Speed of Vertex reinforced jump process on Galton-Watson trees

Xinxin Chen, Xiaolin Zeng

Abstract

We give an alternative proof of the fact that the vertex reinforced jump process on Galton-Watson tree has a phase transition between recurrence and transience as a function of $c$, the initial local time, see [3]. Further, applying techniques in [1], we show a phase transition between positive speed and null speed for the associated discrete time process in the transient regime.

1 Introduction and results

Let $G = (V, E)$ be a locally finite graph endowed with its vertex set $V$ and edge set $E$. Assign to each edge $e = \{u, v\} \in E$ a positive real number $W_e = W_{u,v}$ as its conductance, and assign to each vertex $u$ a positive real number $\phi_u$ as its initial local time. Define a continuous-time $V$ valued process $(Y_t; t \geq 0)$ on $G$ in the following way: At time 0 it starts at some vertex $v_0 \in V$; If $Y_t = v \in V$, then conditionally on $\{Y_s; 0 \leq s \leq t\}$, the process jumps to a neighbor $u$ of $v$ at rate $W_{v,u}L_u(t)$ where

$$L_u(t) := \phi_u + \int_0^t 1_{\{Y_s = u\}} ds.$$ (1)

We call $(Y_t)_{t \geq 0}$ the vertex reinforced jump process (VRJP) on $(G, W)$ starting from $v_0$.

It has been proved in [6] that when $G = \mathbb{Z}$, $(Y_t)$ is recurrent. When $G = \mathbb{Z}^d$ with $d \geq 2$, the complete description of its behavior has not been revealed even though lots of effort has been made, see e.g. [2 3 5 6 7 16].

Here we are interested in the case when $G$ is a supercritical Galton-Watson tree, as we will see, acyclic property of trees largely reduces the difficulty to study this model. In [5] it is shown that the VRJP (for constant parameters $W_{v,u} \equiv 1$ and $\phi_x \equiv 1$) on 3-regular tree has positive speed and satisfies a central limit theorem. Later, Basdevant and Singh [3] gave a precise description of the phase transition of recurrence/transience for VRJP on supercritical Galton-Watson trees. In this paper, our main results, Theorem 2 describes the ballistic case of the VRJP when it is transient on supercritical Galton-Watson trees without leaves. Our proof is based on the random walk in random environment (RWRE) representation result of Sabot and Tarrès [16], and on techniques in the studies of RWRE on trees, especially on a result of Aidekon [1] (see also e.g. [10 11] for more on the studies of RWRE on trees).

Consider a rooted Galton-Watson tree $T$ with offspring distribution $(q_k, k \geq 0)$ such that

$$b := \sum_{k \geq 0} kq_k > 1.$$ 

For some constant $c > 0$, we denote VRJP$(c)$ the process $(Y_t)$ on the Galton-Watson tree $T = (V, E)$ with $W_e \equiv 1$, $\forall e \in E$ and $\phi_x \equiv c$, $\forall x \in V$, starting from the root $\rho$. Hence the behaviors of this
process depends on $G$ and $c$. This definition is equivalent to VRJP with constant edge weight $W$ and initial local time $1$, up to a time change. We first recall the phase transition result obtained in [6]. Let $A$ be an inverse Gaussian distribution of parameters $(1,c^2)$, i.e.

$$P(A \in dx) = 1_{x \geq 0} \frac{c}{\sqrt{2\pi x^3}} \exp \left\{ - \frac{c^2(x-1)^2}{2x} \right\} dx,$$

(2)

The expectation w.r.t. $P(dx)$ is denoted $E$.

**Theorem 1** (Basdevant & Singh). Let $\mu(c) = \inf_{a \in \mathbb{R}} E[A^a] = E[\sqrt{A}]$, then the VRJP$(c)$ on a supercritical GW tree with offspring mean $b$ is recurrent a.s. if and only if $b\mu(c) \leq 1$.

**Remarks 1.** This phase transition was proved in [3] by considering the local times of VRJP. We will give another proof from the point of view of a random walk in random environment (RWRE), as a consequence of Theorem 3.

When $b\mu(c) > 1$, a further question is to study the escape rate of the process. Define the speed of the process $(Y)$ by

$$v(Y) := \liminf_{t \to \infty} \frac{d(\rho, Y_t)}{t} = \lim_{t \to \infty} \frac{d(\rho, Y_t)}{t}$$

(3)

where $d$ is the graph distance, and the last equality will be justified by Lemma [1]. To study the speed, we use the RWRE point of view, relying on a result of Sabot & Tarrès [16], in particular, on the following fact:

Let $(Y_t)$ be a VRJP on a finite graph $G = (V,E)$ with edge weight $(W)$ and initial local time $(\phi)$. If $(Z_t)$ is defined by

$$Z_t := Y_{D^{-1}(t)} \text{ where } D(t) := \sum_{x \in V} (L_x(t)^2 - \phi_x^2),$$

(4)

then $(Z_t)$ is a mixture of Markov jump processes (c.f. also [17]). Moreover, the mixing measure is explicit.

Applying this result to our VRJP$(c)$ on a tree, denote $(\eta_n)_{n \geq 0}$ the discrete time process associated to $(Z_t)$, it turns out that $(\eta_n)$ is a random walk in random environment. In [1], Aidekon gave a sharp and explicit criterion for the asymptotic speed to be positive, for random walks in random environment on Galton-Watson trees such that the environment is site-wise independent and identically distributed. This result cannot apply directly to the time changed VRJP$(c)$, since the quenched transition probability depends also on the environment of the neighbors, see (7).

Aidekon’s idea was to say that, the slowdown comes from wanders in long pipes, therefore, the random walk is roughly a trapped random walk. To study the speed, it is enough to look at the random walk on the traps, that is, pipes. This also explains why the criterion depends on $q_1$, the probability that the GW tree generates one offspring.

In our case the environment is almost i.i.d., the same idea will also work. Compare to [1], we mainly deal with the local dependences of the quenched probability transition. We believe that same type of criterion also holds for a larger type of random walk in random environment, with suitable conditions on the moments of the environment and locality of the transition probabilities.

Let us state our criterion, similar to (3), define

$$v(Z) := \liminf_{t \to \infty} \frac{d(\rho, Z_t)}{t}, \quad v(\eta) := \liminf_{n \to \infty} \frac{d(\rho, \eta_n)}{n}.$$  

(5)
To study the speed, our techniques can only deal with trees without leaves, hence we assume that \( q_0 = 0 \). In addition, we assume that
\[
M := \sum_{k \geq 0} k^2 q_k < \infty.
\]
For any \( r \in \mathbb{R} \), let
\[
\xi_r = \xi_r(c) := \mathbb{E}[A^{-r}].
\]
By (2), \( \xi_r \in (0, \infty) \) for any \( r \). In particular, \( \mu(c) = \xi_{-1/2}(c) \). Our main theorem states that the speed depends on the value of \( q_1 \) and \( c \).

**Theorem 2.** Consider VRJP\((c)\) on a supercritical GW tree such that \( b\mu(c) > 1 \), we have
\[ (1) \lim_{t \to \infty} d(\rho, Z_t) t \quad \text{and} \quad \lim_{n \to \infty} \frac{d(\rho, Z_n)}{n} \text{ exist almost surely}; \]
\[ (2) \text{Assume } q_0 = 0 \text{ and } M < \infty. \text{ If } q_1 \xi_{1/2} < 1, \text{ then } v(\eta) > 0, \text{ } v(Z) > 0; \text{ if } q_1 \xi_{1/2} > 1, \text{ then } v(\eta) = v(Z) = 0. \]

**Remarks 2.** Let \( \Lambda \) be the Lebesgue measure of \{ \( t \in \mathbb{R} : \mathbb{E}[A^{2t}] < 1/q_1 \) \}. \( q_1 \xi_{1/2} < 1 \) is equivalent to the condition \( \Lambda > 1 \). In other words, same as in Theorem 1.1 of [1] if \( \Lambda > 1 \), the walk has positive speed; if \( \Lambda < 1 \), its speed is null.

**Corollary 1.** VRJP\((c)\) \((Y_t)_{t \geq 0}\) on a supercritical GW tree such that \( b\mu(c) > 1 \), admits a speed \( v(Y) \geq 0 \) a.s. If in addition \( q_0 = 0 \), \( M < \infty \) and \( q_1 \xi_{1/2} < 1 \), then \( v(Y) > 0 \).

**Remarks 3.** Our method cannot tackle the critical case \( q_1 \xi_{1/2} = 1 \). Moreover, whether \( q_1 \xi_{1/2} > 1 \) implies \( v(Y) = 0 \) remains unknown, since we do not have estimates on the random time change.

The rest of this paper is organized as follows. In Section 2 we use a result of Sabot & Tarres [16] to recover the RWRE structure of VRJP. Section 3 is devoted to an alternative proof of Theorem 1 as an application of the RWRE point of view. Section 4 establishes the existence of the speed for the RWRE and prove Theorem 2. The proofs of some technical lemmas are left in Appendix.

## 2 RWRE on Galton-Watson tree

### 2.1 Mixture of Markov jump process by changing times

In this subsection, we consider a VRJP \((Y_t)_{t \geq 0}\) on a tree \( T = (V, E) \) rooted at \( \rho \), with edge weights \((W)\) and initial local time \((\phi)\). If \( x \neq \rho \), let \( \overleftarrow{x} \) be the parent of \( x \) on the tree, the associated edge is denoted by \( e_x = (x, \overleftarrow{x}) \) with weight \( W_{e_x} \).

Recall that the time changed version of VRJP \((Z_t)\) defined in (4) is mixture of Markov jump processes with correlated mixing measure. The advantage of considering VRJP on trees is that, the random environment becomes independent.

