mathbf{P}(B(\nu)) = \mathbf{P}(B_0)$ , and from the independence of  $B(\nu)$  and  $\{X_{T_{i\psi\zeta}} = \nu\}$ , we infer that for an arbitrary vertex  $\nu$ , and each  $i \geq 1$ ,

$$(3.4) \quad \mathbf{P}(B(\nu)) = \mathbf{P}(B_i).$$

Now the right-hand side of (3.1) equals

$$(3.5) \quad \mathbf{P}(B_0) \sum_{\nu} \mathbf{P}\left(\bigcap_{j=1}^{k-1} B_{i_j} \cap \{X_{T_{\psi\zeta i_k}} = \nu\}\right) = \mathbf{P}(B_{i_k}) \mathbf{P}\left(\bigcap_{j=1}^{k-1} B_{i_j}\right).$$

Equation (3.3) follows now by iteration.  $\square$

**THEOREM 3.5.** *Assume (2.26). For  $n \geq 1$ , we have that*

$$(3.6) \quad \mathbf{P}(\ell_1 \geq n\psi\zeta) \leq \gamma^{n-1},$$

where  $\gamma$  is defined in (3.2).

**PROOF.** Notice that on the event  $B_i$ , the colored process survives in the subtree  $\Lambda_{X_{T_{i\psi\zeta}}}$ . It follows that  $B_i \subseteq \{\text{level } i\psi\zeta \text{ is a regeneration level}\}$ . Hence

$$\{\ell_1 \geq n\psi\zeta\} \subseteq \bigcap_{i=1}^{n-1} B_i^c,$$

and the theorem now follows from the Lemmas 3.3 and 3.4.  $\square$

**3.2. Moment bounds for the first regeneration time.** Recall the first regeneration time in Definition 3.1, and define

$$\Pi = \sum_{\nu \in \mathcal{G}_b} \mathbb{1}_{\{T(\nu) \leq \tau_1\}}$$

to be the number of distinct vertices visited by time  $\tau_1$ . We denote with  $M(n, q)$  the  $n$ th moment of a geometric variable with parameter  $q$ . We have the following explicit bound on the moments of  $\Pi$ , which implies an explicit bound on the moments of  $\tau_1$  (see Theorem 3.7 below).

PROPOSITION 3.6. Assume (2.26). For  $p \geq 1$ , it holds that

$$(3.7) \quad \mathbf{E}[\Pi^p] \leq 1 + \gamma^{-1}(1 - \gamma^{1/(2\psi\zeta)})^{-1}(M(p, 1 - \gamma^{1/(2\psi\zeta)}) - 1) \times M^{1/2}(2p, 1 - \beta).$$

PROOF. Recall  $\Pi_k = \sum_{v: |v|=k} \mathbb{1}_{\{T(v) < \infty\}}$ ,  $k \geq -1$ , which is the number of vertices visited at level  $k$ , and observe that

$$(3.8) \quad \begin{aligned} \Pi &= 1 + \sum_{n=1}^{\infty} \sum_v \mathbb{1}_{\{T(v) \leq T_n\}} \mathbb{1}_{\{\ell_1=n\}} \leq \sum_{n=1}^{\infty} \sum_{v: |v|<n} \mathbb{1}_{\{T(v) < \infty\}} \mathbb{1}_{\{\ell_1=n\}} \\ &= \sum_{n=1}^{\infty} \sum_{k=-1}^{n-1} \Pi_k \mathbb{1}_{\{\ell_1=n\}}. \end{aligned}$$

We use Jensen’s inequality, and obtain that

$$(3.9) \quad \begin{aligned} \mathbf{E}[\Pi^p] &= \sum_{n=1}^{\infty} \mathbf{E} \left[ \left( \sum_{k=-1}^{n-1} \Pi_k \right)^p \mathbb{1}_{\{\ell_1=n\}} \right] \\ &\stackrel{\text{(Jensen)}}{\leq} \sum_{n=1}^{\infty} (n+1)^{p-1} \sum_{k=-1}^{n-1} \mathbf{E}[\Pi_k^p \mathbb{1}_{\{\ell_1=n\}}]. \end{aligned}$$

