FUNCTIONAL CENTRAL LIMIT THEOREM FOR RANDOM WALKS IN RANDOM ENVIRONMENT DEFINED ON REGULAR TREES

ANDREA COLLEVECCHIO, MASATO TAKEI, AND YUMA UEMATSU

Abstract. We study Random Walks in an i.i.d. Random Environment (RWRE) defined on $b$-regular trees. We prove a functional central limit theorem (FCLT) for transient processes, under a moment condition on the environment. We emphasize that we make no uniform ellipticity assumptions. Our approach relies on regenerative levels, i.e. levels that are visited exactly once. On the way, we prove that the distance between consecutive regenerative levels have a geometrically decaying tail. In the second part of this paper, we apply our results to Linearly Edge-Reinforced Random Walk (LERRW) to prove FCLT when the process is defined on $b$-regular trees, with $b \geq 4$, substantially improving the results of the first author (see Theorem 3 of [4]).

1. INTRODUCTION

Random Walk in Random Environment (RWRE) is a class of self-interacting processes that attracted much attention from probabilists since the seminal work of Kesten, Kozlov, and Spitzer [10] and Solomon [15], in the 70’s. It seems that the initial motivation behind this class of process was related to problems in biology, crystallography and metal physics. The interest in this field grew substantially, and we refer to [3] and [16] for an overview of this beautiful subject.

We study random walks in an i.i.d. random environment defined on $b$-regular trees. We provide a functional central limit theorem (FCLT) for processes that are transient, assuming a moment condition on the environment. Our approach relies on regenerative levels, i.e. levels that are visited exactly once. On the way, we prove that the space between these regenerative levels have a geometrically decaying tail. We emphasize that we make no uniform ellipticity assumptions.

Key words and phrases. Random Walks in Random Environment, Self-interacting random walks, Functional central limit theorem.
To the best of our knowledge, this is the first of this type for RWRE on trees (even under the assumption of uniform ellipticity). A FCLT was proved by Peres and Zeitouni for biased random walks on Galton-Watson trees (see [13]).

In the second part of this paper, we apply our results to Linearly Edge-Reinforced Random Walk (LERRW), a model introduced in [7] and which is described below, after Remark 2. We prove FCLT for LERRW when the process is defined on $b$-regular trees, with $b \geq 4$, substantially improving the results of the first author (see Theorem 3 of [4]). Moreover, our results can be combined with the ones of Zhang [17] and provide upper large deviations results for $b$-regular trees with $b \geq 4$, which could be improved, with extra computations, to $b \geq 3$ (see Remark 5 below).

Fix an integer $b \in \mathbb{N}$. Let $G = (V, E)$ be an infinite $b$-regular tree with root $\rho$. We augment $G$ by adjoining a parent $\rho^{-1}$ to the root $\rho$. In this graph each vertex has degree $b + 1$, with the exception of $\rho^{-1}$ that has degree one. If two vertices $\nu$ and $\mu$ are the endpoints of the same edge, they are said to be neighbours, and this property is denoted by $\nu \sim \mu$. The distance $|\nu - \mu|$ between any pair of vertices $\nu, \mu$, not necessarily adjacent, is the number of edges in the unique self-avoiding path connecting $\nu$ to $\mu$. For any other vertex $\nu$, with $\nu \neq \rho$, we let $|\nu|$ be the distance of $\nu$ from the root $\rho$, i.e. $|\nu| = |\nu - \rho|$. We set $|\rho^{-1}| = -1$.

We write $\nu < \mu$ if $\nu$ is an ancestor of $\mu$, that is if $\nu$ lays on the self-avoiding path connecting $\mu$ to $\rho$. Alternatively, we say that $\mu$ is a descendant of $\nu$. For any vertex $\nu$, denote by $\nu_1, \nu_2, \ldots, \nu_b$ its offspring, and by $\nu^{-1}$ its parent.

For $\nu \in V$, let $A_\nu = (A_{\nu 1}, A_{\nu 2}, \ldots, A_{\nu b})$ to denote the (finite, positive) weights on the edges between $\nu$ and its offspring. For simplicity, we index the weight associated to edge $e$ by the endpoint of $e$ with larger distance from $\rho$. The environment $\omega$ for the random walk on the tree is then defined, for any vertex $\nu$ with offspring $\nu_i$, $1 \leq i \leq b$, by the probabilities

\begin{equation}
\omega(\nu, \nu_i) := \frac{A_{\nu i}}{1 + \sum_{1 \leq j \leq b} A_{\nu j}}; \quad \omega(\nu, \nu^{-1}) := \frac{1}{1 + \sum_{1 \leq j \leq b} A_{\nu j}}.
\end{equation}

We set $\omega(\nu, \mu) = 0$ if $\mu$ and $\nu$ are not neighbours. Given the environment $\omega$, we define the random walk $X = \{X_n, n \geq 0\}$ that starts at $\rho$ to be the Markov chain with $P_\rho(X_0 = x) = 1$, having transition probabilities $P_\omega(X_{n+1} = \mu 1 | X_n = \mu) = \omega(\mu, \mu 1)$.
whenever \( \mu \neq \varrho^{-1} \). We set
\[
P^\omega(X_{n+1} = \varrho \mid X_n = \varrho^{-1}) = 1.
\]
The combined probability measure from which the environment is realized is denoted by \( \mathbb{P} \) and its expectation by \( \mathbb{E} \), and the semi-direct product \( P^x := \mathbb{P} \times P^\omega_x \) represents the annealed measure of the process which starts from vertex \( x \). For simplicity, we use \( P \) and \( P^\omega \) respectively for \( P^\varrho \) and \( P^\varrho_\omega \). For any vertex \( \nu \), set
\[
T_\nu := \inf\{ k \geq 0 : X_k = \nu \}.
\]
Sometimes we use \( T(\nu) \) instead of \( T_\nu \). We are interested in the case (see Assumption A below) where \( P(T(\nu) = \infty) > 0 \). Moreover, for \( n \in \mathbb{N} \), let
\[
T_n := \inf\{ k \geq 0 : |X_k| = n \}.
\]

**Assumption A** From now on, we suppose that \( (A_\nu)_{\nu \in V} \) are i.i.d., and
\[
\inf_{t \in [0,1]} \mathbb{E}[A^t] > 1/b.
\]

In particular, condition (1.2) implies transience of the process, i.e. it visits each vertex only finitely often, a.s.. This result was proved by Lyons and Pemantle [11], and see [2] for a generalization of this result to Markovian environments. Aïdékon ([1], Theorem 1.5) proved that the condition
\[
\mathbb{E} \left[ \left( \sum_{1 \leq i \leq b} A_{\varrho i} \right)^{-1} \right] < \infty,
\]
is sufficient for the transient process \( X \) to have positive speed, i.e. there exists a positive finite constant \( v_b \) such that
\[
\lim_{n \to \infty} \frac{|X_n|}{n} = v_b, \quad \mathbb{P}\text{-a.s.}
\]
Denote by \( \lfloor x \rfloor \) the integer part of \( x \). Our main result is the following.

**Theorem 1** (Annealed FCLT). Under Assumption A, if we make the further assumption
\[
\mathbb{E} \left[ \left( \sum_{1 \leq i \leq b} A_{\varrho i} \right)^{-p} \right] < \infty, \quad \text{for some } p > 2,
\]
then there exists a positive constant \( \sigma_b \) such that
\[
\left( \frac{|X_{nt}| - v_bnt}{\sqrt{n}\sigma_b} \right)_{t \in [0,1]} \Rightarrow (W_t)_{t \in [0,1]},
\]
where $|X_{nt}|$ is the linear interpolation between $|X_{\lfloor nt \rfloor}|$ and $|X_{\lfloor nt \rfloor+1}|$ for non-integer values of $nt$, $(W_t)_t$ is a standard Brownian motion, and $\Rightarrow$ denotes convergence in distribution as $n \to \infty$.