**Theorem 3.** Let \( T = (V, E) \) be a tree rooted at \( \rho \), endowed with edge weights \((W_e)_{e \in E}\) and initial local times \((\phi_x)_{x \in V}\). Let \( (A_x, x \in V \setminus \{ \rho \}) \) be independent random variables defined by
\[
P(A_x \in da) = \mathbb{1}_{\mathbb{R}^+}(a) \phi_x \sqrt{\frac{W_{e_x} \phi_x \phi_{\overleftarrow{x}}}{2\pi a^3}} \exp\left(-W_{e_x} \phi_x \phi_{\overleftarrow{x}} \left(\frac{a - 1}{2a}\right)^2\right) da.
\]
If \( X_t \) is a mixture of Markov jump processes starting from \( \rho \), such that, conditionally on \( (A_x, x \in V \setminus \{ \rho \}) \), \( X_t \) jumps from \( x \) to \( \overleftarrow{x} \) at rate \( \frac{1}{2} W_{e_x} \phi_{\overleftarrow{x}} A_x \), and from \( \overleftarrow{x} \) to \( x \) at rate \( \frac{1}{2} W_{e_x} \phi_x A_x \). Then \( X_t \) and \( Z_t \) (defined in (4)) has the same distribution.
Proof. On trees, VRJP observed at times when it stays on any finite sub-tree \( T_f = (V_f, E_f) \) (also rooted at \( \rho \)) of \( T \), behaves the same as VRJP restricted to \( T_f \); moreover, the restriction is independent of the VRJP outside \( T_f \). Therefore, it is enough to prove the theorem on finite tree \( T_f \). By Theorem 2 of \[15\] (with a slight modification of the initial local time, or a more detailed version in \[15\], appendix B), if we denote

\[
l_x(t) = \int_0^t \mathbb{1}_{Z_s = x} ds,
\]

then

\[
U_x = \frac{1}{2} \lim_{t \to \infty} \left( \log \frac{l_x(t) + \phi_x^2}{l_\rho(t) + \phi_\rho^2} - \log \frac{\phi_x^2}{\phi_\rho^2} \right)
\]

exists a.s. and \( \{U_x : x \in V_f\} \) has distribution (where \( du = \prod_{x \neq \rho} du_x \))

\[
dQ_{\rho,T_f}^{W,\phi}(u) = \prod_{x=0 \atop x \neq \rho} \frac{\phi_x}{\sqrt{2\pi}} e^{-\sum_{x \in V_f} u_x - \sum_{(x,y) \in E_f} \frac{1}{2} W_{x,y}(e^{u_x-u_y} \phi_x^2 + e^{u_y-u_x} \phi_y^2 - 2 \phi_y)} \sqrt{\prod_{\{x,y\} \in E_f} W_{x,y} e^{u_x+u_y}} du.
\]

Now, conditionally on \( (U_x) \), \( Z_t \) is a Markov process which jumps at rate (from \( x \) to \( z \)) \( \frac{1}{2} W_{x,z} e^{U_z-U_x} \).

For \( e_x = (x, \xleftarrow{\to}) \in T_f \), if we apply the change of variable \( y_{e_x} = (u_x - \log \phi_x) - (u_x - \log \phi_x) \), then (note that \( u \mapsto y \) is a diffeomorphism and \( dy = du \)) the density of \( (y) \) writes

\[
dQ_{\rho,T_f}^{W,\phi}(y) = \prod_{e_x = (x, \xleftarrow{\to}) \in E_f} \sqrt{W_{e_x} \phi_x \phi_{\xleftarrow{\to}}} 2\pi \exp \left( \frac{1}{2} (y_{e_x} - W_{e_x} \phi_x \phi_{\xleftarrow{\to}} (e^{y_{e_x}} + e^{-y_{e_x}}) - 2) \right) dy.
\]

Plugging \( a_x = e^{-y_{e_x}} \) entails that \( a_x \) is Inverse Gaussian distributed with parameter \( (1, W_{e_x} \phi_x \phi_{\xleftarrow{\to}}) \) i.e.

\[
dQ_{\rho,T_f}^{W,\phi}(a) = \prod_{x \in V_f \setminus \{\rho\}} 1_{a_x > 0} \sqrt{W_{e_x} \phi_x \phi_{\xleftarrow{\to}}} 2\pi a_x \exp(-W_{e_x} \phi_x \phi_{\xleftarrow{\to}} (a_x - 1)^2) da_x
\]

Finally note that

\[
\frac{1}{2} W_{x,z} e^{u_x-u_z} = \begin{cases} \frac{1}{2} W_{x,z} \frac{\phi_x}{\phi_{\xleftarrow{\to}}} & \text{if } z = \xleftarrow{\to} \smallsmile \xrightarrow{\to} \\ \frac{1}{2} W_{x,z} \frac{\phi_{\xleftarrow{\to}}}{\phi_x} & \text{if } \xleftarrow{\to} = x. \end{cases}
\]

\[\square\]

For VRJP\((c)\) on a GW tree, the theorem immediately implies:

**Corollary 2.** On a sampled GW tree \( T = (V, E) \), the time changed VRJP\((c)\) (\( Z_t \)) is a random walk in environment given by \( (A_x, x \in V \setminus \{\rho\}) \), where \( (A_x) \) are i.i.d. inverse Gaussian distributed with parameters \((1, c^2)\), and conditionally on the environment, the process jumps at rate

\[
\begin{cases} \frac{1}{2A_x} & \text{from } x \to \xleftarrow{\to} \\ \frac{1}{2A_x} & \text{from } \xrightarrow{\to} x \to x. \end{cases}
\]

(6)

### 2.2 RWRE on Galton Watson tree and notations

In the sequel, let \( T = (V, E) \) be a Galton-Watson tree with offspring distribution \( \{q_k ; k \geq 0\} \). Recall that \((\eta_n)_{n \geq 0}\) denotes the discrete time process associated to \( (Z_t) \) (or \( (Y_t) \)), which is a random walk in random environment.
Note that there are two levels of randomnesses in the environment. First, we sample a GW tree, \( T \), whose law is denoted by \( GW(dT) \). Then, given the tree \( T \) (rooted at \( \rho \)), we define \( \omega = \{ A_x, x \in V \setminus \{ \rho \} \} \) as in Corollary 2 whose law is \( \prod_{x \in T\setminus\{\rho\}} \mathbb{P}(dA_x) \), which we denote abusively \( \mathbb{P}(d\omega) \). Finally, given \( (w, T) \), the Markov jump process \( (Z_t; t \geq 0) \) is defined by its jump rate in (6).

For convenience, we artificially add a vertex \( \overrightarrow{\rho} \) to \( T \), designing the parent of the root. Let \( A_\rho \) be another copy of \( A \), independent of all others. Now, (abusively) let \( \omega = (A_\rho, x \in V) \) be the enlarged environment. Given \( (\omega, T) \), define the new Markov chain \( \eta \), which is a random walk on \( V \cup \{ \overrightarrow{\rho} \} \), with transition probabilities

\[
\begin{align*}
\left\{ \begin{array}{l}
    p(x, \overrightarrow{x}) \propto 1/A_x \\
p(x, z) \propto A_z \\
p(\overrightarrow{\rho}, \rho) = 1
\end{array} \right. \\
\text{where } \overrightarrow{z} = x \in V
\end{align*}
\]

(7)

This modification will not change the recurrence/transience behavior of the RWRE \( \eta \) nor its speed in the transient regime. We will always work with this modification in the sequel.

Let us now introduce the notation of quenched and annealed probabilities. Given the environment \((\omega, T)\), let \( P_{\omega,T}^x \) denote the quenched probability of the random walk \( \eta \) with \( \eta_0 = x \in V \) a.s. Denote by \( \mathbb{P}_x^T, \mathbb{Q}, \mathbb{P}_\rho \) the measures:

\[
\mathbb{P}_x^T(\cdot) := \int P_{\omega,T}^x(\cdot) \mathbb{P}(d\omega),
\]

\[
\mathbb{Q}(\cdot) := \int 1_{\{\cdot\}} \mathbb{P}(d\omega) GW(dT)
\]

\[
\mathbb{P}_\rho(\cdot) := \int \mathbb{P}_\rho^T(\cdot) GW(dT),
\]

and the associated expectations are denoted by \( E_{\mathbb{P}_x^T}, E_{\mathbb{Q}}^T, E_{\mathbb{P}_\rho} \) and \( \mathbb{E} \). For brevity, we omit the starting point if the random walk starts from the root; that is, we write \( P^{\omega,T}, \mathbb{P}^T \) and \( \mathbb{P} \) for \( P_{\omega,T}^x, \mathbb{P}_x^T \) and \( \mathbb{P}_\rho \). Notice that \( \mathbb{P} \) is the annealed law of \( \eta \).

For any vertex \( x \), let \( |x| = d(\rho, x) \) be the generation of \( x \) and denote by \( [\rho, x] \) the unique shortest path from \( x \) to the root \( \rho \), and \( x_i \) (for \( 0 \leq i \leq |x| \)) the vertices on \( [\rho, x] \) such that \( |x_i| = i \). In particular, \( x_0 = \rho \) and \( x_0 = x \). In words, \( x_i \) (for \( i < |x| \)) is the ancestor of \( x \) at generation \( i \). Also denote \( \llbracket \rho, x \rrbracket := \{ \rho \} \cup [\rho, x] \) and \( \llbracket \rho, x \rrbracket := \{ \rho \} \cup [\rho, x] \). Moreover, for \( u, v \in T \), we write \( u < v \) if \( u \) is an ancestor of \( v \).

### 3 Phase transition: an alternative proof of Theorem 1

The ideas follow from Lyons and Pemantle [13], by means of random electrical network.

**Proof of Theorem 1** The RWRE is equivalent to an electrical network with random conductances:

\[
C_{x} := C(x, \overrightarrow{x}) = (\prod_{u \in \llbracket \rho, x \rrbracket} A_u)^2 A_x, \forall x \in V \setminus \{ \rho \}.
\]

We omit the proof of the transient case which is quite similar to that in Lyons and Pemantle [13], however, we will detail the recurrence case. That is, we will show that if \( b\mu(c) \leq 1 \), then the RWRE is recurrent a.s.
First consider the case $b\mu(c) < 1$, note that
\[
E_Q \left[ \sum_{n \geq 1} \sum_{|x|=n} C_{e_x}^{1/4} \right] = \sum_{n \geq 1} \int \left( \int \sum_{|x|=n} C_{e_x}^{1/4} \mathbf{P}(d\omega) \right) GW(dT)
\]
\[
= \sum_{n \geq 1} \int \sum_{|x|=n} E[A^{1/2}]^{n-1} E[A^{1/4}] GW(dT)
\]
\[
= \sum_{n \geq 1} b^n E[A^{1/2}]^{n-1} E[A^{1/4}].
\]
Because $\mu(c) = E[A^{1/2}] < 1/b$, we have, for some constants $c_1, c_2 \in \mathbb{R}^+$
\[
E_Q \left[ \sum_{n \geq 1} \sum_{|x|=n} C_{e_x}^{1/4} \right] \leq c_1 \sum_{n \geq 0} (b\mu(c))^n \leq c_2 < \infty,
\]
which implies that
\[
\sum_{n \geq 1} \sum_{|x|=n} C_{e_x}^{1/4} < \infty, \quad \text{Q-a.s.}
\]
As a result, there exists a stationary probability a.s., moreover $\eta$ is positive recurrent.

Turning to the case $b\mu(c) = 1$, let $\Pi_n := \{e_x : |x| = n\}$ be a sequence of cutsets. Observe that
\[
W_n := \sum_{|x|=n} \prod_{u \in |\rho, x|} A_{1/2}^{1/2} = \sum_{|x|=n} C_{e_x}^{1/4} A_x^{1/4}.
\]
is a martingale with respect to $\mathcal{F}_n = \sigma(\{A_x, |x| \leq n\})$. By Biggin’s theorem ([4 12]), it converges a.s. to zero. More precisely, we use the equivalent condition (iv) in page 2 of [12], since $t \mapsto E(A_t)$ attain its minimum at $t = \frac{1}{2}$, in terms of $m(\alpha), m'(\alpha)$ in [12], we actually have $\alpha = -\frac{1}{2}$ and $m(\alpha) = 1, m'(\alpha) = 0$, therefore, the criterion (iv) is never satisfied.