First notice that Lemma 2.4, proved for random walk in a random environment, holds also for once edge-reinforced random walk with the same proof. We first use Cauchy–Schwarz’s inequality, and then Lemma 2.4 together with Lemma A.1 from the Appendix to obtain that the right-hand side of the last display is smaller than

$$(3.10) \quad \begin{aligned} &\sum_{n=1}^{\infty} (n+1)^{p-1} \sum_{k=-1}^{n-1} \mathbf{E}[\Pi_k^{2p}]^{1/2} \mathbf{P}(\ell_1 = n)^{1/2} \\ &\leq M^{1/2}(2p, 1 - \beta) \sum_{n=1}^{\infty} (n+1)^p \mathbf{P}(\ell_1 \geq n)^{1/2}. \end{aligned}$$

Finally, with Theorem 3.5, we obtain that

$$(3.11) \quad \begin{aligned} &\sum_{n=1}^{\infty} (n+1)^p \mathbf{P}(\ell_1 \geq n)^{1/2} \\ &\leq \gamma^{-1} \sum_{n=2}^{\infty} n^p \gamma^{(n-1)/(2\psi\zeta)} \\ &= \gamma^{-1}(1 - \gamma^{1/(2\psi\zeta)})^{-1}(M(p, 1 - \gamma^{1/(2\psi\zeta)}) - 1). \end{aligned}$$

Claim (3.7) now follows by collecting the results in (3.9) to (3.11).  $\square$

We are now ready to state the main result of this subsection.

**THEOREM 3.7.** *Assume (2.26) and that  $\mathbb{E}[A^{-p-\varepsilon}] < \infty$  for some  $p \geq 1$  and  $\varepsilon > 0$ . It holds that*

$$(3.12) \quad \mathbb{E}[\tau_1^p] \leq \frac{\pi^2}{6} \mathbb{E}[L(\rho)^{p+\varepsilon}]^{1/q} \mathbb{E}[\Pi^{2(p-1)q'}]^{1/(2q')} \mathbb{E}[\Pi^{4q'}]^{1/(2q')} < \infty,$$

where  $q = 1 + \varepsilon/p$ , and  $q' = 1 + p/\varepsilon$  is the dual of  $q$ .

**PROOF.** Label with  $\sigma_i$  the  $i$ th vertex that is visited by the process  $\mathbf{X}$ . By Jensen’s inequality, we find

$$(3.13) \quad \begin{aligned} \mathbb{E}[\tau_1^p] &= \mathbb{E} \left[ \left( \sum_{i=1}^{\Pi} L(\sigma_i) \right)^p \right] \leq \mathbb{E} \left[ \Pi^{p-1} \sum_{i=1}^{\Pi} L(\sigma_i)^p \right] \\ &= \sum_{i=1}^{\infty} \mathbb{E}[\Pi^{p-1} L(\sigma_i)^p \mathbb{1}_{\{\Pi \geq i\}}]. \end{aligned}$$

By Hölder’s inequality, and by stationarity, the right-hand side of the last display is smaller than

$$(3.14) \quad \mathbb{E}[L(\rho)^{p+\varepsilon}]^{1/q} \mathbb{E}[\Pi^{2(p-1)q'}]^{1/(2q')} \sum_{i=1}^{\infty} \mathbf{P}(\Pi \geq i)^{1/(2q')}.$$

By Chebychev’s inequality, we find that

$$(3.15) \quad \sum_{i=1}^{\infty} \mathbf{P}(\Pi \geq i)^{1/(2q')} \leq \sum_{i=1}^{\infty} i^{-2} \mathbb{E}[\Pi^{4q'}]^{1/(2q')} = \frac{\pi^2}{6} \mathbb{E}[\Pi^{4q'}]^{1/(2q')}.$$

Putting (3.13), (3.14) and (3.15) together, we obtain the claim.  $\square$

**3.3. An invariance principle and bounds on the covariance.** For ORRW, an invariance principle is known (see Theorem 3 in Durrett, Kesten and Limic [13]). For RWRE, an annealed invariance principle easily follows from the results of Aidékon [2]. We further refer to Peres and Zeitouni [21] for a quenched invariance principle for biased random walks on Galton–Watson trees. Define

$$B^n = \frac{1}{\sqrt{n}} (|X_{[n]}| - [n]v), \quad \beta_t^n = B_t^n + (nt - [nt])(B_{[t+1]}^n - B_t^n), \quad n \geq 1,$$

that is,  $\beta$  is the polygonal interpolation of  $k/n \rightarrow B_{k/n}^n, k \geq 0$ . We endow the space  $C(\mathbb{R}_+, \mathbb{R})$  of continuous functions with the topology of uniform convergence on compacts, and with its Borel  $\sigma$ -algebra.