**Remark 2.** We are not assuming that the random variables $A_{E_i}$ are bounded or bounded away from 0, i.e. the so-called uniform ellipticity assumption.

**Remark 3.** We would like to add few words about the topology under which the convergence in (1.5) takes place. We consider the space of càdlàg functions on $[0,1]$ equipped with the Borel $\sigma$-algebra generated by the Skorokhod topology.

We apply our results to Linearly Edge-Reinforced Random Walk (LERRW) on trees, which is defined as follows. To each edge of the tree, assign initial weight one. These weights are updated depending on the behaviour of the process. LERRW takes values on the vertices of $\mathcal{G}$, at each step it jumps to vertices which are neighbors of the present one, say $x$. The probability to pick a particular neighbor is proportional to the weight of the edge connecting that vertex to $x$. Each time the process traverses an edge, its weight is increased by one. See [14] for a surprising connection between LERRW and the Zirnbauer $H^{3/2}$ model. When $\mathcal{G}$ is a tree, we can use a random walk in i.i.d. random environment to study LERRW. LERRW on the binary tree is transient and has positive speed, even though does not satisfy (1.3) (see [1]). Our result is the following and improves Theorem 3 of [4].

**Theorem 4.** Let $X$ be LERRW on a $b$-regular tree, with $b \geq 4$. Then $X$ satisfies (1.5) for some choice of $\sigma_b$.

**Remark 5.** If we replace (1.4) with

$$\mathbb{E} \left[ \left( \sum_{1 \leq i \leq b} A_{E_i} \right)^{-p} \right] < \infty, \quad \text{for some } p > 1,$$

then we can prove the finiteness of certain moments of certain regenerative times, which is enough in order to obtain an upper large deviation result for the speed for the case $b \geq 3$, according to a paper of Zhang [17]. Notice that the previous known result on this was given by Zhang [17] for $b \geq 70$.

2. Regenerative times and structure of the proof of Theorem 1

From now on, $p$ will be used to denote the exponent that satisfies condition (1.4). Under the assumptions of Theorem 1 (more precisely (1.2)), the process $X$ is transient. It is natural to introduce in this context the so-called regenerative times.
Definition 6. Set $\tau_0 = 0$. For $m \in \mathbb{N}$ define recursively,

$$
\tau_m = \inf \left\{ k > \tau_{m-1} : \sup_{j<k} |X_j| < |X_k| \leq \inf_{j\geq k} |X_j| \right\}.
$$

For each $m \in \mathbb{Z}_+$, let $\ell_m = |X_{\tau_m}|$.

The elements of the process $(\ell_i)_i$ are called cut levels (or regenerative levels). The regenerative times $(\tau_i)_i$ are the hitting times of the cut levels. Under the measure $P$, the sequences $((\ell_k - \ell_{k-1}, \tau_k - \tau_{k-1}))_{k \geq 1}$ are independent and, except for the first one, distributed like $(\ell_1, \tau_1)$ under $P (\cdot | T(\rho^{-1}) = \infty)$. Moreover, based on a result of Zerner (see Lemma 3.2.5 in [16]), it is not difficult to prove that $E[\ell_2 - \ell_1] < \infty$. We prove that $\ell_2 - \ell_1$, under Assumption A, has an exponential tail. To our knowledge, this result is new.

Theorem 7. Under Assumption A, for any $b \geq 2$, we have that

$$
P(\ell_2 - \ell_1 \geq k) \leq a^k,
$$

for some constant $a \in (0, 1)$.

Afterwards, we prove that under the assumption [14] we have

$$
E[(\tau_2 - \tau_1)^2] < \infty.
$$

Set $Y_i := \ell_i - \ell_{i-1} - v_b (\tau_i - \tau_{i-1})$. We have, for $\tau_m \leq n < \tau_{m+1},$

$$
\frac{|X_n| - n v_b}{\sqrt{n}} \geq \frac{\ell_m - \tau_{m+1} v_b}{\sqrt{\tau_I}},
$$

where $I$ equals $m + 1$ if the numerator $\ell_m - \tau_{m+1} v_b \geq 0$ and $m$ otherwise. Hence,

$$
\frac{|X_n| - n v_b}{\sqrt{n}} \geq \sqrt{\frac{m}{\tau_I}} \left( \frac{1}{\sqrt{m}} \sum_{i=1}^m Y_i - v_b \frac{\tau_m - \tau_{m+1}}{\sqrt{m}} \right).
$$

As $\tau_m = \sum_{i=1}^m (\tau_i - \tau_{i-1})$, in virtue of the strong law of large numbers, $\tau_I/m$ converges a.s. to a positive finite constant. Moreover, (2.1) guarantees that $\sum_{i=1}^m Y_i/(\sqrt{m})$ weakly converges to a normal$(0, \sigma)$, for some finite constant $\sigma > 0$, and $(\tau_m - \tau_{m+1})/\sqrt{m}$ converges in probability to 0. Hence, assuming (2.1), and using Slutsky Lemma, the right hand side of (2.3) converges weakly to a normal$(0, K)$ for some $K \in (0, \infty)$. Similarly

$$
\frac{|X_n| - n v_b}{\sqrt{n}} \leq \sqrt{\frac{m+1}{\tau_J}} \left( \frac{1}{\sqrt{m+1}} \sum_{i=1}^{m+1} Y_i + v_b \frac{\tau_m - \tau_{m+1}}{\sqrt{m+1}} \right),
$$

where $J$ equals $m + 1$ if the numerator $\ell_m - \tau_{m+1} v_b \geq 0$ and $m$ otherwise.
where $J$ equals $m$ if $\ell_{m+1} - \tau_m v_b \geq 0$ and $m + 1$ otherwise. The right-hand side of (2.4) converges to a normal$(0, \sigma)$. The procedure to step from the ordinary central limit theorem to the functional one is classical, and we refer to section 4 of [9].

3. Extension Processes

Here, we define a construction that is closely related to the ones introduced in [5] and [6]. This construction allows to decouple the behaviour of the process on subtrees, even when the process is transient. This will allow us to build a family of coupled processes which are independent when defined on disjoint subsets of the tree, and usefully correlated to $X$.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ denote a probability space on which

$$(3.1) \quad Y = (Y(\nu, \mu, k) : (\nu, \mu) \in V^2, \text{with } \nu \sim \mu, \text{ and } k \in \mathbb{Z}_+)$$

is a family of independent exponential random variables with mean 1, and where $(\nu, \mu)$ denotes an ordered pair of vertices. Below, we use these collections of random variables to generate the steps of $X$. Moreover, we define a family of coupled walks using the same collection of ‘clocks’ $Y$.

Define, for any $\nu, \mu \in V$ with $\nu \sim \mu$, the quantities

$$(3.2) \quad r(\nu, \mu) = \mathbb{I}_{\{\mu = \nu - 1\}} + \sum_{i=1}^{b} A_{\nu i} \mathbb{I}_{\{\mu = \nu_i\}}.$$

As it was done in [6], we are now going to define a family of coupled processes on the subtrees of $G$. For any rooted subtree $G' = (V', E')$ of $G$, the root $g'$ of $G'$ is defined as the vertex of $V'$ with smallest distance to $g$. Let us define the extension $X^{(G')}_{g'}$ on $G'$ as follows. Set $X_{0}^{(G')} = g'$. For $\nu \in V'$, a collection of nonnegative integers $k = (k_{\mu})_{\mu : [\nu, \mu] \in E'}$, and $n \geq 0$, let

$$A_{k, n, \nu}^{(G')} = \left\{ X_{n}^{(G')} = \nu \right\} \cap \bigcap_{\mu : [\nu, \mu] \in E'} \left\{ \#\{1 \leq j \leq n : (X_{j-1}^{(G')}, X_{j}^{(G')}) = (\nu, \mu)\} = k_{\mu} \right\}.$$

Note that the event $A_{k, n, \nu}^{(G')}$ deals with jumps along oriented edges. For $\nu, \nu'$ such that $[\nu, \nu'] \in E'$ and for $n \geq 0$, on the event

$$(3.3) \quad A_{k, n, \nu}^{(G')} \cap \left\{ \nu' = \arg \min_{\mu : [\nu, \mu] \in E'} \left\{ \sum_{k=0}^{k_{\mu}} \frac{Y(\nu, \mu, k)}{r(\nu, \mu)} \right\} \right\},$$
we set $X_{n+1}^{(G')} = \nu'$, where the function $r$ is defined in (3.2) and the clocks $Y$’s are from the same collection $Y$ fixed in (3.1).