We are going to show that Q-a.s.,
\[
\liminf_{n \to \infty} \sum_{|x|=n} C_{e_x}^{1/4} = 0,
\]
in particular, this will imply that Q-a.s. $\inf_{w \in \Pi, \text{cutset}} \sum_{e_x \in \Pi} C_{e_x} = 0$. By the trivial half of the max-flow min-cut theorem, the corresponding network admits no flow a.s. Hence, the random walk is a.s. recurrent. Observes that
\[
\sum_{|x|=n} C_{e_x}^{1/4} = \sum_{|x|=n} \prod_{u \in |\rho, x|} A_{1/2}^{1/2} A_x^{1/4} (A_x \geq 1) + \sum_{|x|=n} \prod_{u \in |\rho, x|} A_{1/2}^{1/2} A_x^{1/4} (A_x < 1)
\]
\[
= \sum_{|x|=n} \prod_{u \in |\rho, x|} A_{1/2}^{1/2} A_x^{1/4} (A_x \geq 1) + \sum_{|y|=n-1} \prod_{u \in |\rho, y|} A_{1/2}^{1/2} A_x^{1/4} (A_x < 1)
\]
\[
\leq W_n + \sum_{|y|=n-1} \prod_{u \in |\rho, y|} A_{1/2}^{1/2} \nu_y,
\]
where $\nu_y$ denotes the number of children of $y$. Letting $n$ go to infinity yields that
\[
0 \leq \liminf_{n \to \infty} \sum_{|x|=n} C_{e_x}^{1/4} \leq \liminf_{n \to \infty} \sum_{|y|=n-1} \prod_{u \in |\rho, y|} A_{1/2}^{1/2} \nu_y.
\]
For any $K \geq 1$, separating the sum over vertices $y$ according to $\{\nu_y < K\}$ or $\{\nu_y \geq K\}$, the last term is bounded by
\[
\lim_{n \to \infty} KW_{n-1} + \liminf_{n \to \infty} \sum_{|y|=n-1} \prod_{u \in \rho, y} A_u^{1/2} \nu_y 1_{\{\nu_y \geq K\}} = \liminf_{n \to \infty} \sum_{|y|=n-1} \prod_{u \in \rho, y} A_u^{1/2} \nu_y 1_{\{\nu_y \geq K\}}.
\]

By Fatou’s lemma,
\[
\mathbb{E}_Q \left( \liminf_{n \to \infty} \sum_{|y|=n-1} \prod_{u \in \rho, y} A_u^{1/2} \nu_y 1_{\{\nu_y \geq K\}} \right) 
\leq \liminf_{n \to \infty} \mathbb{E}_Q \left( \sum_{|y|=n-1} \prod_{u \in \rho, y} A_u^{1/2} \nu_y 1_{\{\nu_y \geq K\}} \right) = \mathbb{E}_Q[\nu_{\rho}, \nu_{\rho} \geq K],
\]

since for all $|y| = n - 1$, $\nu_y$ is independent of $\prod_{u \in \rho, y} A_u^{1/2}$ and $\mathbb{E}_Q \left( \sum_{|y|=n-1} \prod_{u \in \rho, y} A_u^{1/2} \right) = 1$.

Consequently, for any $K \geq 1$,
\[
\mathbb{E}_Q \left[ \liminf_{n \to \infty} \sum_{|x|=n} C_{e_x}^{1/4} \right] \leq \mathbb{E}_Q[\nu_{\rho}, \nu_{\rho} \geq K].
\]

As $b = \mathbb{E}_Q[\nu_y] < \infty$, letting $K \to \infty$ gives
\[
\mathbb{E}_Q \left[ \liminf_{n \to \infty} \sum_{|x|=n} C_{e_x}^{1/4} \right] = 0.
\]

This implies (8). \qed

4 Speed when transient

Turning to the positivity of $v(Z)$ and $v(\eta)$, note that the processes $(Z_t)$ and $(\eta_n)$ are mixture of Markov processes but $(Y_t)$ is not, in fact, $(Y_t)$ escapes faster than $(Z_t)$, in particular, when $v(Z) > 0$, we have $v(Y) > 0$. But we are not sure whether $v(Z) = 0$ implies $v(Y) = 0$.

4.1 Regeneration structure

In this section, we show that, when the process $(\eta_n)$ (or $(Z_t)$) is transient, its path can be cut into independent pieces, using the notion of regeneration time. As a consequence, the speed $v(\eta)$, $v(Z)$ exists a.s. as a limit (not just a lim inf).

On a tree, when a random walk traverses an edge for the first and last time simultaneously, we say it regenerates since it will now remain in a previously unexplored sub-tree. For any vertex $x$, let $D(x) = \inf\{k \geq 1, \; \eta_{k-1} = x, \eta_k = \overset{\leftarrow}{x}\}$, write $\tau_n = \inf\{k \geq 0, \; |\eta_k| = n\}$ and define the regeneration time recursively by
\[
\begin{cases}
\Gamma_0 = 0 \\
\Gamma_n = \Gamma(n) = \inf\{k > \Gamma_{n-1}; \; d(\eta_k) \geq 3, D(\eta_k) = \infty, \tau_{|\eta_k|} = k\}
\end{cases}
\]

where $d(x)$ is the degree of the vertex $x$. 7
Lemma 1. Let $S(\cdot) = \mathbb{P}(\cdot | d(\rho) \geq 3, \mathcal{D}(\rho) = \infty)$, if $\eta$ is transient, then

i) For any $n \geq 1$, $\Gamma_n < \infty$ \textit{P}.a.s.

ii) Under $\mathbb{P}$, $(\Gamma_{n+1} - \Gamma_n, |\eta_{\Gamma_{n+1}}| - |\eta_n|, A_{\Gamma_{n+1}})_{n \geq 1}$ are independent and distributed as $(\Gamma_1, |\eta_1|, A_1)$ under $S$.

iii) $E_{S}(|\eta_1|) < \infty$.

We feel free to omit the proof because it is analogue to ‘Fact’ in [1] p.10. In addition, Lemma 1 also holds without assuming $d(\eta_k) \geq 3$ in the definition of $\Gamma_n$, but we will need this assumption later in the proof of Lemma 7.

By strong law of large numbers, one immediately sees that there exist two constants $c_4 \geq c_3 \geq 1$ such that $\mathbb{P}$.a.s.,

$$\lim_{n \to \infty} \frac{|\eta_n|}{n} = c_3 \in [1, \infty), \quad \lim_{n \to \infty} \frac{\Gamma_n}{n} = c_4 \in [c_3, \infty].$$

In addition, for any $n \geq 1$, there exists a unique $u(n) \in \mathbb{N}$ such that

$$\Gamma_{u(n)} \leq n < \Gamma_{u(n)+1}$$

and $|\eta_{u(n)}| \leq |\eta_n| < |\eta_{u(n)+1}|$. Letting $n$ go to infinity, (in particular $u(n) \to \infty$) in

$$\frac{|\eta_n|}{n} < \frac{|\eta_{u(n)+1}|}{\Gamma_{u(n)+1}} \leq \frac{|\eta_{u(n)}|}{\Gamma_{u(n)}} = \frac{|\eta_{u(n)+1}|}{u(n)} \frac{u(n)}{\Gamma_{u(n)}}.$$

We have $\mathbb{P}$.a.s.

$$\frac{|\eta_n|}{n} \to v(\eta) := \frac{c_3}{c_4} \in [0, 1].$$

For $Z_t$, the same arguments can be applied. As a consequence of the i.i.d. decomposition, $v(Z) = \lim_{t \to \infty} \frac{|Z_t|}{t}$ exists a.s. The existence of $v(Y) = \lim_{n \to \infty} \frac{|Y|}{n}$ can be justified by performing the time change $D(t)$ between consecutive regenerative epochs.

4.2 The auxiliary one dimensional process

The RWRE can also be defined on the deterministic graph $\mathbb{H} = \{-1, 0, 1, \ldots\}$, on which many quantities are viable by explicit computations. The strategy is to compare the random walk on a tree to the random walk on the half line, in the forthcoming sections we will explain how these comparisons will be done. In this section we list some properties of the one dimensional random walk, their proofs can be found in Appendix A.

Let $\tilde{\eta}_n$ be the random walk on the half line $\mathbb{H} = \{-1, 0, 1, \ldots\}$ in the random environment $\omega = (A_k, k \geq 0)$ which are i.i.d. copies of $A$ under $\mathbb{P}$, with transition probability according to (7); that is,

$$p(i, i+1) = \frac{A_{i+1}}{1/A_i + A_{i+1}} \quad i \geq 0$$

$$p(i, i-1) = \frac{1/A_i}{1/A_i + A_{i+1}} \quad i \geq 0$$

$$p(-1, 0) = 1$$

Similarly we denote $\tilde{P}_i^\omega, \tilde{\mathbb{P}}_i, \tilde{E}_i^\omega, \tilde{E}_i$ respectively the quenched and annealed probability/expectation for such process starting from $i$, and for any $n \in \mathbb{H}$, define the following stopping times

$$\tilde{\tau}_n = \inf \{ k \geq 0, \tilde{\eta}_k = n \}, \quad \tilde{\tau}_n^* = \inf \{ k \geq 1, \tilde{\eta}_k = n \}.$$
Let $F_1, F_2 > 0$ be two expressions which can depend on any variable, but in particular on $n$. If there exists $f : \mathbb{N} \to \mathbb{R}^+$ with $\lim_{n \to \infty} \frac{1}{n} \log f(n) = 0$ such that $F_1 f(n) \geq F_2$, then we denote $F_1 \gtrsim_n F_2$ ($F_1$ greater than $F_2$ up to polynomial constant).

Recall that $A$ is Inverse Gaussian distributed with parameter $(1, \sigma^2)$, define the rate function associated to $\log A$ by

$$I(x) = \sup_{t \in \mathbb{R}} \{tx - \log E(A^t)\},$$

also define

$$t^* = \sup\{t \in \mathbb{R}, E(A^t)q_1 \leq 1\}.$$  

Here are the list of estimates in dimension one.

**Lemma 2.** For any $z > 0$ and $0 < z_1 < 1$, we have, for any $0 < a < 1$

$$\tilde{P}_0(\tilde{t}_n \wedge \tilde{t}_- > m|A_0 \in [a, \frac{1}{a}]) \gtrsim_n \exp\left\{-n \left(z_1 I(\frac{z}{2z_1}) + (1 - z_1)I(\frac{-z}{2(1 - z_1)})\right)\right\}$$

where $m \in \mathbb{N}$ is such that $n = \left\lfloor \frac{\log m}{z}\right\rfloor$.

**Lemma 3.** Denote

$$L' = \sup_{z > 0, \; 0 < z_1 < 1} \left\{\log q_1 z - z_1 I\left(\frac{z}{2z_1}\right) - \frac{1 - z_1}{z} I\left(\frac{-z}{2(1 - z_1)}\right)\right\},$$

we have $L' = t^* + \frac{1}{2}$.