**PROPOSITION 3.8.** *The  $C(\mathbb{R}_+, \mathbb{R})$ -valued random variable  $\beta^n$  converge under  $\mathbf{P}$  in law to a Brownian motion  $B$ , with covariance*

$$K = \mathbb{E}[(\ell_1 - v\tau_1)^2 | D = \infty] \mathbb{E}[\tau_1 | D = \infty]^{-1}.$$



PROOF. For ORRW, we refer to Theorem 3 in [13]. For RWRE, observe that the second moment of  $\tau_1$ , and thus of  $\ell_1$ , is finite, as follows from Propositions 2.1 and 2.2 in Aidékon [2]. Since  $\mathbf{P}[D = \infty] = 1 - \beta > 0$ , also  $\mathbf{E}[\ell_1^2 | D = \infty] \leq \mathbf{E}[\tau_1^2 | D = \infty] < \infty$ . Further it is well known that

$$(3.16) \quad (\tau_{i+1} - \tau_i, \ell_{i+1} - \ell_i)_{i \geq 1} \quad \text{is an i.i.d. sequence under } \mathbf{P} \quad \text{and}$$

$$\text{for } i \geq 1 \quad (\tau_{i+1} - \tau_i, \ell_{i+1} - \ell_i)$$

has same law under  $\mathbf{P}$  as  $(\tau_1, \ell_1)$  under  $\mathbf{P}(\cdot | D = \infty)$

(see [14]) (see also [18] for a similar statement for biased random walks on Galton–Watson trees). With the help of this i.i.d. structure, the proof of the invariance principle is now quite standard (see, e.g., Theorem 3 in Durrett, Kesten and Limic [13] and also Theorem 3.3 in Shen [23]).  $\square$

With the help of Theorems 3.5 and 3.7, we obtain explicit bounds on the covariance  $K$  via the following proposition. For RWRE (resp., ORRW) denote with  $w$  the right-hand side in inequality (2.12) [resp., (2.49)], so that  $v \geq w$ . Let  $a$  be the smallest even integer larger or equal to  $[3/w] + 1$ . As  $w \leq 1$ , we have  $a \geq 4$ .

PROPOSITION 3.9. *In the case of RWRE, we assume that (2.26) holds and that  $\mathbb{E}[A^{-2-\varepsilon}] < \infty$  for some  $\varepsilon > 0$ . In the case of ORRW we choose  $\psi$  satisfying (2.46). Then we have the following common upper bound on the covariance  $K$ :*

$$(3.17) \quad K \leq (1 - \alpha_\psi)^{-1} (\mathbf{E}[\ell_1^2] + \mathbf{E}[\tau_1^2]) \quad \text{for RWRE and ORRW,}$$

and the following lower bound:

$$(3.18) \quad K \geq b(1 - \alpha_\psi) \mathbf{E}[\tau_1]^{-1} \mathbb{E}[\omega(\rho, \vec{\rho}_1)^{a/2}] \mathbb{E}[\omega(\vec{\rho}_1, \rho)^{a/2-1} (1 - \omega(\vec{\rho}_1, \rho))] \quad \text{for RWRE,}$$

$$K \geq (1 - \alpha_\psi) \mathbf{E}[\tau_1]^{-1} \left(\frac{b}{b + \delta}\right)^2 \left(\frac{\delta}{b + \delta}\right)^{a/2-1}$$

$$\times \left(\frac{\delta}{b - 1 + 2\delta}\right)^{a/2-1} \quad \text{for ORRW.}$$

PROOF. We start with the upper bound. We use the trivial bound  $(d - b)^2 \leq d^2 + b^2$ ,  $d, b \geq 0$ , and  $v \leq 1$  to obtain that

$$(3.19) \quad K \leq \mathbf{E}[\ell_1^2 | D = \infty] + \mathbf{E}[\tau_1^2 | D = \infty] \leq (1 - \beta)^{-1} (\mathbf{E}[\ell_1^2] + \mathbf{E}[\tau_1^2]).$$