We define $X = X^{(G)}$ to be the extension on the whole tree. It is easy to check, from memoryless property of exponential random variables, that this provides a construction of the RWRE $X$ on $G$. This continuous-time embedding is classical and it is inspired by Rubin’s construction, after Herman Rubin (see the Appendix in Davis [8]). If we consider proper subtrees $G'$ of $G$, one can check that, with these definitions, the steps of $X$ on the subtree $G'$ are given by the steps of $X^{(G')}$. Notice that for any two subtrees $G'$ and $G''$ whose edge sets are disjoint, the extensions $X^{(G')}$ and $X^{(G'')}$ are independent as they are defined by two disjoint sub-collections of $Y$.

**Definition 8.** For any vertex $\nu \in V$, define $fc(\nu)$, called the first-child of $\nu$, as the a.s. unique minimizer of $Y(\nu, \nu_i, 0)/r(\nu, \nu_i)$ over the collection of offspring $(\nu_i)$ of $\nu$. For definiteness, the root $\rho$ and its parent $\rho^{-1}$ are not first children.

Notice that a first child is not necessarily visited by the process $X$. If the latter visits $fc(\nu)$, then it is the first among the children of $\nu$ to be visited. The random vertices $X_{T_n}$, for $n \geq 1$, are all first children.

### 4. Proof of Theorem 7

For any vertex $\nu$, with $\nu \neq \rho^{-1}$, denote by $\Lambda_\nu$ the tree composed by $\nu$, $\nu^{-1}$, the descendants of $\nu$ and the edges connecting them. This tree is isomorphic to the original tree $G$. Consider the extension $X^{\Lambda_\nu}$. Set

$$T^{(\nu)}_i := \inf\{n > 0: |X^{\Lambda_\nu}_n| - |\nu| = i\},$$

i.e. the hitting times of this extensions to level that has distance $i$ from $|\nu|$. Using condition (1.2), combined with arguments from [2], for all large $n$ we have

$$b^n P(T^{(\nu)}_{i-1} > T^{(\nu)}_i) > 1.$$  

(4.1)

In fact under condition (1.2), it is proved that

$$\lim_{n \to \infty} b^n P(T^{(\nu)}_{i-1} > T^{(\nu)}_i) = \infty.$$  

(4.2)

(See proof of Theorem 2.1 in [2]). Fix $n^* \in \mathbb{N}$ which satisfies (4.1). We now construct a branching process as follows. We color green the vertices $\nu$ at level $n^*$ which are visited before $X$ returns to $\rho^{-1}$. A vertex $\nu$ at level $jn^*$, for some integer $j \geq 2$, is colored green, if both
• its ancestor $\mu$ at level $(j - 1)n^*$ is green, and
• the extension over the path $[\mu^{-1}, \nu]$, hits $\nu$ before returning to $\mu^{-1}$.

The green vertices evolve as a Galton–Watson tree, with offspring mean $b^{n^*} P(T_{-1}^{(\nu)} > T_{n^*}^{(\nu)}) > 1$. Hence this random tree is supercritical, and thus the probability of there being an infinite number of green vertices is positive.

Denote by $\text{GR}(\Lambda_\nu)$ the set of green vertices on the tree $\Lambda_\nu$. Next, we want to define a sequence of events which we show to be independent and which are closely related to the event that a given level is a cut level. Fix $\xi \in \mathbb{N}$. For any vertex $\nu \in V$, define the random subset of vertices $\Theta_\nu \subset V$ as follows. Vertex $\mu \in \Theta_\nu$, iff

- $\mu$ is a descendant of $\nu$;
- the distance between $\mu$ and $\nu$ is a multiple of $\xi n^*$;
- $\mu$ is a first child.

**Definition 9.** Define $\text{GR}^- (\Lambda_\nu)$ to be the set of green vertices obtained from $\text{GR}(\Lambda_\nu)$ by deleting elements of $\Theta_\nu$ and their descendants. Define the event

$$E(\nu) := \{|\text{GR}^- (\Lambda_\nu)| = \infty\}.$$

**Proposition 10.** For any $\xi$ large enough, $P(E(\nu)) > 0$.

**Proof.** As we observed, the green vertices evolve as a supercritical Galton–Watson tree. The event $E(\nu)$ is the survival event for a certain subtree of the Galton-Watson tree of green vertices, obtained by pruning. Choose $\xi$ large enough that

$$b^{\xi n^* - 1}(b - 1)P(T_{-1}^{(\nu)} > T_{\xi n^*}^{(\nu)}) > 1.$$  

This is possible because of (4.2). Color brown the descendants of $\nu$ at distance $\xi n^*$ from $\nu$ which are green vertices and are not first children. Recursively, color brown the descendants of $\nu$ at distance $k\xi n^*$,

- which are green vertices and not first children, and
- whose ancestors at level $(k - 1)\xi n^*$ from $\nu$ are brown.

The random set of vertices of brown vertices evolve like the population of a branching process with mean offspring larger than

$$b^{\xi n^* - 1}(b - 1)P(T_{-1}^{(\nu)} > T_{\xi n^*}^{(\nu)}).$$

Hence it is supercritical by virtue of (4.3), which in turn implies that $P(E(\nu)) > 0$. \[\blacksquare\]
Proposition 11. The events $D_k := E(X_{T_{i\xi^*}})$, with $k \in \mathbb{N}$, are independent.

Proof. Fix indices $i_1 < i_2 < \ldots < i_n$. It is enough to prove that

$$P\left(\bigcap_{k=1}^{n} D_{i_k}\right) = P\left(\bigcap_{k=1}^{n-1} D_{i_k}\right) P(D_{i_n}).$$

To prove (4.4), we condition on the possible values of $X_{T_{i\xi^*}}$. To simplify notation, set $\iota_n = i_n \xi^*$. Then

$$P\left(\bigcap_{k=1}^{n} D_{i_k}\right) = \sum_{\nu \in V: |\nu| = \iota_n} P\left(\bigcap_{k=1}^{n} D_{i_k} \mid X_{T_{i\nu}} = \nu\right) P\left(X_{T_{i\nu}} = \nu\right).$$

Conditionally on $\{X_{T_{i\nu}} = \nu\}$, $D_{i_n}$ is determined by the collection of exponentials $Y(x, y, k)$ where both $x, y$ are vertices of $\Lambda_\nu$ and $k \in \mathbb{N}$. On the other hand, conditionally on $\{X_{T_{i\nu}} = \nu\}$, the event $\bigcap_{k=1}^{n-1} D_{i_k}$ depends on a disjoint set of exponentials. Hence $D_{i_n}$ and $\bigcap_{k=1}^{n-1} D_{i_k}$ are, given $\{X_{T_{i\nu}} = \nu\}$, conditionally independent, i.e.

$$P\left(\bigcap_{k=1}^{n} D_{i_k} \mid X_{T_{i\nu}} = \nu\right) = P\left(\bigcap_{k=1}^{n-1} D_{i_k} \mid X_{T_{i\nu}} = \nu\right) P(D_{i_n} \mid X_{T_{i\nu}} = \nu).$$

Finally, notice that by a simple symmetry argument, we have

$$P(D_{i_n} \mid X_{T_{i\nu}} = \nu) = P(D_{i_n}).$$

Proof of Theorem 7. First, notice that if $D_1$ holds, then $i\xi^*$ is a cut level, as after time $T_{i\xi^*}$, the process $X$ never visits level $i\xi^* - 1$ again. Hence,

$$P(\ell_1 \geq n) \leq P\left(\bigcap_{i=1}^{\left\lfloor n/(\xi^*)\right\rfloor} D_i^c\right) = P\left(D_i^{\left(\left\lfloor n/(\xi^*)\right\rfloor\right)}\right),$$

proving the theorem.