**Lemma 4.** Define, for $i \in \mathbb{H}$ and any stopping time $\tau$, $\tilde{G}^\tau(i, i) = \tilde{E}^\omega_i(\sum_{k=0}^{\tau} I_{\tilde{y}_k = i})$. Let $0 \leq Y_1 < Y_2 < y < Y_3$ be points on the half line, we have, for any $0 \leq \lambda \leq 1$,

$$\tilde{P}_{Y_1}^{\omega}(\tilde{\tau}_u < \tilde{\tau}_{Y_1} - 1)\tilde{G}^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}}(y, y) \leq \tilde{E}_{Y_1}^{\omega}\tilde{E}_{Y_3}^{\omega}\tilde{E}_{Y_3}^{\omega} \leq S_{\lambda, [Y_1, Y_2]} \left(1 + A_{Y_{2+1}}^\lambda \left(1 + \tilde{E}_{Y_{2+1}}^{\omega}[\tilde{\tau}_{Y_2} \wedge \tilde{\tau}_{Y_3}]^\lambda\right)\right).$$

where

$$S_{\lambda, [Y_1, Y_2]} := 1 + 2A_{Y_1}^\lambda \sum_{Y_1 < u \leq Y_2} \prod_{Y_1 < u < z} A_u^2 A_z^2 + A_{Y_1}^\lambda \prod_{Y_1 < u \leq Y_2} A_u^{2\lambda}.$$ 

**Lemma 5.** If $0 \leq \lambda < (t^* - \frac{1}{2}) \wedge 1$, then there exists sufficiently small $\delta > 0$ such that for all $n_1 > 0$

$$E\left(\left(1 + \frac{1}{A_{n_1}^\lambda}\right)\left(1 + \frac{1}{A_n}A_0\tilde{E}_0^{\omega}[\tilde{\tau}_- \wedge \tilde{\tau}_n]^\lambda\right)\right) \lesssim_n (q_1 + \delta)^{-n}.$$

### 4.3 Null speed regime

In this section we prove a part of (2) (when the speed is zero) in Theorem 2.

**Proposition 1.** Recall the definition of $t^*$ in \cite{10}, if $q_1 E(A^{-1/2}) > 1$, then $1 < t^* < \frac{3}{2}$ and

$$\limsup_n \frac{\log |\eta_n|}{\log n} \leq t^* - \frac{1}{2}.$$

In particular, if $q_1 E(A^{-1/2}) > 1$, then $\mathbb{P}$-a.s., $v(\eta) = 0$; in fact,

$$|\eta_n| = n^{(t^* - 1/2) + o(1)} = o(n), \; n \to \infty.$$
Remarks 4. Similar arguments can be carried out for the continuous time process \((Z_t)\), i.e. if \(q_1 \mathbb{E}(A^{-1/2}) > 1\), then
\[
\limsup_t \frac{\log |Z_t|}{\log t} \leq t^* - \frac{1}{2}.
\] (13)

Let us state an estimate on the tail distribution of the regeneration time \(\Gamma_1\) under \(S(\cdot)\):

Lemma 6. For \(N\) large enough, in the case \(1 < t^* < 3/2\), for any \(\lambda \in (t^* - \frac{1}{2}, 1)\), there is \(\epsilon > 0\) such that
\[
S(\Gamma_1 > N^{1/\lambda}) \geq N^{-1+\epsilon}.
\] (14)

With the help of the above lemma, we prove Proposition 1.

Proof of Proposition 1. Note that \(t \mapsto \mathbb{E}(A^t)\) is a convex function, and it is symmetric w.r.t. the line \(t = \frac{1}{2}\), where it takes the minimum, in particular \(\mathbb{E}(A^{-1/2}) = \mathbb{E}(A^{3/2})\). As we have assumed that \(q_1 \mathbb{E}(A^{-1/2}) > 1\), it follows that \(t^* < \frac{3}{2}\). On the other hand, since \(\mathbb{E}(A) = 1\), obviously \(t^* > 1\). For any \(\lambda \in (t^* - 1/2, 1)\), by Lemma 6 there exists \(\epsilon > 0\) such that
\[
\mathbb{P}(\max_{2 \leq k \leq n} (\Gamma_k - \Gamma_{k-1}) \leq n^{1/\lambda}) = S(\Gamma_1 \leq n^{1/\lambda})n^{-1} 
\leq (1 - n^{-1+\epsilon})n^{-1} \lesssim n \exp(-n^\epsilon).
\]
Therefore,
\[
\sum_{n \geq 2} \mathbb{P}(\max_{2 \leq k \leq n} (\Gamma_k - \Gamma_{k-1}) \leq n^{1/\lambda}) < \infty.
\]
By Borel-Cantelli lemma, \(\mathbb{P}\text{-a.s.}\), for all \(n\) large enough,
\[
\Gamma_n \geq \max_{2 \leq k \leq n} (\Gamma_k - \Gamma_{k-1}) \geq n^{1/\lambda}.
\]
It follows that \(\mathbb{P}\text{-a.s.}, \liminf_n \frac{\log \Gamma_n}{\log n} \geq \frac{1}{\lambda}\). As \(\liminf_n \frac{\log \tau_n}{\log n} \geq \liminf_n \frac{\log \Gamma_n}{\log n}\) (see (3.1) in [1]), we have
\[
\limsup_n \frac{\log \tau_n}{\log n} \leq \lambda \frac{1}{t^* - \frac{1}{2}} < 1, \mathbb{P}\text{-a.s.}
\]
\[
\square
\]
It remains to prove Lemma 6. In fact, when \( q_1 \) is large, it is more likely that there will be some long branch constituting vertices of degree two on the GW tree, especially starting from the root. These branches will slow down the process and entail zero velocity. The following lemma gives a comparison between the tail distribution of the regeneration time \( \Gamma_1 \) and the probability that the process wanders on these branches (which is a one dimensional random walk in random environment, that is, \( (\hat{\eta}_n) \)).

**Lemma 7.** For any \( m \geq 1 \), \( 0 < a < 1 \), we have

\[
S(\Gamma_1 > m) \geq c_5 \sum_{n=1}^{\infty} q^n \hat{P}_0(\hat{\tau}_n = m) \geq c_5 \sum_{n=1}^{\infty} q^n \hat{P}_0(\hat{\tau}_n > m) | A_0 \in [a, \frac{1}{a}]).
\]

Now we prove Lemma 6 with the help of Lemma 7 and some results on the one dimensional RWRE.

**Proof of Lemma 6.** By Lemma 2, one sees that for all \( z > 0 \), \( 0 < z_1 < 1 \) and \( m \) such that \( n = \lfloor \frac{\log m}{z} \rfloor \),

\[
\hat{P}_0(\hat{\tau}_n \wedge \hat{\tau}_1 > m | A_0 \in [a, \frac{1}{a}]) \geq \exp(-n \left( z_1 I\left( \frac{z}{2z_1} \right) + (1 - z_1) I\left( \frac{-z}{2(1 - z_1)} \right) \right))
\]

where we recall that \( I(x) = \sup_{t \in \mathbb{R}} \{tx - \log \mathbb{E}(A^t)\} \). For large \( m \), by Lemma 7 for all \( m \geq 1 \), \( 0 < a < 1 \)

\[
S(\Gamma_1 > m) \geq c_5 \max_{n,\lfloor \frac{\log m}{z} \rfloor} q^n \hat{P}_0(\hat{\tau}_n \wedge \hat{\tau}_1 > m) | A_0 \in [a, \frac{1}{a}])
\]

then by Lemma 2 for all \( m \geq 1 \), \( 0 < a < 1 \), \( z > 0 \), \( 0 < z_1 < 1 \)

\[
S(\Gamma_1 > m) \geq \exp\left(-\log m \frac{I\left( \frac{z}{2z_1} \right) + (1 - z_1) I\left( \frac{-z}{2(1 - z_1)} \right) - \log q_1}{z} \right)
\]

That is, for all \( z > 0 \), \( 0 < z_1 < 1 \)

\[
S(\Gamma_1 > e^{nz}) \geq \exp\left(nz \left( \frac{\log q_1}{z} - \frac{z_1}{z} I\left( \frac{z}{2z_1} \right) - \frac{1 - z_1}{z} I\left( \frac{-z}{2(1 - z_1)} \right) \right) \right)
\]

It follows from the proof of Lemma 3 that we can take \( z_1 = \frac{1}{2}, z = (t \mapsto \log \mathbb{E}(A^t))'(t^*) > 0 \)

\[
S(\Gamma_1 > e^{nz}) \geq e^{nz(-t^* + 1/2)}.
\]

Recall that \( 1 < t^* < \frac{3}{2} \), thus for any \( \lambda \in (t^* - \frac{1}{2}, 1) \), there exists a \( \epsilon > 0 \) such that (denote \( N = \lfloor e^{nz} \rfloor \))

\[
S(\Gamma_1 > N^{1/\lambda}) \geq N^{-1+\epsilon}.
\]

\( \square \)

It remains to prove the comparison Lemma 7. We define, for \( x \neq \frac{1}{\rho} \),

\[
\tau_x = \inf\{ n \geq 0; \, \eta_n = x \}, \, \tau_x^* = \inf\{ n > 0; \, \eta_n = x \}, \, \beta(x) = P_x^w(T_\infty = \infty)
\]

11
Note that for any \( x \in T \), \( \beta(x) \) depends only on the sub-tree \( T_x \) rooted at \( x \) and the environment \( \{ A_y(\omega); y \in T_x \} \), let us denote \( \beta \) a generic r.v. distributed as \( \beta(\rho) \), by transient assumption, \( \beta > 0 \) a.s. and \( \mathbb{E}(\beta) > 0 \).

Moreover, by Markov property,
\[
\beta(x) = \sum_{y: y = x} p(x, y)[P^\omega_{\rho,T}(\tau_x = \infty) + P^\omega_{\rho,T}(\tau_x < \infty)\beta(x)] \\
= \sum_{y: y = x} p(x, y)[\beta(y) + (1 - \beta(y))\beta(x)].
\]

Note that \( \beta(x) > 0 \), \( \mathbb{P} \)-a.s. hence,
\[
\frac{1}{\beta(x)} = 1 + \frac{1}{A_x \sum_y A_y \beta(y)}.
\] (15)

In particular, \( \beta(x) \) is increasing as a function of \( A_x \).