The upper bound (3.17) follows from Proposition 2.6. Let us now turn to the lower bound (3.18) for random walk in random environment. We use the following approach:

$$(3.20) \quad \mathbf{E}[(\ell_1 - v\tau_1)^2 | D = \infty] \geq \mathbf{E}[(\ell_1 - v\tau_1)^2 \mathbb{1}_{\{v\tau_1 \geq \ell_1 + 1\}} | D = \infty]$$

$$\geq \mathbf{P}[v\tau_1 \geq \ell_1 + 1 | D = \infty],$$

where the last inequality comes from the fact that on the event  $\{v\tau_1 \geq \ell_1 + 1\}$  we have  $(\ell_1 - v\tau_1)^2 \geq 1$ . Hence

$$(3.21) \quad \begin{aligned} K &\geq \mathbf{P}(v\tau_1 \geq \ell_1 + 1 | D = \infty) \mathbf{E}[\tau_1 | D = \infty]^{-1} \\ &\geq \mathbf{P}(v\tau_1 \geq \ell_1 + 1, D = \infty) \mathbf{E}[\tau_1]^{-1}. \end{aligned}$$

Next we find a suitable subset of  $\{v\tau_1 \geq \ell_1 + 1\}$  whose probability is easy to compute. Consider the event

$$C \stackrel{\text{def}}{=} \left\{ T_2 = a, D(X_{T_2}) = \infty, \bigcup_{i=1}^b \{X_j \in \{\rho, \vec{\rho}_i\}, \forall j \leq T_2 - 1\} \right\}.$$

If this event holds, then the walk, started at the root  $\rho$ , visits level two first at time  $a$  and, after this time, never goes back to level 1. Moreover before time  $T_2$ , the process  $\mathbf{X}$  visits only the vertices  $\rho$  and  $\vec{\rho}_i$  for some  $i$ , and hence it does not return to  $\overleftarrow{\rho}$ . As  $a \geq 4$ , it jumps at least once from  $\vec{\rho}_1$  to  $\rho$ , so that level one cannot be a cut level and  $\ell_1 = 2$ . As  $a \geq [3/w] + 1 \geq [3/v] + 1$ , we have

$$C \subset \{\ell_1 = 2, \tau_1 \geq [3/v] + 1, D = \infty\}.$$

On the event  $\{\ell_1 = 2, \tau_1 \geq [3/v] + 1\}$  we have that  $v\tau_1 \geq 3$ , hence  $v\tau_1 - \ell_1 \geq 1$ . In other words,

$$(3.22) \quad C \subset \{v\tau_1 \geq \ell_1 + 1, D = \infty\}.$$

We first focus on the RWRE case. Let us now compute the probability of the event  $C$ . The Markov property implies that

$$\mathbf{P}_\omega(C) = \sum_{i=1}^b \omega(\rho, \vec{\rho}_i)^{a/2} \omega(\vec{\rho}_i, \rho)^{a/2-1} (1 - \omega(\vec{\rho}_i, \rho)) \mathbf{E}_\omega[\mathbf{P}_{X_{T_2}, \omega}(D = \infty)].$$

The random variables  $\omega(\rho, \vec{\rho}_i)$ ,  $\omega(\vec{\rho}_i, \rho)(1 - \omega(\vec{\rho}_i, \rho))$  and  $\mathbf{E}_\omega[\mathbf{P}_{X_{T_2}, \omega}(D = \infty)]$  are independent, since they are measurable w.r.t. disjoint parts of the environment. We use, in addition, stationarity to find that

$$\mathbf{P}(C) = b \mathbb{E}[\omega(\rho, \vec{\rho}_1)^{a/2}] \mathbb{E}[\omega(\vec{\rho}_1, \rho)^{a/2-1} (1 - \omega(\vec{\rho}_1, \rho))] \mathbf{E}[\mathbf{P}_{X_{T_2}, \omega}(D = \infty)].$$

Again, by independence and stationarity,

$$\begin{aligned} \mathbf{E}[\mathbf{P}_{X_{T_2}, \omega}(D = \infty)] &= \sum_{\nu} \mathbf{E}[\mathbf{P}_{\nu, \omega}(D = \infty), X_{T_2} = \nu] \\ &= \sum_{\nu} \mathbf{P}_{\nu}(D = \infty) \mathbf{P}(X_{T_2} = \nu) \\ &= \mathbf{P}(D = \infty) = 1 - \beta. \end{aligned}$$

It follows that

$$(3.23) \quad \begin{aligned} & \mathbf{P}(v\tau_1 - \ell_1 \geq 1, D = \infty) \\ & \geq \mathbf{P}(C) = b\mathbb{E}[\omega(\rho, \vec{\rho}_1)^{a/2}] \mathbb{E}[\omega(\vec{\rho}_1, \rho)^{a/2-1} (1 - \omega(\vec{\rho}_1, \rho))] \\ & \quad \times (1 - \beta). \end{aligned}$$