5. Finite second moment between cut times

Our plan is to prove, in order, that

- the $q$-th moment of the number of distinct vertices visited by time $\tau_1$ grows as a power function, with degree $q$, for all $q > 1$.
- The $p$-th moment of the total number of visits to $\varrho$ by the process $X$ is finite.
Let $\Pi_n$ be the cardinality of the range, that is i.e. the number of distinct vertices, of $X$ by time $T_n$. The number of vertices visited at level $i$ is bounded by $Y_i$, where $(Y_i)_i$ is a sequence of i.i.d. geometric random variables. We recall that the process is transient. Hence, each time it jumps to an unvisited vertex $\nu$ there is a fixed positive probability that the process never visits again $\nu^{-1}$.

Hence for any $q > 1$,

$$E[\Pi_n^q] \leq E \left[ \left( \sum_{i=1}^{n} Y_i \right)^q \right] \leq n^q E[Y_1^q].$$

In other words, for any $q > 1$, we have

$$(5.1) \quad E[\Pi_n^q] = O(n^q).$$

Define $\Pi = \Pi_{\ell_1}$, that is the number of different vertices visited by the time the process hits the first cut level.

**Lemma 12.** For any $q > 1$, we have that

$$(5.2) \quad E[\Pi^q] < \infty.$$

**Proof.** Notice that

$$E[\Pi^q] = \sum_{n=0}^{\infty} E[\Pi_n^q 1_{\{\ell_1 = n\}}] \leq C \sum_{n=0}^{\infty} n^q P(\ell_1 \geq n)^{1/2} < \infty,$$

where we used Cauchy-Schwarz’s inequality, (5.1) and Theorem 7.

Define, for $\nu \in V$,

$$L_{\nu} := \sum_{k=0}^{\infty} 1_{\{X_k = \nu\}}, \quad \beta_{\nu}(\omega) := P^\nu(\ell_1 = \infty),$$

which respectively are, the total time spent in $\nu$ and the quenched probability that the walk never returns to $\nu^{-1}$.

**Remark 13.** Under the measure $P_\omega$, the random variable $L_{\omega^{-1}}$ is distributed as $\text{Geometric}(\beta_{\omega}(\omega))$, i.e.

$$P_\omega(L_{\omega^{-1}} = k) = \left(1 - \beta_{\omega}(\omega)\right)^k \beta_{\omega}(\omega) \quad \text{for} \ k \geq 0.$$

**Proposition 14.** Under the assumptions of Theorem 7 we have that

$$E \left[ \left( \beta_{\omega} \right)^{-p} \right] < \infty,$$

where $p > 2$ satisfies (1.4).
Proof. This proof is inspired by the proof of Lemma 2.2 in [1]. We include the steps for completeness.

\[ \beta_\nu(\omega) = \sum_{i=1}^{b} \omega(\nu, \nu_i) \beta_{\nu_i}(\omega) + \sum_{i=1}^{b} \omega(\nu, \nu_i)(1 - \beta_{\nu_i}(\omega)) \beta_\nu(\omega). \]  

From (5.3), it follows that

\[ \frac{1}{\beta_\nu} = 1 + \frac{1}{\sum_{i=1}^{b} A_{\nu_i}\beta_{\nu_i}} \leq 1 + \min_{1 \leq i \leq b} \frac{1}{A_{\nu_i}\beta_{\nu_i}}. \]  

Consider a random path generated as follows. We set \( v_0 = \zeta \), and we define \( v_k \), with \( k \geq 1 \), recursively. Suppose that \( v_j \) for \( j \leq k \) are defined. Set \( v_{k+1} \) to be one of the maximizers \( x \mapsto A_x \), where \( x \) ranges over the offspring of \( v_k \). If there is more than one maximizer, we choose among them uniformly at random. Define \( C(v_k) \) the set of offspring of \( v_k \) different from \( v_{k+1} \). Fix \( \varepsilon > 0 \).

\[ E_n := \bigcap_{k=1}^{\infty} \bigcap_{y \in C(v_k)} \{ A_y \beta_y(\omega) < \varepsilon \}. \]

We set \( E_0^c = \emptyset \). Notice that \( E_{n+1} \subset E_n \) and that on the event \( E_{n+1}^c \cap E_n \) we have

\[ \min_{y \in C(v_n) \cup \{ v_{n+1} \}} \frac{1}{A_y \beta_y(\omega)} \leq \frac{1}{\varepsilon}. \]

Combining these two facts with (5.4), we infer the following. Denote by \( (y(i))_i \) the offspring of \( v_n \),

\[ \frac{\mathbb{I}_{E_n}}{\beta_{v_n}(\omega)} \leq 1 + \min_{i \in \mathbb{N}} \frac{\mathbb{I}_{E_{n+1}} \mathbb{I}_{E_n}}{A_{y(i)} \beta_{y(i)}(\omega)} + \min_{i \in \mathbb{N}} \frac{\mathbb{I}_{E_{n+1}}}{A_{y(i)} \beta_{y(i)}(\omega)} \]

\[ \leq 1 + \frac{1}{\varepsilon} \mathbb{I}_{E_{n+1}} \mathbb{I}_{E_n} + \min_{y \in C(v_n)} \frac{\mathbb{I}_{E_{n+1}}}{A_y \beta_y(\omega)} \]

\[ \leq 1 + \frac{1}{\varepsilon} \mathbb{I}_{E_{n+1}} \mathbb{I}_{E_n} + \frac{1}{\varepsilon} \mathbb{I}_{E_{n+1}} \frac{1}{A_{v_{n+1}} \beta_{v_{n+1}}(\omega)} \]

\[ \leq 1 + \frac{1}{\varepsilon} + \frac{\mathbb{I}_{E_{n+1}}}{A_{v_{n+1}} \beta_{v_{n+1}}(\omega)}. \]

Following [1] (proof of Lemma 2.2), by the i.i.d. structure of the environment, we have \( \mathbb{P}(E_n) = \mathbb{P}(E_1)^n \). By reiterating (5.5), we obtain

\[ \frac{1}{\beta_\phi} \leq \left( 1 + \frac{1}{\varepsilon} \right) \left( 1 + \sum_{n=1}^{\infty} B(n) \right), \]
where \( B(n) = \mathbb{1}_{E_n} \prod_{j=1}^{n} (A_{v_j})^{-1} \). Notice that for any sequence of non-negative numbers \((b_n)_n\), we have

\[
\left( \sum_{n=1}^{\infty} b_n \right)^p \leq \sum_{n=1}^{\infty} 2^{(p-1)n} b_n^p.
\]

In order to prove (5.7), it is enough to notice that

\[
\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} 2^{-n} 2^n b_n,
\]

and apply Jensen’s inequality. Next, we combine (5.6) and (5.7), to get

\[
\frac{1}{\beta_{\varrho}^p} \leq \left( 1 + \frac{1}{\varepsilon} \right)^p \left( 1 + \sum_{n=1}^{\infty} B(n) \right)^p
\]

\[
\leq \left( 1 + \frac{1}{\varepsilon} \right)^p 2^{p-1} \left( 1 + \sum_{n=1}^{\infty} 2^{(p-1)n} B(n)^p \right).
\]

We have

\[
E[B(n)^p] = E \left[ \mathbb{1}_{E_1} \left( A_{v_1} \right)^{-p} \right]^n =: c^n.
\]