Proof of Lemma 7. For any vertex \( x \), let \( h(x) \) be the first descendant of \( x \) such that \( d(h(x)) \geq 3 \). Let \( k_0 = \inf\{k \geq 2 : q_k > 0\} \). According to the definition of \( \Gamma_1 \), one observes that when \( \eta_i \neq \rho \),
\[
\Gamma_1 \geq \tau^*_\rho \land \tau_{h(\eta_i)}.
\]

In fact, we are going to consider the following events
\[
E_0 = \{d(\rho) = k_0 + 1, A_\rho \geq a, A_{\rho_i} \in [a, \frac{1}{a}], \forall 1 \leq i \leq k_0\} \text{ where } \rho_i \text{ are children of } \rho,
\]
\[
E_1 = E_0 \cap \{\eta_i \neq \rho, m < \tau^*_\rho < \tau_{h(\eta_i)} , \eta_{\tau^*_\rho + 1} \neq \{\rho, \eta_i\} \} \cap \{\eta_n \neq \rho; \forall n \geq \tau^*_\rho + 1\},
\]
\[
E_2 = E_0 \cap \{\eta_i \neq \rho, m < \tau_{h(\eta_i)} < \tau^*_\rho \} \cap \{\eta_n \neq \rho; \forall n \geq \tau_{h(\eta_i)} + 1\}.
\]

As \( \Gamma_1 \geq \tau^*_\rho \land \tau_{h(\eta_1)} \), we have \( E_1 \cup E_2 \subset E_0 \cap \{D(\rho) = \infty, \Gamma_1 > m\} \) and \( E_1 \cap E_2 = \emptyset \). So,
\[
\mathbb{P}(E_0 \cap \{D(\rho) = \infty, \Gamma_1 > m\}) \geq \mathbb{P}(E_1) + \mathbb{P}(E_2).
\]

For \( E_1 \), by strong Markov property at \( \tau^*_\rho \) and weak Markov property at time 1,
\[
P^\omega_{\rho,T}(E_1) = \mathbf{1}_{E_0} P^\omega_{\rho,T}(\{\eta_1 \neq \rho, m < \tau^*_\rho < \tau_{h(\eta_1)} , \eta_{\tau^*_\rho + 1} \neq \{\rho, \eta_i\} \} \cap \{\eta_n \neq \rho; \forall n \geq \tau^*_\rho + 1\})
\]
\[
= \mathbf{1}_{E_0} \sum_{i=1}^{k_0} p(\rho, \rho_i) P^\omega_{\rho_i,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} p(\rho, \rho_j) \beta(\rho_j).
\]

Given \( E_0 \), \( p(\rho, \rho_i) \geq \frac{a^2}{k_0 + 1} =: c_6 \) for any \( 1 \leq i \leq k_0 \). So,
\[
P^\omega_{\rho,T}(E_1) \geq \mathbf{1}_{E_0} \sum_{i=1}^{k_0} c_6 P^\omega_{\rho_i,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \sum_{j \neq i} c_6 \beta(\rho_j)
\]
\[
= c_6^2 \mathbf{1}_{E_0} \sum_{i=1}^{k_0} \sum_{j \neq i} P^\omega_{\rho_i,T}(m - 1 < \tau_\rho < \tau_{h(\rho_i)}) \beta(\rho_j),
\]
where $P_{\rho_i}^{\omega, T}(m - 1 < \tau_\rho < h(\rho))$ is a function of $A_y$, $\rho_i \leq y \leq h(\rho_i)$ which is independent of $\beta(\rho_j) = \frac{A_{\rho_j}}{1 + \sum_{z \neq \rho_j} A_{\rho, \beta(z)}}$ for $i \neq j$. Recall that given $E_0$, we have $P(A_\rho \in [a, a^{-1}]) := c_7 > 0$.

Now for any couple $(i, j)$, one has
\[
\mathbb{E}[\mathbb{I}_{E_0} P_{\rho_i}^{\omega, T}(m - 1 < \tau_\rho < h(\rho_i)) \beta(\rho_j) ] \\
\geq c_7^{k_0 - 1} q_{\rho_j} \mathbb{E}[\mathbb{I}_{A_{\rho_j} \in [a, a^{-1}]} P_{\rho_i}^{\omega, T}(m - 1 < \tau_\rho < h(\rho_i)) \mathbb{E}[\beta(\rho_j) \mathbb{I}_{A_{\rho_j} \in [a, a^{-1}]}] \\
\geq c_7^{k_0 - 1} \mathbb{E}[\mathbb{I}_{E_0} P_{\rho_i}^{\omega, T}(m - 1 < \tau_\rho < h(\rho_i)) \mathbb{E}[\beta(\rho) \mathbb{I}_{A_{\rho} \in [a, a^{-1}]}].
\]

Note that $\mathbb{E}[\beta(\rho)] > 0$. So we could choose $a < 1$ such that $\mathbb{E}[\beta(\rho) \mathbb{I}_{A_{\rho} \in [a, a^{-1}]}) > 0$. Taking sum over $i, j$ shows that
\[
P(\mathbb{P}(E_1)) \geq \frac{c_2}{c_7} (k_0 - 1) \mathbb{E}\left( \mathbb{I}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega, T}(m - 1 < \tau_\rho < h(\rho_i)) \right). \tag{16}
\]

Similarly for $E_2$, by Markov property,
\[
P_{\rho_i}^{\omega, T}(E_2) = \mathbb{I}_{E_0} P_{\rho_i}^{\omega, T}(\{\eta_l \neq \rho, m < h(\eta_l) < \tau_\rho \} \cap \{\eta_l \neq \hat{h}(\eta_l) ; \forall n \geq h(\eta_l) + 1\}) \\
= \mathbb{I}_{E_0} \sum_{i=1}^{k_0} p(\rho, \rho_i) P_{\rho_i}^{\omega, T}(m - 1 < h(\rho_i) < \tau_\rho) \beta(h(\rho_i)) \\
\geq c_6 \mathbb{E}\left[ E_0 \sum_{i=1}^{k_0} P_{\rho_i}^{\omega, T}(m - 1 < h(\rho_i) < \tau_\rho) \beta(h(\rho_i)) \right].
\]

To get rid of dependence between $P_{\rho_i}^{\omega, T}(m - 1 < h(\rho_i) < \tau_\rho)$ and $\beta(h(\rho_i))$, we note that
\[
P_{\rho_i}^{\omega, T}(m - 1 < h(\rho_i) < \tau_\rho) \geq P_{\rho_i}^{\omega, T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho, \eta_{\tau_{h(\rho_i)} + 1} = h(\rho_i))
\]

which by Markov property is $P_{\rho_i}^{\omega, T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho) \frac{A_{h(\rho_i)}}{A_{h(\rho_i)} + A(h(\rho_i))}$. This term and $\beta(h(\rho_i))$ are both increasing on $A_{h(\rho_i)}$. FKG inequality conditionally on $\{T; A_u, u \neq h(\rho_i)\}$ entails
\[
\mathbb{P}(E_2) \geq c_6 \mathbb{E}\left[ \mathbb{I}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega, T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho) \frac{A_{h(\rho_i)}}{A_{h(\rho_i)} + A(h(\rho_i))} \right] \times \mathbb{E}(\beta(\rho)) \\
= c_6 \mathbb{E}[\mathbb{I}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega, T}(m - 1 < \tau_{h(\rho_i)} < \tau_\rho) \mathbb{I}_{A_{h(\rho_i)} \geq a}] \mathbb{E}\left[ \frac{a A_{\rho}}{1 + a A_{\rho}} \right], \tag{17}
\]

with $c_8 := c_6 \mathbb{E}(\beta(\rho)) > 0$. Clearly, $\mathbb{E}[\frac{a A_{\rho}}{1 + a A_{\rho}}] > 0$ for any $a > 0$. Combining (16) with (17) yields that
\[
\mathbb{P}(E_1) + \mathbb{P}(E_2) \geq c_9 \mathbb{E}\left[ \mathbb{I}_{E_0} \sum_{i=1}^{k_0} P_{\rho_i}^{\omega, T}(\tau_\rho \wedge \tau_{h(\rho_i)} > m - 1) \mathbb{I}_{A_{h(\rho_i)} \geq a} \right] \\
\geq c_9 k_0 Q(\mathbb{E}) \mathbb{P}(\tau_\rho \wedge \tau_{h(\rho)} > m - 1, A_{h(\rho)} \geq a | A_\rho \in [a, \frac{1}{a}]) \\
\geq c_{10} \mathbb{P}(\tau_\rho \wedge \tau_{h(\rho)} > m - 1; |h(\rho)| \geq 2 | A_\rho \in [a, \frac{1}{a}]). \tag{18}
\]
where \( h(\rho) \) is the grand parent of \( h(\rho) \). Let us go back to \( S(\Gamma_1 > m) \). As \( \mathbb{P}(d(\rho) \geq 3, D(\rho) = \infty) > 0 \), recall that

\[
S(\Gamma_1 > m) = \mathbb{P}(\Gamma_1 > m | d(\rho) \geq 3, D(\rho) = \infty) \\
\geq \mathbb{P}(E_0 \cap \{ D(\rho) = \infty, \Gamma_1 > m \}) \\
\geq \mathbb{P}(E_1) + \mathbb{P}(E_2).
\]

By (18), taking \( c_5 = c_{10} \), we have

\[
S(\Gamma_1 > m) \geq c_5 \mathbb{P}(\tau_{\rho} \wedge \tau_{h(\rho)} > m - 1, |h(\rho)| \geq 2|A_\rho| \in [a, \frac{1}{a}]) \\
= c_5 q_1^2 \sum_{n=1}^{\infty} q_1^n \mathbb{P}_0(\hat{\tau}_{-1} \wedge \hat{\tau}_n > m - 1 | A_0 \in [a, \frac{1}{a}]).
\]

\[
\square
\]

### 4.4 Positive speed on big tree and asymptotic of \(|Z_t|\) on small tree

This subsection is devoted to the proof of the following propositions, firstly when the tree is big (i.e. \( q_1 \) small), the RWRE has positive speed; when the tree is small (\( q_1 \) large), we can compute exactly the asymptotic behavior of \(|\eta_n|\) and \(|Z_t|\).

**Proposition 2.** If \( q_1 \mathbb{E}(A^{-1/2}) < 1 \), then

\[
v(\eta) > 0 \text{ and } v(Z) > 0.
\]

As a consequence, also \( v(Y) > 0 \).

**Proposition 3.** Assume that \( q_1 \mathbb{E}(A^{-1/2}) > 1 \), we have \( \mathbb{P}\text{-a.s.} \)

\[
\lim_{n \to \infty} \frac{\log |\eta_n|}{\log n} = \lim_{t \to \infty} \frac{\log |Z_t|}{\log t} = t^* - 1/2 \in (1/2, 1)
\]

where \( t^* = \sup \{ t \in \mathbb{R}, \mathbb{E}(A^t) q_1 \leq 1 \} \).

Let us give some definitions and heuristics before proving these propositions, write, for \( n \geq 0 \),

\[
\tau_n(\eta) = \inf \{ k \geq 0; |\eta_k| = n \} \text{ and } \tau_n(Z) = \inf \{ t \geq 0; |Z_t| = n \}
\]

the hitting times of the \( n \)-th generation for \( \eta \) and \( Z \) respectively. As a consequence of the law of large numbers, \( \mathbb{P}\text{-a.s.} \),

\[
\lim_{n \to \infty} \frac{\tau_n(\eta)}{n} = \frac{1}{v(\eta)} \text{ and } \lim_{n \to \infty} \frac{\tau_n(Z)}{n} = \frac{1}{v(Z)}.
\]

The study of the speed is reduced to the study of \( \tau_n(\eta) \) and \( \tau_n(Z) \). For any \( x \in T, n \geq -1 \), let \( N_x \) and \( N_n \) denote the time spent by the walk \( \eta \) at \( x \) and at the \( n \)-th generation respectively:

\[
N(x) = \sum_{k \geq 0} \mathbb{1}_{\eta_k = x}, \quad N_n = \sum_{|x| = n} N(x),
\]

observe that

\[
\tau_n(\eta) \leq \sum_{k=-1}^{n} N_k, \quad E^{x,T}[\tau_n(Z)|\eta] \leq \sum_{x:-1 \leq |x| \leq n} N_x \frac{A_x}{1 + A_x B_x},
\]

where \( A_x \) and \( B_x \) are the adjacency and connecting probabilities of the tree at \( x \).
where \( B_x := \sum_{y : y = x} A_y \).