The lower bound (3.18) for RWRE now follows from (3.21), (3.23) and Proposition 2.6. Let us now turn to the proof of the lower bound (3.18) for ORRW. We follow the same strategy as above, and we see that (3.21) and (3.22) hold. It remains to compute the probability of the event  $C$

$$\mathbf{P}(C) = (1 - \beta) \left(\frac{b}{b + \delta}\right)^2 \left(\frac{\delta}{b + \delta}\right)^{a/2-1} \left(\frac{\delta}{b - 1 + 2\delta}\right)^{a/2-1}.$$

By proceeding as in (3.23) and above, and with the help of Proposition 2.15, the proof of (3.18) is complete.  $\square$

APPENDIX

LEMMA A.1. *Let  $M(n, q)$  denote the  $n$ th moment of a geometric random variable with parameter  $q$ . Then for  $n \geq 1$ ,  $M(n, q) \leq c_n q^{-n}$ , for some constant  $c_n$  that only depends on  $n$ .*

PROOF. We define  $g(q, n) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} k^n (1 - q)^{k-1}$ , and notice that  $M(n, q) = \sum_{k=1}^{\infty} k^n q (1 - q)^{k-1} = qg(q, n)$ . Since  $0 < q < 1$ , it is enough to show that there are coefficients  $a_i^{(n)}$  such that

$$(A.1) \quad g(q, n) = \frac{\sum_{i=1}^n a_i^{(n)} q^{n-i}}{q^{n+1}} = \sum_{i=1}^n a_i^{(n)} q^{-i-1}.$$

We prove (A.1) by induction. As  $g(q, 1) = 1/q^2$ , (A.1) holds for  $n = 1$ . Suppose now (A.1) holds for  $n - 1$ . We have

$$(A.2) \quad \begin{aligned} g(q, n) - g(q, n - 1) &= \sum_{k=1}^{\infty} k^{n-1} (k - 1) (1 - q)^{k-1} \\ &= (1 - q) \frac{d}{d(1 - q)} \sum_{k=1}^{\infty} k^{n-1} (1 - q)^{k-1} \\ &= (1 - q) \frac{d}{d(1 - q)} g(q, n - 1), \end{aligned}$$

where  $\frac{d}{dx}$  denotes the derivative with respect  $x$ . By the induction hypothesis,

$$(A.3) \quad \frac{d}{d(1 - q)} g(q, n - 1) = \sum_{i=1}^{n-1} (i + 1) a_i^{(n-1)} q^{-i-2},$$

and hence, using (A.1) to (A.3),

$$(A.4) \quad g(q, n) = na_{n-1}^{(n-1)} q^{-n-1} + \sum_{i=2}^{n-1} i(a_{i-1}^{(n-1)} - a_i^{(n-1)})q^{-i-1} - a_1^{(n-1)} q^{-2}.$$

This shows (A.1), and the proof is finished.  $\square$

**Acknowledgments.** We would like to thank two anonymous referees for their comments and for pointing out to us the lower bound for the speed (2.10) established by Aidékon in [3].