In virtue of the definition of \( E_1 \) and the integrability condition (1.4), we can choose \( \varepsilon \) small enough that \( c < 1/2^{p-1} \). Hence, by taking expectations in (5.8), we get

\[
E \left[ \frac{1}{\beta_{\varrho}^p} \right] < \infty.
\]

**Proposition 15.** We have

\[
E \left[ L_{\varrho}^p \right] < \infty.
\]

**Proof.** In virtue of Remark 13 combined with Proposition 24 in the Appendix, we have

\[
E[L_{\varrho}^{p-1}] \leq 1 + C_p E \left[ \frac{\beta_{\varrho}}{(-\ln(1-\beta_{\varrho}))^{p+1}} \right]
\]

\[
\leq 1 + C_p E \left[ \frac{1}{\beta_{\varrho}^p} \right] < \infty,
\]

where in the last step we used the fact that \( x/(-\ln(1-x)) \in (0, 1) \) for any \( x \in (0, 1) \), and Proposition 14. Recall that the subtrees \( \Lambda_{v} \) was introduced at the beginning of Section 4. Next, denote by \( \Lambda_i \) the tree composed by \( \varrho, \varrho i, \ldots \).
all the descendants of \( \varrho \), and the edges connecting them. Denote by \( \tilde{L}_i \) the number of visits to \( \varrho \) by the extension \( X^\Lambda_i \), i.e.

\[
\tilde{L}_i := \sum_{k=0}^{\infty} 1_{\{X^\Lambda_i^k = \varrho\}}.
\]

Under the measure \( \mathbb{P} \), \( \tilde{L}_i \) and \( L^\varrho - 1 \) are equally distributed. Hence, \( \mathbb{E}[\tilde{L}^\varrho_i] < \infty \).

Finally

\[
L^\varrho \leq 1 + L^\varrho - 1 + \sum_{i=1}^{b} \tilde{L}_i,
\]

proving our result.

\textbf{Lemma 16.} There exists a collection of random variables \( (\bar{L}_\nu)_\nu \), such that \( \bar{L}_\nu \sim L^\varrho \) and \( L_\nu \leq \bar{L}_\nu \), a.s., for all \( \nu \in V \), with \( \nu \neq \varrho^{-1} \).

\textit{Proof.} Consider the extension \( X^\Lambda_\nu \). Set

\[
L_\nu := \sum_{k=0}^{\infty} 1_{\{X^\Lambda_\nu^k = \nu\}}.
\]

By the definition of the extension, \( \bar{L}_\nu \) shares the same distribution as \( L^\varrho \).

Moreover, \( \bar{L}_\nu \sim L^\varrho \) as \( X \) observed while in \( \Lambda_\nu \) and the extension coincide up to the last time the former process leaves \( \Lambda_\nu \) and never returns to it.

Let \( (\sigma_i)_i \) be the sequence vertices visited by the process \( X \), ordered chronologically. More precisely, \( \nu = \sigma_i \) if and only if \( \nu \) is the \( i \)-th distinct vertex visited by \( X \) and \( \ell_1 > |\nu| \).

\textbf{Proposition 17.} For any \( p > 0 \), there exists \( C_p \in (0, \infty) \), such that

\[
\sup_{k \in \mathbb{N}} \mathbb{E}[L^p_{\sigma_k}] < C_p.
\]

\textit{Proof.} Consider the random variable \( \bar{L}_\nu \) defined in (5.11). Notice that \( \bar{L}_\nu \) and the event \( \{\sigma_k = \nu\} \) are independent. In fact the latter is generated using the collection of Poisson processes

\[
\{Y(x, y) : \text{none of } x \text{ and } y \text{ are descendants of } \nu\},
\]

where the processes \( Y \) were introduced in (3.11) and each vertex \( \nu \) is NOT considered a descendant of itself, while the extension is generated using a disjoint collection of Poisson processes.
collection of Poisson processes. Hence
\[
\mathbb{E}[L_{\sigma_k}^p] = \mathbb{E}\left[ \sum_{\nu} L_{\nu}^p 1_{\sigma_k = \nu} \right] 
\leq \sum_{\nu} \mathbb{E}\left[ L_{\nu}^p 1_{\sigma_k = \nu} \right]
\]
\[
= \sum_{\nu} \mathbb{E}\left[ L_{\nu}^p \right] \mathbb{P}(\sigma_k = \nu)
\]
\[
= \mathbb{E}\left[ L_0^p \right] < \infty.
\]
Hence, by taking \( C_p = \mathbb{E}\left[ L_0^p \right] \) we conclude our proof. \( \square \)

**Proposition 18.** \( \mathbb{E}[\tau_1^2] < \infty. \)

**Proof.** The vertices \( \sigma_1, \sigma_2, \ldots, \sigma_\Pi \) are the vertices visited before time \( \tau_1 \). We have
\[
\mathbb{E}[\tau_1^2] \leq \mathbb{E}\left[ \left( \sum_{k=1}^{\Pi} L_{\sigma_k} \right)^2 \right] \leq \Pi \sum_{k=1}^{\Pi} L_{\sigma_k}^2.
\]
Let \( q > 1 \) the conjugate of \( p/2 \), i.e. \( (1/q) + (2/p) = 1 \). By Hölder’s inequality, and using Proposition 17 we have that
\[
\mathbb{E}\left[ \Pi \sum_{k=1}^{\Pi} L_{\sigma_k}^2 \right] = \mathbb{E}\left[ \sum_{k=1}^{\infty} L_{\sigma_k}^2 \mathbb{I}_{\{\Pi \geq k\}} \right]
\]
\[
\leq \sum_{k=1}^{\infty} \mathbb{E}\left[ L_{\sigma_k}^p \right]^{2/p} \mathbb{E}\left[ \Pi^q \mathbb{I}_{\{\Pi \geq k\}} \right]^{1/q}
\]
\[
\leq (C_p)^p \sum_{k=1}^{\infty} \mathbb{E}\left[ \Pi^q \mathbb{I}_{\{\Pi \geq k\}} \right]^{1/q} < \infty,
\]
where \( C_p \) is the same as in Proposition 17. \( \square \)

Finally, notice that
\[
\mathbb{E}[(\tau_2 - \tau_1)^2] = \mathbb{E}[\tau_1^2 | T_{\nu^{-1}} = \infty] \leq \mathbb{E}[\tau_1^2] \mathbb{P}(T_{\nu^{-1}} = \infty) < \infty,
\]
which proves (2.1).

### 6. Linearly Edge-Reinforced Random Walks

Define a discrete time process \( X \) as follows. It takes as values the vertices of \( G \). Initially all the edges are given weight one, and \( X_0 = \nu \). Whenever an edge is traversed, i.e. the process jumps from one of its endpoints to the other,
the weight of the edge is increased by one, and the process jumps to nearest neighbors with probabilities proportional to the weights of the connecting edges. Notice, that this process can be represented as a RWRE with the following environment (see [12]). To each vertex \( \nu \neq q^{-1} \), assign an independent \((b+1)\)-dimensional random vector \( Z_\nu = (Z^{(\nu)}_0, Z^{(\nu)}_1, \ldots, Z^{(\nu)}_b) \), distributed as a Dirichlet distribution, with parameters \((1, 1/2, 1/2, \ldots, 1/2)\). The distribution of the vector assigned to \( q \) is still a Dirichlet distribution but with different parameters. This exception does not affect our analysis, as we are interested in a limit theorem. Set

\[
A_{\nu i} = \frac{Z^{(\nu)}_i}{Z^{(\nu)}_0}.
\]

Theorem 1 can be applied to this process to yield a functional central limit theorem for \( b \geq 5 \). In fact, condition (1.2) is satisfied (see Pemantle [12]). We have

\[
\mathbb{E} \left[ \left( \sum_{1 \leq i \leq b} A_{\nu i} \right)^{-p} \right] = \mathbb{E} \left[ \left( \frac{Z^{(\nu)}_0}{1 - Z^{(\nu)}_0} \right)^p \right]
\]

\[
= \int_0^1 \frac{1}{\Gamma(b/2)} \Gamma(b/2 + 1) \left( \frac{x}{1-x} \right)^p \left(1 - x\right)^{(b/2) - 1} dx
\]

\[
= \frac{B(1 + p, b/2 - p)}{B(1, b/2)}
\]

which is finite if and only if \( b/2 - p > 0 \), i.e. \( b > 2p > 4 \). Hence condition (1.4) is satisfied for \( b \geq 5 \).