In what follows, we actually study \( N_n \) for large \( n \) to show that \( \lim \inf n^{-1} \sum_{k=1}^{n} N_k < \infty \), \( \mathbb{P} \)-a.s. The heuristics is the following. Fix some \( n_0, K_0 \) (to choose later), pick some vertex \( y \) at the \( n \)-th generation, if \( y \) roughly lies in a subtree of height \( n_0 \) with more than \( K_0 \) leaves, then the random walk will immediately go down, thus \( \mathbb{E}(N_y) \) will be small c.f. Figure 2 left. Otherwise, we seek a down going path \( \hat{y}, \ldots, y, \ldots, \tilde{y} \) such that every vertex in this path does not branch much except for the two ends, and we need these two ends have more than \( K_0 \) descendants after \( n_0 \) generations. In such configuration, we can compare the random walk to the one dimensional one, and once the walker reaches one of the ends, it immediately leaves our path \( \hat{y}, \ldots, \tilde{y} \) c.f. Figure 2 right.

![Figure 2: Two cases to bound \( \mathbb{E}(N_y) \).](image)

If the root has more than \( K_0 \) descendants after \( n_0 \) generations, then we can always find \( \hat{y} \). Otherwise, we need to take \( n \) large and use the Galton Watson structure. To handle this issue, let us introduce the following notations. For the GW tree \( T \), let \( Z_n^T \) be the number of vertices at the \( n \)-th generation. By Lemma 4.1 of [1], we have for any \( K_0 \geq 1 \),

\[
\mathbb{E}_{GW}(Z_n^T 1_{Z_n^T \leq K_0}) \leq K_0 n^{1/K_0} q_1^{n-K_0}.
\]

Let \( r \in (q_1, 1) \) be some real number to be chosen later, let

\[
n_0 = n_0(K_0, r) := \inf \{ n \geq 1, \mathbb{E}_{GW}(Z_n^T 1_{Z_n^T \leq K_0}) \leq r^n \},
\]

which is thus a finite integer. In fact, \( K_0 \) will be chosen according to Corollary 3. Define (recall that we write \( u < v \) if \( u \) is an ancestor of \( v \).)

\[
Z^n_T (u, n) = |\{ x \in T; u < x, |x| = |u| + n \}|.
\]

Let \( T_{n_0} \) be a tree induced from \( T \) in the following way: starting from the root \( \rho \), \( y \) is a child of \( x \) in \( T_{n_0} \) if \( x < y \) and \( |y| = |x| + n_0 \). Define a subtree \( \mathcal{W} \) of \( T_{n_0} \) by

\[
\mathcal{W} = \{ x \in T_{n_0}; \forall u \in T_{n_0}, u < x \Rightarrow Z^n_T (u, n_0) \leq K_0 \}.
\]

Let \( W_k \) be the population of the \( k \)-th generation of \( \mathcal{W} \), \( \mathcal{W} \) is a sub critical Galton Watson tree of mean offspring \( \mathbb{E}_{GW}(Z_0^T 1_{Z_0^T \leq K_0}) \leq r^{n_0} \); in particular, for any \( k \geq 0 \), \( \mathbb{E}_{GW}(W_k) \leq r^{kn_0} \).

For any \( y \in T \), let \( y_0 \) be the youngest ancestor of \( y \) in \( T_{n_0} \). For \( n \geq n_0 \), let \( j = \lfloor \frac{n}{n_0} \rfloor \geq 1 \) so that \( jn_0 \leq n < (j+1)n_0 \). Define

\[
N_{n,1} = \sum_{|y|=n} N(y) 1_{Z^n(y,y_0) > K_0}, \quad N_{n,1}^* = \sum_{|y|=n} N(y) \frac{A_y}{1 + A_y B_y} 1_{Z^n(y,y_0) > K_0} \tag{21}
\]
Figure 3: An example in the case $K_0 = n_0 = 2$.

\[ N_{n,2} = \sum_{|y|=n} N(y) \mathbb{1}_{z_T(y_0, n_0) \leq K_0, y_0 \notin W}, \quad N^*_{n,2} = \sum_{|y|=n} N(y) \frac{A_y}{1 + A_y B_y} \mathbb{1}_{z_T(y_0, n_0) \leq K_0, y_0 \notin W} \]  

(22)

**Lemma 8.** There exist \( r \in (q_1, 1) \) and \( K_0 > 0 \), such that, with the definitions of \( n_0, N_{n,1}, N^*_{n,1} \) above, for some constant \( L > 0 \), for any \( n \geq n_0 \)

\[ \mathbb{E}(N_{n,1}) \leq L, \quad \mathbb{E}(N^*_{n,1}) \leq L. \]  

(23)

**Lemma 9.** With the same assumption as in Lemma 8, let \( t^* \) be defined as in (10), then

\[ \mathbb{E}(N^\lambda_{n,2}) \leq L, \quad \mathbb{E}((N^*_{n,2})^\lambda) \leq L \]  

(24)

holds in either of the following cases

1) \( \lambda \in [0,1] \) and \( t^* > 3/2 \)

2) \( \lambda \in [0, t^* - \frac{1}{2}) \) and \( t^* \leq 3/2 \).

We are prepared to prove Proposition 2 and Proposition 3.

**Proof of Proposition 2** Since \( q_1 \mathbb{E}(A^{-1/2}) < 1 \), \( t^* > 3/2 \). We choose \( \lambda = 1 \). As \( W \) is finite a.s., if \( \chi = (\text{height}(W) + 1)n_0 \) (where for a finite tree \( T \), height\((T) := \max_{x \in T}|x|\)), then

for all \( n \geq \chi \), \( N_n \leq N_{n,1} + N_{n,2} \).

By Lemma 8, 9, for any \( n \geq n_0 \),

\[ \mathbb{E}(N_n, n \geq \chi) \leq 2L. \]

Thus,

\[ \liminf_{n \to \infty} \mathbb{E} \left[ \frac{\sum_{i=\chi}^n N_n}{n} \right] \leq 2L. \]

By Fatou’s lemma, a.s.

\[ \liminf_{n \to \infty} \frac{\sum_{k=-1}^n N_k}{n} = \liminf_{n \to \infty} \frac{\sum_{k=\chi}^n N_k}{n} < \infty. \]

Therefore,

\[ \frac{1}{v(\eta)} = \liminf_{n \to \infty} \frac{\tau_n}{n} \leq \liminf_{n \to \infty} \frac{\sum_{k=-1}^n N_k}{n} < \infty. \]
This implies that $v(\eta) > 0$.

The case for $Z_i$ can be treated in a similar manner with $N_n^*$ instead of $N_n$. Finally, to prove $v(Y) > 0$, it is enough to recall $Z_{D(t)} = Y_t$ where $D(t) = \sum x(L_x(t)^2 - c^2)$ and note that
\[
\frac{D(t)}{t} = \frac{\sum x(L_x(t) + c)(L_x(t) - c)}{\sum x(L_x(t) - c)} \geq 2c > 0.
\]

It follows that
\[
v(Y) = \lim_{t \to \infty} \frac{|Y_t|}{t} = \lim_{t \to \infty} \frac{|Z_{D(t)}|}{t} \geq v(Z) \liminf_{t \to \infty} \frac{D(t)}{t} \geq 2cv(Z).
\]

Proof of Proposition 3. If $q_t \mathbb{E}(A^{-1/2}) \geq 1$, $\lambda < t^* - 1/2 \leq 1$. Let $N_i(Z)$ be the time spent at the $i$-th generation by $(Z_t)$. Let $\Gamma_k(Z)$ be the regenerative times corresponding to $(Z_t)_{t \geq 0}$. Let $u(n)$ be the unique integer such that $\Gamma_{u(n)} \leq \tau_n(Z) < \Gamma_{u(n)+1}$. Then,
\[
\frac{\Gamma_{u(n)}(Z)^\lambda}{n} \leq \sum_{k \leq u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^{\lambda} = \sum_{k \leq u(n)} \left(\sum_{i=\max(Z_{k-1},Z_k)}^{\max(Z_k,n)} N_i(Z)\right)^\lambda
\]
\[
\leq \sum_{i \leq n} N_i(Z)^\lambda.
\]

Taking limit yields that
\[
\liminf_{n \to \infty} \frac{\Gamma_{u(n)}(Z)^\lambda}{n} \leq \liminf_{n \to \infty} \sum_{k \leq u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^{\lambda} \leq \liminf_{n \to \infty} \sum_{i \leq \lambda} N_i(Z)^\lambda.
\]

Applying Jensen’s inequality then Lemma 8 and Lemma 9 implies that
\[
\mathbb{E}[N_n(Z)^\lambda; n \geq \lambda] \leq \mathbb{E}[\mathbb{E}^u[T(N_n(Z); n \geq \lambda|\eta)]^\lambda \leq \mathbb{E}[(N_{n,1} + N_{n,2})^\lambda; n \geq \lambda] \leq 2L.
\]

It follows from Fatou’s lemma that
\[
\liminf_{n \to \infty} \frac{\Gamma_{u(n)}(Z)^\lambda}{n} \leq \liminf_{n \to \infty} \sum_{k \leq u(n)} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^{\lambda} < \infty.
\]

By law of large numbers,
\[
\lim_{n \to \infty} \frac{n}{u(n)} = \mathbb{E}_S[|Z\Gamma_1(Z)|] < \infty, \text{ and } \lim_{n \to \infty} \frac{\sum_{k \leq n} (\Gamma_k(Z) - \Gamma_{k-1}(Z))^{\lambda}}{n} = \mathbb{E}_S[\Gamma_1(Z)^\lambda].
\]

Therefore there exists a constant $C \in (0, \infty)$ such that
\[
\liminf_{n \to \infty} \frac{\Gamma_n(Z)^\lambda}{n} < C.
\]

Note that $|Z_t| \geq \#\{k : \Gamma_k(Z) < t\}$. So we get $|Z_t| \geq t^\lambda/C$ for all sufficiently large $t$. We hence deduce that
\[
\liminf_{t \to \infty} \frac{\log |Z_t|}{\log t} \geq \lambda.
\]

Letting $\lambda \uparrow t^* - 1/2$ yields
\[
\liminf_{n \to \infty} \frac{\log |Z_t|}{\log t} \geq t^* - 1/2.
\]

The result follows from Remark 4. Similar arguments can be applied to $\lim_{n \to \infty} \frac{\log |Z|}{\log n}$. □
It remains to show the main Lemmas 8, 9. Let us first state some preliminary results. As the walk is transient, the support of the random walk should be slim. This is formulated in the following lemma:

**Lemma 10.** There exists a constant $c_{11} > 0$ such that for any $n \geq 1$, $E\left( \sum_{|x|=n} \mathbb{I}_{\tau_x < \infty} \right) \leq c_{11}$.