## REFERENCES

- [1] AIDÉKON, E. (2008). Transient random walks in random environment on a Galton–Watson tree. *Probab. Theory Related Fields* **142** 525–559. [MR2438700](#)
- [2] AIDÉKON, E. (2010). Large deviations for transient random walks in random environment on a Galton–Watson tree. *Ann. Inst. H. Poincaré Probab. Statist.* **46** 159–189. [MR2641775](#)
- [3] AIDÉKON, E. (2009). Marches aléatoires en milieu aléatoire et marches branchantes. Ph.D. thesis, Univ. Paris VI. Available at <http://tel.archives-ouvertes.fr/tel-00426925/fr/>.
- [4] CHEN, D. (2002). Estimating the speed of random walks. In *Applied Probability (Hong Kong, 1999)*. *AMS/IP Stud. Adv. Math.* **26** 17–23. Amer. Math. Soc., Providence, RI. [MR1909874](#)
- [5] COLLEVECCHIO, A. (2006). Limit theorems for reinforced random walks on certain trees. *Probab. Theory Related Fields* **136** 81–101. [MR2240783](#)
- [6] COLLEVECCHIO, A. (2006). On the transience of processes defined on Galton–Watson trees. *Ann. Probab.* **34** 870–878. [MR2243872](#)
- [7] COLLEVECCHIO, A. (2009). Limit theorems for vertex-reinforced jump processes on regular trees. *Electron. J. Probab.* **14** 1936–1962. [MR2540854](#)
- [8] DAI, J. J. (2005). A once edge-reinforced random walk on a Galton–Watson tree is transient. *Statist. Probab. Lett.* **73** 115–124. [MR2159246](#)
- [9] DAVIS, B. (1990). Reinforced random walk. *Probab. Theory Related Fields* **84** 203–229. [MR1030727](#)
- [10] DAVIS, B. and VOLKOV, S. (2002). Continuous time vertex-reinforced jump processes. *Probab. Theory Related Fields* **123** 281–300. [MR1900324](#)
- [11] DEMBO, A., GANTERT, N., PERES, Y. and ZEITOUNI, O. (2002). Large deviations for random walks on Galton–Watson trees: Averaging and uncertainty. *Probab. Theory Related Fields* **122** 241–288. [MR1894069](#)
- [12] DURRETT, R. (1996). *Probability: Theory and Examples*, 2nd ed. Duxbury Press, Belmont, CA. [MR1609153](#)
- [13] DURRETT, R., KESTEN, H. and LIMIC, V. (2002). Once edge-reinforced random walk on a tree. *Probab. Theory Related Fields* **122** 567–592. [MR1902191](#)
- [14] GROSS, T. (2004). Marche aléatoire en milieu aléatoire sur un arbre. Ph.D. thesis, Univ. Paris VI.
- [15] HU, Y. and SHI, Z. (2007). Slow movement of random walk in random environment on a regular tree. *Ann. Probab.* **35** 1978–1997. [MR2349581](#)
- [16] HU, Y. and SHI, Z. (2007). A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. *Probab. Theory Related Fields* **138** 521–549. [MR2299718](#)
- [17] LYONS, R. and PEMANTLE, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20** 125–136. [MR1143414](#)

- [18] LYONS, R., PEMANTLE, R. and PERES, Y. (1996). Biased random walks on Galton–Watson trees. *Probab. Theory Related Fields* **106** 249–264. [MR1410689](#)
- [19] MENSHIKOV, M. and PETRITIS, D. (2002). On random walks in random environment on trees and their relationship with multiplicative chaos. In *Mathematics and Computer Science, II (Versailles, 2002)* 415–422. Birkhäuser, Basel. [MR1940150](#)
- [20] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79 (electronic). [MR2282181](#)
- [21] PERES, Y. and ZEITOUNI, O. (2008). A central limit theorem for biased random walks on Galton–Watson trees. *Probab. Theory Related Fields* **140** 595–629. [MR2365486](#)
- [22] SELLKE, T. (1994). Reinforced random walk on the d-dimensional integer lattice. Technical Report 94–26, Dept. Statistics, Purdue Univ.
- [23] SHEN, L. (2003). On ballistic diffusions in random environment. *Ann. Inst. H. Poincaré Probab. Statist.* **39** 839–876. [MR1997215](#)
- [24] SZNITMAN, A.-S. (2004). Topics in random walks in random environment. In *School and Conference on Probability Theory. ICTP Lect. Notes, XVII* 203–266 (electronic). Abdus Salam Int. Cent. Theoret. Phys., Trieste. Available at [www.ictp.it/~pub\\_off/lectures/lns017/Sznitman/Sznitman.ps.gz](http://www.ictp.it/~pub_off/lectures/lns017/Sznitman/Sznitman.ps.gz). [MR2198849](#)
- [25] VIRÁG, B. (2000). On the speed of random walks on graphs. *Ann. Probab.* **28** 379–394. [MR1756009](#)
- [26] ZEITOUNI, O. (2004). Random walks in random environment. In *Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1837** 189–312. Springer, Berlin. [MR2071631](#)
- [27] ZEITOUNI, O. (2006). Random walks in random environments. *J. Phys. A* **39** R433–R464. [MR2261885](#)

DIPARTIMENTO DI MATEMATICA APPLICATA  
UNIVERSITÀ CA' FOSCARI  
SAN GIOBBE, CANNAREGIO 873  
30121 VENEZIA  
ITALY  
E-MAIL: [collevec@unive.it](mailto:collevec@unive.it)

MAX PLANCK INSTITUTE  
FOR MATHEMATICS IN THE SCIENCES  
COMMISSARIAT AUX ASSURANCES  
7, BOULEVARD ROYAL  
L-2449 LUXEMBOURG  
GD DE LUXEMBOURG  
E-MAIL: [tomschmitz1@gmail.com](mailto:tomschmitz1@gmail.com)