Remark 19. As in [12], let us consider the more general situation where each time the process traverses an edge, its weight is increased by \( \Delta > 0 \). In this case \( Z_\nu = (Z^{(\nu)}_0, Z^{(\nu)}_1, \ldots, Z^{(\nu)}_b) \) is distributed as a Dirichlet distribution with parameters \(((1 + \Delta)/(2\Delta), 1/(2\Delta), 1/(2\Delta), \ldots, 1/(2\Delta))\), and

\[
\mathbb{E} \left[ \left( \sum_{1 \leq i \leq b} A_{\nu i} \right)^{-p} \right] = \frac{B\left(1 + \Delta + p, \frac{b}{2\Delta} - p\right)}{B\left(1 + \Delta, \frac{b}{2\Delta}\right)}
\]

Thus condition (1.4) is satisfied if and only if \( 0 < \Delta < b/4 \).

Now we turn to the proof of Theorem 4 for the case \( b = 4 \). For any \( \nu \in V \), with \( |\nu| \geq 0 \), define

\[
T^+_\nu := \inf\{k > 0 \colon X_k = \nu\}, \quad \gamma_\nu(\omega) := P^\omega_\nu(T^{-1}_\nu = \infty, T^+_\nu = \infty).
\]

Recall the definition of \( \omega(\nu, \nu i) \) given in (1.1), and that for LERRW on the \( b \)-regular tree and for \( |\nu| \geq 1 \), we have \( \omega(\nu, \nu i) \) is distributed as Beta(1/2, (b+1)/2).
1)/2), while \( \omega(\nu, \nu^{-1}) \) is a Beta(1, \( b/2 \)). The reasoning presented in this section follows closely the one given in section 7 in [1].

**Remark 20.** We fix \( \varepsilon \in (0, 1/3) \), and we can assume \( \omega(\varrho, \varrho^{-1}) \leq 1 - \varepsilon \) without loss of generality. In fact, our goal is to prove a limit theorem, and the process is transient. Hence the influence of \( \omega(\varrho, \varrho^{-1}) \) vanishes in the limit.

**Proposition 21.** For LERRW on the 4-regular tree, there exists a positive finite constant \( C \) such that for any vertex \( \nu \) with \(|\nu| \geq 0\), we have

\[
\mathbb{E} \left[ \left( \frac{\mathbb{1}_{\{\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon\}}}{\gamma_{\nu}(\omega)} \right)^{\frac{28}{9}} \right] \leq C.
\]

**Proof.** From now on, fix \( \delta = 35/18 \), \( \chi = 9/28 \), and \( \alpha = 71/280 \). From the proof of Lemma 7.1 in [1], we have

\[
\mathbb{E} \left[ (\beta_{\nu})^{-\delta} \right] < \infty.
\]

Moreover

\[
\frac{1}{\gamma_{\nu}(\omega)} = \sum_{i=1}^{4} \frac{1}{\omega(\nu, \nu_i) \beta_{\nu_i}} \leq \min_{1 \leq i \leq 4} \frac{1}{\omega(\nu, \nu_i) \beta_{\nu_i}}.
\]

By Fubini’s theorem, we have

\[
\mathbb{E} \left[ \left( \frac{\mathbb{1}_{\{\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon\}}}{\gamma_{\nu}(\omega)} \right)^{\frac{1}{\chi}} \right] \leq \mathbb{E} \left[ \left( \min_{1 \leq i \leq 4} \frac{1}{\omega(\nu, \nu_i) \beta_{\nu_i}} \right)^{\frac{1}{\chi}} \mathbb{1}_{\{\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon\}} \right] = \int_{0}^{\infty} \mathbb{P}(\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon, \min_{1 \leq i \leq 4} (\omega(\nu, \nu_i) \beta_{\nu_i})^{-1/\chi} \geq n)dn.
\]

Notice that

\[
\{(\omega(\nu, \nu_i) \beta_{\nu_i})^{-1} \geq n^\chi\} \subset \{\omega(\nu, \nu_i)^{-1} \geq n^\alpha\} \cup \{(\beta_{\nu_i})^{-1} \geq n^{\chi - \alpha}\} = E_1^i \cup E_2^i
\]
for each of $i$. On the event \{\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon\}, there exists $1 \leq i \leq 4$ such that \(\omega(\nu, \nu_i) \geq \varepsilon/4\). By symmetry,

\begin{equation}
\mathbb{P}(\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon, \min_{1 \leq i \leq 4} \omega(\nu, \nu_i) - \beta_{\nu_i}^{-1/\chi} \geq n) \leq 4 \mathbb{P}(\omega(\nu, \nu_4) \geq \varepsilon/4, \min_{1 \leq i \leq 4} \omega(\nu, \nu_i) - \beta_{\nu_i}^{-1/\chi} \geq n) \quad \text{(union bound)}
\end{equation}

\begin{equation}
= 4 \mathbb{P}(\omega(\nu, \nu_4) \geq \varepsilon/4, (\beta_{\nu_4})^{-1} \geq n^{\chi} \omega(\nu, \nu_4), \min_{1 \leq i \leq 3} (\omega(\nu, \nu_i) - \beta_{\nu_i})^{-1/\chi} \geq n) \leq 4 \mathbb{P}((\beta_{\nu_4})^{-1} \geq n^{\chi} \varepsilon/4, \min_{1 \leq i \leq 3} (\omega(\nu, \nu_i) - \beta_{\nu_i})^{-1/\chi} \geq n)
\end{equation}

\begin{equation}
\leq 4 \sum_{(k_1, k_2, k_3) \in \{1, 2\}^3} \mathbb{P}((\beta_{\nu_4})^{-\delta} \geq n^{\delta \chi}(\varepsilon/4) \delta, \cap_{i=1}^3 E_i^{k_i}) \quad \text{(independence)}
\end{equation}

\begin{equation}
\leq 4c_0 n^{-\delta \chi} \sum_{(k_1, k_2, k_3) \in \{1, 2\}^3} \mathbb{P}(\cap_{i=1}^3 E_i^{k_i}) \quad \text{(Markov's ineq. and (6.1))}
\end{equation}

\begin{equation}
= 4c_0 n^{-\delta \chi} \left[ \mathbb{P}(E_1^2)^3 + 3\mathbb{P}(E_1^2)^2 \mathbb{P}(E_1^4) + 3\mathbb{P}(E_1^2) \mathbb{P}(E_1^2 \cap E_1^4) + \mathbb{P}(\cap_{i=1}^3 E_i^4) \right].
\end{equation}

In the last equality we used independence and symmetry. We have $\mathbb{P}(E_1^2) \leq c_1 n^{-\delta(\chi - \alpha)}$ again by Markov’s inequality. Since $\omega(\nu, \nu_3)$ is a Beta(1/2, 5/2), we have $\mathbb{P}(E_1^4) \leq c_2 n^{-\alpha/2}$. Notice that conditionally to $\omega(\nu, \nu_3)$, the random variable $\omega(\nu, \nu_2)/\{1 - \omega(\nu, \nu_3)\}$ is distributed as a Beta(1/2, 2). Hence

$$\mathbb{P}(E_1^2 \cap E_1^4) = \mathbb{P}(E_1^4) \mathbb{P}(E_1^2 | E_1^4) \leq c_3 n^{-\alpha}.$$ In the same way, we have $\mathbb{P}(\cap_{i=1}^3 E_i^4) \leq c_4 n^{-3\alpha/2}$. Therefore, using (6.2), we have

$$\mathbb{P}(\omega(\nu, \nu^{-1}) \leq 1 - \varepsilon, \min_{1 \leq i \leq 4} (\omega(\nu, \nu_i) - \beta_{\nu_i})^{-1/\chi} \geq n) \leq 3 \sum_{i=0}^3 c_i n^{-\delta \chi} n^{-\delta(\chi - \alpha) i} n^{-(3-i)\alpha/2}.$$