The following lemma shows that, the escape probability is relatively large. In fact, we cannot show that $E\left( \frac{1}{\beta(p)} \right) < \infty$ for all $q_1 > 0$, however as the GW tree branches anyway, there will be a large copies of independent sub-trees, we show $E\left( \frac{1}{\sum_{i=1}^{\infty} \beta_i} \right) < \infty$ instead.

**Lemma 11.** Consider i.i.d. copies of GW trees $T^{(i)}$ rooted at $\rho^{(i)}$ with independent environment $\omega^{(i)}$, for each $T^{(i)}$, define $\beta_i = P_{\rho^{(i)}}^T(\tau_{\rho^{(i)}} = \infty)$. There exists an integer $K = K(q_1, c) \geq 1$ such that

$$E\left( \sum_{i=1}^{K} \beta_i \right) \leq c_{12} < \infty$$

and

$$E\left( \sum_{i=1}^{K} A_{\rho^{(i)}} \beta_i \right) < c_{12} < \infty.$$

Moreover, if $q_1 \xi_2 < 1$, then $E\left( \frac{1}{\beta(p)} \right) \leq c_{12} < \infty$ and $E\left( \frac{1}{A_{\rho^{(i)}} \beta_i} \right) < c_{12} < \infty$.

**Remarks 5.** In fact, if $q_1 E(A^{-2}) < 1$, a proof similar to Proposition 2.3 of [1] shows that $\eta$ has positive speed, in particular, the VRJP on any regular tree (except $\mathbb{Z}$) admits positive speed.

**Corollary 3.** There exists $K_0 \geq K$, such that for $k \in \{2, 4\}$

$$\mathbb{E}\left( \sum_{i=1}^{K_0} A_{\rho^{(i)}}^k \beta_i \right) < c_{13} < \infty.$$

The proof of Lemma 10, 11 and Corollary 3 will be postponed to the Appendix B, let us state the consequence of these preliminary results. Recall that $Z^n_T$ is the population at generation $n$, and that for any $x \in T$, $\tau_x$ is the first hitting time, $\tau_x^*$ the first return time to $x$. For $u, v \in T$ such that $u < v$ define

$$p_1(u, v) = P_{\rho}^u T (\tau_{\rho} = \infty, \tau_u^* = \infty, \tau_v = \infty)$$

**Lemma 12.** For any $n \geq 2$ and $k \in \{1, 2, 4\}$, consider $K_0$ as in Corollary 3, we have

$$\mathbb{E}\left( \sum_{|u|=n} \frac{1}{p_1(\rho, u)^k} \right) < c_{14}^2 < \infty.$$

In addition,

$$\mathbb{E}\left( \sum_{|u|=n} \frac{1}{p_1(\rho, u)^k} | A_{\rho} \right) < c_{14}^2 (1 + \frac{1}{A_{\rho}}). \quad (26)$$

**Proof of Lemma 12** Fix $n \geq 2$, let $Y_0 := \inf\{l \geq 1; Z_l > K_0\}$, then $\{Z^n_T > K_0\} = \{Y_0 \leq n\}$. For any $u \in T$ such that $|u| \geq Y_0$, let $U$ be its ancestor at the $Y_0$-th generation. By Markov property,

$$p_1(\rho, u) \geq \sum_{|y|=Y_0-1} P_{\rho}^w T (\tau_y < \tau^*_u) P_{\rho}^w T (\tau_{\rho} = \infty, \tau_U = \infty) \geq \sum_{|y|=Y_0-1} \prod_{i=0}^{Y_0-2} p(y_i, y_{i+1}) P_{\rho}^w T (\tau_{\rho} = \infty, \tau_U = \infty) \quad (27)$$
where \( \{y_0(=\rho), y_1, \ldots, y_{T_0-1}(=y)\} \) is the unique path connecting \( \rho \) and \( y \). Note that if \( U = y \), then

\[
P_y^{\omega,T}(\tau_{y_y} = \infty, \tau_U = \infty) = \sum_{z: z = y, z \neq U} p(y, z)\beta(z) + \sum_{z: z = y, z \neq U} p(y, z)(1 - \beta(z))P_y^{\omega,T}(\tau_{y_y} = \infty, \tau_U = \infty).
\]

Otherwise

\[
P_y^{\omega,T}(\tau_{y_y} = \infty, \tau_U = \infty) = \sum_{z: z = y} p(y, z)\beta(z) + \sum_{z: z = y} p(y, z)(1 - \beta(z))P_y^{\omega,T}(\tau_{y_y} = \infty, \tau_U = \infty)
\]

It follows that in both cases,

\[
P_y^{\omega,T}(\tau_{y_y} = \infty, \tau_U = \infty) = \frac{\sum_{z: z = y} 1\text{ if } z \neq U \text{ and } p(y, z)\beta(z)}{p(y, y) + \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } p(y, z)\beta(z)} \geq \frac{1 + A_y A_U + \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_y A_z \beta(z)}{1 + A_y 1 + A_U 1 + \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z)}
\]

Plugging it into (27) yields that

\[
p_1(\rho, u) \geq \frac{1}{1 + A_U} \min_{|y| = Y_0-1} \left( \prod_{i=0}^{Y_0-2} p(y_i, y_{i+1}) \frac{A_y}{1 + A_y} \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z) \right) \geq \frac{1}{1 + A_U} \min_{|y| = Y_0-1} \left( \prod_{i=0}^{Y_0-2} p(y_i, y_{i+1}) \frac{A_y}{1 + A_y} \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z) \right)
\]

Thus, for \( k \in \{1, 2, 4\},

\[
\frac{1}{p_1(\rho, u)^k} \leq (1 + A_U)^k \left( \frac{1}{1 + A_U} \min_{|y| = Y_0-1} \left( \prod_{i=0}^{Y_0-2} p(y_i, y_{i+1}) \frac{A_y}{1 + A_y} \right) \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z) \right)^k.
\]

Given the tree \( T \), by integrating w.r.t. \( P(d\omega) \), we have

\[
1_{n \geq Y_0} \sum_{|u|=n} 1_{p_1(u)^k} \leq E^T \left( \frac{1}{\min_{|y| = Y_0-1} \left( \prod_{i=0}^{Y_0-2} p(y_i, y_{i+1}) \frac{A_y}{1 + A_y} \right)^k \sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z) \right)^k
\]

\[
\times Z^T(U, n - Y_0) E^T [(1 + A_U)^k] E^T \left( \frac{1}{\sum_{z: z = y} 1\text{ if } z \neq U \text{ and } A_z \beta(z) \right)^k \right).
\]
It follows from Lemma 11 for $k = 1$ or Corollary 3 for $k = 2, 4$ that

\[
\mathbb{E}_Q \left( \mathbb{I}_{n \geq \gamma_0} \sum_{|u| = n} \frac{1}{p_1(\rho, u)^k} \left| \gamma_0, Z_l; 0 \leq l \leq \gamma_0 \right. \right) 
\leq c_{15} \mathbb{I}_{n \geq \gamma_0} \mathbb{E}_T \left[ \min_{y = \gamma_0-1} \left( \prod_{i=0}^{\gamma_0-2} p(y_i, y_{i+1}) \frac{A_y}{A_y A_{y+1}} \right)^k \right] \times \sum_{|U| = \gamma_0} \mathbb{E}\left[(1 + A)^k\right] b^{n - \gamma_0} 
\leq c_{16} \mathbb{I}_{n \geq \gamma_0} \sum_{|y| = \gamma_0-1} \mathbb{E}_T \left[ \prod_{i=0}^{\gamma_0-2} \left( 1 + A_{y_i} \right) \frac{1 + B_{y_i}}{A_{y_i} A_{y_i+1}} \right]^k \sum_{|U| = \gamma_0} b^{n - \gamma_0}.
\]

By independence of $A_x, x \in T$, we see that

\[
\mathbb{E}_T \left[ \prod_{i=0}^{\gamma_0-2} \left( 1 + A_{y_i} \right) \frac{1 + B_{y_i}}{A_{y_i} A_{y_i+1}} \right]^k \leq c_{17}^{\gamma_0-1},
\]

with $c_{17} \in (1, \infty)$. Consequently,

\[
\mathbb{E}_Q \left( \mathbb{I}_{n \geq \gamma_0} \sum_{|u| = n} \frac{1}{p_1(\rho, u)^k} \right) \leq \mathbb{E}_Q \left( c_{16} \mathbb{I}_{n \geq \gamma_0} \sum_{|y| = \gamma_0-1} c_{17}^{\gamma_0-1} \sum_{|U| = \gamma_0} b^{n - \gamma_0} \right)
\leq c_{16} K_0 \mathbb{E}_Q \left( \mathbb{I}_{n \geq \gamma_0} c_{17}^{n-1} Z_n^T \right)
\leq c_{18} (c_{17} b)^n < \infty.
\]

(26) follows in the same way. \hfill \square

**Proof of Lemma 3** We only bound $\mathbb{E}(N_{n,1})$, the argument for $\mathbb{E}(N^{*}_{n,1})$ is similar. For any $y \in T$ at the $n$-th generation such that $Z^T_n(y_0, n_0) > K_0$, let $Y$ be the youngest ancestor of $y$ such that $Z^T(Y, n_0) > K_0$. Clearly, $y_0 \leq Y \leq y$. So,

\[
N_{n,1} = \sum_{|y| = n} N(y) \mathbb{I}_{Z^T_n(y_0, n_0) > K_0} \leq \sum_{|y| = n} N(y) \mathbb{I}_{y_0 \leq Y \leq y}.
\]

Taking expectation w.r.t. $E_{\rho}^{\omega, T}$ implies that

\[
E_{\rho}^{\omega, T}(N_{n,1}) \leq \sum_{|y| = n} E_{\rho}^{\omega, T}(N(y)) \mathbb{I}_{y_0 \leq Y \leq y} = \sum_{|y| = n} P_{\rho}^{\omega, T}(\tau_y < \infty) E_{\rho}^{\omega, T}(N(y)) \mathbb{I}_{y_0 \leq Y \leq y}.
\]

Applying the Markov property at $\tau_Y$ to $E_{\rho}^{\omega, T}(N(y))$, we have

\[
E_{\rho}^{\omega, T}(N(y)) = G^{\tau_Y}(y, y) + P_{\rho}^{\omega, T}(\tau_y < \infty) E_{\rho}^{\omega, T}(\tau_y < \infty) E_{\rho}^{\omega, T}(N(y))
\]

where (write $\{(\tau_Y \wedge \infty) > \tau^*_y\} = \{\tau^*_y < \infty \text{ and } \tau^*_y < \tau_Y\}$ for short)

\[
G^{\tau_Y}(y, y) = E_{\rho}^{\omega, T}(\sum_{k=0}^{\tau_Y} \mathbb{I}_{\eta_k = y}) = \frac{1}{1 - P_{\rho}^{\omega, T}(\tau_y < \infty) E_{\rho}^{\omega, T}(\tau_y < \infty)}.
\]
Consequently,
\[
E_{y}^{\omega,T}(N(y)) = \frac{G^{\gamma_{y}}(y, y)}{1 - P_{y}^{\omega,T}(\tau_{y} < \infty)P_{y}^{\omega,T}(\tau_{Y} < \infty)} \leq \frac{G^{\gamma_{y}}(y, y)}{1 - P_{y}^{\omega,T}(\tau_{y} \neq \infty)} = \frac{G^{\gamma_{y}}(y, y)}{P_{y}^{\omega,T}(\tau_{Y} = \infty)}.
\]