With our choice of $\alpha, \delta$ and $\chi$, we have $\delta \chi + 3\alpha/2 > 1$ and $\delta \chi + 3\delta(\chi - \alpha) > 1$, and this completes the proof. ■

**Proposition 22.** For LERRW on the 4-regular tree, $E[L_y^3] < \infty$ for any $y \in V$.

**Proof.** Define, for $\nu \in V$ and $t \geq 0$,

$$L_\nu(t) := \sum_{k=0}^t \mathbb{1}_{\{X_k = \nu\}},$$

$$L_\nu^+(t) := \sum_{k=0}^{t-1} \mathbb{1}_{\{X_k, X_{k+1} = (\nu^{-1}, \nu)\}},$$
\[ L_\nu^-(t) := \sum_{k=0}^{t-1} \mathbb{1}_{\{ (X_k, X_{k+1}) = (\nu, \nu^{-1}) \}}. \]

Noting that \( L_\nu^-(t) \leq L_\nu^+(t) \) for \( \nu \neq \rho \), we have
\[
L_\nu(t) = \mathbb{1}_{\{ \nu = \rho \}} + L_\nu^+(t) + \sum_{i=1}^{b} L_{\nu_i}^-(t) \leq 1 + L_\nu^+(t) + \sum_{i=1}^{b} L_{\nu_i}^+(t).
\]

Let \( L_y^+ := L_y^+(\infty) \). To prove \( \mathbb{E}[L_y^3] < \infty \), it is enough to show \( \mathbb{E}[(L_y^+)^3] < \infty \).

Recall that \( \omega(\rho, \rho^{-1}) \leq 1 - \varepsilon \) (see Remark 20). For a vertex \( y \), and let \( Y \) be the youngest ancestor of \( y \), with \( \omega(Y, Y^{-1}) \leq 1 - \varepsilon \). More precisely \( Y \) is the vertex \( z \) on the unique self-avoiding path connecting \( y \) to \( \rho \), which has maximum distance from \( \rho \) and satisfies \( \omega(z, z^{-1}) \leq 1 - \varepsilon \). We have \( \omega(g, g^{-1}) > 1 - \varepsilon \) for all ancestors \( g \) of \( y \) with \( |Y| < |g| < |y| \). Let \( Y_y^+ := \inf \{ i > T_Y : X_i = y \} \), that is the hitting time of \( y \) after \( T_Y \). Notice that, on \( \{ X_0 = y \} \),
\[
L_y^+ = L_y^+(T_Y) + \mathbb{1}_{\{T_Y < \infty\}} \mathbb{1}_{\{ T_Y^y < \infty \}} \tilde{L}_y^+,
\]
where \( \tilde{L}_y^+ \) is an independent copy of \( L_y^+ \). Coupling with the simple random walk on the path \([Y, y]\), we prove next that
\[
P^y(\omega(T_y^+ < T_Y) \leq \frac{2}{3}.
\]
In fact, if \( Y \notin \{ y, y^{-1} \} \), we have that
\[
P^y(\omega(T_y^+ < T_Y) \leq \max_{a \in [0, 1]} \left( a \cdot 1 + (1 - a) \frac{\varepsilon}{1 - \varepsilon} \right),
\]
where the term \( \varepsilon/(1-\varepsilon) \) is derived using a coupling with a biased random walk. As \( \varepsilon/(1-\varepsilon) < 1 \), we have that the maximum is attained at \( a = \varepsilon \), which implies
\[
P^y(\omega(T_y^+ < T_Y) \leq \varepsilon + (1 - \varepsilon) \cdot \frac{\varepsilon}{1 - \varepsilon} = 2\varepsilon \leq \frac{2}{3},
\]
proving (6.4) in the case \( Y \notin \{ y, y^{-1} \} \). If \( Y = y^{-1} \), then
\[
P^y(\omega(T_y^+ < T_Y) \leq \varepsilon \leq \frac{2}{3},
\]
while if \( Y = y \) then
\[
P^y(\omega(T_y^+ < T_Y) = 0 \leq \frac{2}{3}.
\]

Hence \( L_y^+(T_Y) \) is stochastically dominated by a geometric distribution with average 3, that is \( \mathbb{E}[L_y^+/(T_Y)] \leq 1/(1 - 2/3) = 3 \). This implies that there exist positive constants \( c_5, c_6 \) such that
\[
\mathbb{E}^y[L_y^+(T_Y)^2] \leq c_5, \quad \mathbb{E}^y[L_y^+(T_Y)^3] \leq c_6.
\]
Therefore, by taking expectations in both sides of (6.3), and using strong
Markov property, we have

\[ E^\gamma_y[L^+_y] \leq E^\gamma_y[L^+_y(T_y)] + P^\gamma_y(T_y < \infty)P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y] \]

\[ \leq E^\gamma_y[L^+_y(T_y)] + P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y]. \]

After rearranging,

\[ E^\gamma_y[L^+_y] \leq \frac{E^\gamma_y[L^+_y(T_y)]}{1 - P^\gamma_y(T_y < \infty)} \leq \frac{E^\gamma_y[L^+_y(T_y)]}{\gamma_y(\omega)} \leq \frac{3}{\gamma_y(\omega)}. \]

As for the second moment, we have

\[ E^\gamma_y[(L^+_y)^2] \leq E^\gamma_y[L^+_y(T_y)^2] + 2E^\gamma_y[L^+_y(T_y)]P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y] \]

\[ + P^\gamma_y(T_y < \infty)E^\gamma_y[(L^+_y)^2], \]

and

\[ E^\gamma_y[(L^+_y)^2] \leq \frac{E^\gamma_y[L^+_y(T_y)^2] + 2E^\gamma_y[L^+_y(T_y)]P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y]}{\gamma_y(\omega)} \]

\[ \leq \frac{c_5 + 6 \cdot \frac{3}{\gamma_y(\omega)}}{\gamma_y(\omega)} \leq \frac{c_7}{\gamma_y(\omega)^2}. \]

Turning to the third moment, we have

\[ E^\gamma_y[(L^+_y)^3] \leq E^\gamma_y[L^+_y(T_y)^3] + 3E^\gamma_y[L^+_y(T_y)^2]P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y] \]

\[ + 3E^\gamma_y[L^+_y(T_y)]P^\gamma_y(T_y < \infty)E^\gamma_y[(L^+_y)^2] + P^\gamma_y(T_y < \infty)E^\gamma_y[(L^+_y)^3], \]

and

\[ E^\gamma_y[(L^+_y)^3] \leq \frac{c_6 + 3c_5P^\gamma_y(T_y < \infty)E^\gamma_y[L^+_y] + 9P^\gamma_y(T_y < \infty)E^\gamma_y[(L^+_y)^2]}{\gamma_y(\omega)} \]

\[ \leq \frac{c_8}{\gamma_y(\omega)^3}. \]

Recall that \( \chi = 9/28 \). Using the Markov property and Hölder’s inequality,

\[ E[(L^+_y)^3] = E\left[ E^\gamma_y[(L^+_y)^3] \right] = E\left[ P^\gamma_y(T_y < \infty)E^\gamma_y[(L^+_y)^3] \right] \]

\[ \leq E\left[ \frac{c_8}{\gamma_y(\omega)^3} \right] \]

\[ = \sum_{z \in \{z, y^{-1}\}} E\left[ \mathbb{1}_{(Y = z)} \mathbb{1}_{(\omega(z^{-1}) \leq 1 - \epsilon)} \left( \frac{c_8}{\gamma_z(\omega)^3} \right) \right] \]

\[ \leq c_8 \sum_{z \in \{z, y^{-1}\}} P(Y = z)^{1 - 3\chi} \mathbb{E}\left[ \left( \frac{\mathbb{1}_{(\omega(z^{-1}) \leq 1 - \epsilon)}}{\gamma_z(\omega)} \right)^{\frac{1}{\chi}} \right]^{3\chi}, \]
For each fixed ray \( \sigma = (\nu_i)_{i \in \mathbb{N}} \) where \( \nu_{i+1} \sim \nu_i \) and \( |\nu_{i+1}| = |\nu_i| + 1 \), we have that the process \( \omega(\nu_{i+1}, \nu_i) \) is composed by i.i.d. random variables. Hence \( \mathbb{P}(Y = z) \leq \mathbb{P}(\omega(g1, g) > 1 - \varepsilon)^{|y| - |z|} \) for any ancestor \( z \) of \( y \), and we have

\[
\mathbb{E}[(L_y^+)^3] \leq c_9 \sum_{n=0}^{\infty} \mathbb{P}(\omega(g1, g) > 1 - \varepsilon)^{(1-3\lambda)n} < \infty.
\]

This completes the proof.

The next result is a by-product of Proposition 22 combined with the proof of Proposition 18.

**Proposition 23.** Consider LERRW on the 4-regular tree. For any \( p \in (0, 3) \), we have \( \mathbb{E}[\tau_p^+ < \infty]. \)

**Proof.** It is enough to prove the proposition for \( p \in (1, 3) \). We have

\[
E[\tau_p^+] \leq E\left[ \left( \prod_{k=1}^{\Pi} L_{\sigma_k} \right)^p \right] \leq E\left[ \prod_{k=1}^{\Pi} L_{\sigma_k}^p \right] .
\]

Choose \( t, q > 1 \) conjugates, i.e. \((1/t) + (1/q) = 1\), and \( tp < 3 \). By Hölder’s inequality, and using Proposition 17, we have that

\[
E\left[ \prod_{k=1}^{\Pi} L_{\sigma_k}^p \right] \leq \sum_{k=1}^{\infty} E\left[ L_{\sigma_k}^{tp} \right]^{1/t} E\left[ (p-1)^{q-1} \mathbb{I}_{\{\Pi \geq k\}} \right]^{1/q} \leq (C_{tp})^t \sum_{k=1}^{\infty} E\left[ (p-1)^{q-1} \mathbb{I}_{\{\Pi \geq k\}} \right]^{1/q} < \infty,
\]

where \( C_{tp} \) is the same as in Proposition 17.

7. **Appendix**

**Proposition 24.** Let \( p > 0 \). Consider a random variable \( Y \) geometrically distributed with parameter \( \theta \in (0, 1) \), and probability mass function

\[
\mathbb{P}(Y = k) = (1 - \theta)^k \theta, \quad \text{with } k \geq 0.
\]

If we set \( \lambda := -\ln(1 - \theta) \), we have

\[
\mathbb{E}[Y^p] \leq C_p \frac{\theta}{\lambda^{p+1}} + 1,
\]
where $C_p$ is a positive finite constant that depends on $p$ but not on $\theta$.

Proof. First notice that the function $f: x \mapsto x^p e^{-\lambda x}$, for $x \geq 0$, achieves its unique maximum at $x^* = p/\lambda$. As $f$ is non-negative, and it is decreasing in the interval $[x^*, \infty)$, we have the following estimate

$$\sum_{k=0}^{\infty} f(k) \leq x^* f(x^*) + \int_{0}^{\infty} f(u) du.$$ 

Hence

$$\mathbb{E}[Y^p] = \sum_{k=0}^{\infty} k^p (1 - \theta)^k \theta 
\leq 1 + \theta \frac{p}{\lambda} \left( \frac{p}{\lambda} \right)^p e^{-\theta} + \theta \int_{0}^{\infty} x^p e^{-x \lambda} dx 
= 1 + \theta \frac{p}{\lambda} \left( \frac{p}{\lambda} \right)^p e^{-\theta} + \theta \frac{\Gamma(p+1)}{\lambda^{p+1}} 
= : 1 + C_p \frac{\theta}{\lambda^{p+1}}.$$ 

Acknowledgement. A.C. is grateful to Yokohama National University for its hospitality, and he was supported by ARC grant DP180100613, Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS). CE140100049, and YNU iROUTE project. M.T. is partially supported by JSPS Grant-in-Aid for Young Scientists (B) No. 16K21039. The authors thank the anonymous referee for detailed comments. Finally they thank Amanoya for offering a very nice environment, where part of this research was carried.

References

[1] Aïdèkon, E. (2008). Transient random walks in random environment on a Galton-Watson tree. *Probability Theory and Related Fields* **142**(3-4), 525-559.

[2] Barbour, A. and Collevecchio, A. (2017). General random walk in a random environment defined on Galton–Watson trees. *Ann. Inst. H. Poincaré Probab. Statist.* **53**(4), 1657–1674.

[3] Bogachev, L. V. (2006). Random walks in random environments. *Encyclopedia of Mathematical Physics* **4**, 353–397.

[4] Collevecchio, A. (2006). Limit theorems for reinforced random walks on certain trees. *Probability Theory and Related Fields* **136**(1), 81–101.

[5] Collevecchio, A. (2009). Limit theorems for vertex-reinforced jump processes on regular trees. *Electron. J. of Probab.* **14**(66), 1936–1962.
[6] Collevecchio, A., Kious, D., and Sidoravicius, V. (2018). The branching-ruin number and the critical parameter of once-reinforced random walk on trees. *Communications on Pure and Applied Mathematics*, Forthcoming.

[7] Coppersmith, D. and Diaconis, P. (1986). Random walks with reinforcement. *Unpublished manuscript*.

[8] Davis, B. (1990). Reinforced random walk. *Probability Theory and Related Fields* **84**, 203–229.

[9] Durrett, R., Kesten, H., and Limic, V. (2002). Once edge-reinforced random walk on a tree, *Probability Theory and Related Fields* **122**(4), 567–592.

[10] Kesten, H., Kozlov, M. V., and Spitzer, F. (1975). A limit law for random walk in a random environment, *Comp. Math.* **30**, 145–168.

[11] Lyons, R. and Pemantle, R. (1992). Random walk in a random environment and first-passage percolation on trees. *Ann. Probab.* **20**(1), 125–136.

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

[13] Peres, Y. and Zeitouni, O. (2008). A Central Limit Theorem for biased random walks on Galton-Watson trees. *Probab. Theory Relat. Fields* **14**(3–4), 595–629.

[14] Sabot, C. and Tarrès, P. (2015). Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. *J. Eur. Math. Soc.* **17**(9), 2353–2378.

[15] Solomon, F. (1975). Random walks in random environments. *Ann. Probab.* **3**, 1–31.

[16] Zeitouni, O. (2004). Random walks in random environment. *Lecture Notes in Math.* **1837**, 189–312, Springer.

[17] Zhang, Y. (2014). Large deviations in the reinforced random walk model on trees. *Probability Theory and Related Fields* **160**(3–4), 655–678.

Andrea Collevecchio, School of Mathematical Sciences, Monash University, Melbourne

*E-mail address*: Andrea.Collevecchio@monash.edu

Masato Takei, Department of Applied Mathematics, Faculty of Engineering, Yokohama National University

*E-mail address*: takei-masato-fx@ynu.ac.jp

Yuma Uematsu, Department of Systems Integration, Graduate School of Engineering, Yokohama National University

*E-mail address*: uematsu-yuuma-zs@ynu.jp