We bound \(G^{\gamma_{y}}(y, y)\) first. As \(P_{y}^{\omega,T}(\tau_{Y} < \infty) \leq \sum_{z \in \mathbb{Z}} p(y, z) + p(y, y) P_{y}^{\omega,T}(\tau_{y} < (\tau_{Y} \wedge \infty))\),
\[
1 - P_{y}^{\omega,T}(\tau_{Y} < \infty) \geq p(y, y) \left(1 - P_{y}^{\omega,T}(\tau_{y} < \tau_{Y})\right).
\]
By Lemma 4.4 of [1] and (38), the right hand side of the above inequality is larger than
\[
p(y, y) \left(1 - \tilde{p}_{y}^{\omega,T}(\tilde{\tau}_{y} < \tilde{\tau}_{Y})\right).
\]
where we identify \(\tilde{p}_{y}^{\omega,T}\) to the probability of \((\tilde{n}, \tilde{z})\) on the segment \([y, \tilde{y}]\). Therefore,
\[
G^{\gamma_{y}}(y, y) \leq \left(1 + A_{y} \sum_{Y<z<y} A_{z} \prod_{z<u<y} A_{u}^{2}\right)\left(1 + A_{y} B_{y}\right) =: V_{y,Y}.
\]
Consequently,
\[
E_{y}^{\omega,T}(N(y)) \mathbb{I}_{Z_{T}(y_{0}, n_{0}) > K_{0}} \leq P_{y}^{\omega,T}(\tau_{Y} < \infty) \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}, y_{0} \leq Y \leq y}.
\]
Summing over all possibilities of \(Y\) yields that (recall that \(j = \lfloor \frac{y}{n_{0}} \rfloor\))
\[
E_{y}^{\omega,T}(N_{1,n}) \leq \sum_{l=j_{n_{0}}-1}^{n} \sum_{|Y|=l} P_{y}^{\omega,T}(\tau_{Y} < \infty) \sum_{|y|=n_{0}, Y \leq y} \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}}
\]
where the last inequality holds because \(P_{y}^{\omega,T}(\tau_{Y} < \infty) \leq P_{y}^{\omega,T}(\tau_{Y} < \infty)\) and \(P_{y}^{\omega,T}(\tau_{Y} = \infty) \geq P_{y}^{\omega,T}(\tau_{Y} = \infty) = \). Summing over the value of \(Y\) yields that
\[
E_{y}^{\omega,T}(N_{1,n}) \leq \sum_{l=j_{n_{0}}-1}^{n-1} \sum_{|x|=l} P_{y}^{\omega,T}(\tau_{x} < \infty) \sum_{Y=|x|}^{\sum_{|y|=n_{0}, Y \leq y} \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty, \tau_{Y} \neq \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}}.
\]
As conditionally on \(T\), \(P_{y}^{\omega,T}(\tau_{x} < \infty)\) and \(\sum_{Y=|x|}^{\sum_{|y|=n_{0}, Y \leq y} \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty, \tau_{Y} = \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}}\) are independent,
\[
E(N_{1,n}) \leq \mathbb{E} \left( \sum_{l=j_{n_{0}}-1}^{n-1} \sum_{|x|=l} E^{T}(P_{y}^{\omega,T}(\tau_{x} < \infty))E^{T} \left( \sum_{Y=|x|}^{\sum_{|y|=n_{0}, Y \leq y} \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty, \tau_{Y} = \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}} \right) \right)
\]

= \sum_{l=j_{n_{0}}-1}^{n-1} \mathbb{E}(\sum_{|x|=l} \mathbb{I}_{\tau_{x}<\infty}) \mathbb{E} \left( \sum_{|Y|=1}^{\sum_{|y|=n_{0}, Y \leq y} \frac{V_{y,Y}}{P_{y}^{\omega,T}(\tau_{Y} = \infty, \tau_{Y} = \infty)} \mathbb{I}_{Z_{T}(Y, n_{0}) > K_{0}} \right) \right).
\]
Note that for any $|Y| = 1$, 
\[
\frac{\sum_{y=n-l\leq y \leq Y} V_{y,Y}}{P^{\omega,T} (\tau^* = \infty, \tau^- = \infty)} \mathbb{1}_{Z^{T}(Y,n_0) > K_0}
\]
are i.i.d. By Lemma 10
\[
\mathbb{E}(N_{n,1}) \leq b_c 11 \sum_{l=1}^{n-1} A_{n-l}
\]
where
\[
A_{n-l} = \mathbb{E}\left( \frac{\sum_{y=n-l-1} V_{y,\rho}}{P^{\omega,T} (\tau^* = \infty, \tau^- = \infty)} \mathbb{1}_{Z^{T}(\rho,n_0) > K_0} \right).
\]
By Cauchy-Schwartz inequality,
\[
A_{n-l} \leq \sqrt{\mathbb{E}\left( \left( \sum_{y=n-l-1} V_{y,\rho} \right)^2 \right) \mathbb{E}\left[ \frac{\mathbb{1}_{Z^{T}(\rho,n_0) > K_0}}{P^{\omega,T} (\tau^* = \infty, \tau^- = \infty)} \right]^2}
\]
Recall that $Z_n^T$ denotes the number of vertices at the $n$-th generation of the tree $T$, using Lemma 12 then applying again Cauchy-Schwartz inequality to \( \left( \sum_{y=n-l-1} V_{y,\rho} \right)^2 \) implies that
\[
A_{n-l}^2 \leq c_{14} \mathbb{E}\left( Z_{n-l-1}^T \sum_{y=n-l-1} V_{y,\rho}^2 \right)
\leq c_{19} \mathbb{E}_{GW}[c_{20}^{n-l-1} \left( Z_{n-l-1}^T \right)^2],
\]
where the second inequality follows from $\mathbb{E}[V_{y,\rho}] \leq c_{20} |y|$. Plugging it into (28) implies that
\[
\mathbb{E}(N_{n,1}) \leq b_c 11 \sqrt{c_{19}} \sum_{l=1}^{n-1} \mathbb{E}_{GW}[c_{20}^{n-l-1} \left( Z_{n-l-1}^T \right)^2] \leq c_{21} \sum_{k=0}^{n_0} \sqrt{c_{20}^{k} \mathbb{E}_{GW}\left[ (Z_k^T)^2 \right]} \leq c_{22},
\]
since $\mathbb{E}_{GW}\left[ (Z_k^T)^2 \right] < \infty$. Analogously, for $N^*_{n,1}$ we get that
\[
E^{(\omega,T)}(N^*_{n,1}) \leq \sum_{l=0}^{n_0} \sum_{x \in 0} P^{\omega,T} (\tau_x < \infty) \sum_{y: Y = x} \frac{\sum_{y=n,Y \leq Y} V_{y,Y}}{P^{\omega,T} (\tau^* = \infty, \tau^- = \infty)} \mathbb{1}_{Z^{T}(Y,n_0) > K_0}.
\]
And recounting on the same arguments gives a finite upper bound for $\mathbb{E}[N^*_{n,1}]$. \quad \Box

**Proof of Lemma** \(\Box\) Again we only give the proof for $E(N^*_{n,2})$. For $y \in T$, as $Z^T(y_0,n_0) \leq K_0$ and $y_0 \notin W$, we can find the youngest ancestor $Y_1$ of $y$ in $T_{n_0}$ such that $Z^T(Y_1,n_0) > K_0$, automatically $Y_1 < y_0$. Let $Y_2$ be the youngest descendant of $Y_1$ in $T_{n_0}$ such that it is an ancestor of $y$. Let $Y_3$ be the youngest descendant of $y$ in $T_{n_0}$ such that $Z^T(Y_3,n_0) > K_0$.

For any $0 < \lambda \leq 1$,
\[
E^{(\omega,T)}(N^*_{n,2}) \leq E^{(\omega,T)} \left[ \sum_{y=n} N(y)^\lambda \mathbb{1}_{Z^{T}(y_0,n_0) \leq K_0, y_0 \notin W} \right]
\leq \sum_{y=n} \mathbb{1}_{Z^{T}(y_0,n_0) \leq K_0, y_0 \notin W} P^{(\omega,T)} (\tau_x < \infty) \left( E^{(\omega,T)}[N(y)] \right)^\lambda.
\]

In what follows, we identify $\tilde{P}^{\omega}$ with the distribution of a one-dimensional random walk $\tilde{y}$ on the path $[Y_1, Y_3]$. Let us state the following lemmas which will be used in (29).
Lemma 13. For any \( y \in T \) such that \( Y_1 < Y_2 < y < Y_3 \), let \( y^* \) be the unique child of \( y \) which is also ancestor of \( Y_3 \). Then,

\[
\left( E_y^{\omega, T}[N(y)] \right)^\lambda \leq \left( \frac{1 + A_y B_y}{1 + A_y A_y} \right) \lambda \frac{2}{p_1(Y_1, Y_2) p_y^{\omega, Y_3}} \frac{2}{\tilde{G}_{Y_1}^{\eta, \eta} (y, y^*)} \frac{1}{\tilde{p}_1(Y_1, Y_2)}.
\]  

(30)

where \( \tilde{G}_{Y_1}^{\eta, \eta} (y, y^*) \) is the Green function associated with \( \tilde{\eta}_n \).

Lemma 14.

\[
P^{\omega, T} (\tau_y < \infty) \leq P^{\omega, T} (\tau_{Y_1} < \infty) \tilde{P}_{Y_1}^{\omega} (\tilde{\tau}_y < \tilde{\tau}_{Y_1-1}) \frac{1}{\tilde{p}_1(Y_1, Y_2)} \frac{1}{\tilde{p}_1(Y_1, Y_2)}.
\]  

(31)

The proofs of Lemmas 13 and 14 can be found in section 5.2 of [1] with slight modifications, so we feel free to omit them (see (5.10) and (5.11) therein). Now plugging (30) and (31) into (29) yields that

\[
E_y^{\omega, T}[N_{n, 2}^\lambda] \leq \sum_{|y| = n, Y_1 = n, Y_2 = \rho} \frac{2 P^{\omega, T} (\tau_{Y_1} < \infty) \tilde{p}_1(Y_1, Y_2) p_{Y_3}^{\omega} (\tau_{Y_3} = \infty, \tau_{Y_3}^- = \infty)}{p_1(Y_1, Y_2)^2} \left( \frac{1 + A_y B_y}{1 + A_y A_y} \tilde{E}_{Y_1}^{\omega} [\tilde{\tau}_{Y_1} \land \tilde{\tau}_{Y_3}] \right)^\lambda.
\]  

By Lemma 4 one sees that

\[
E_y^{\omega, T}[N_{n, 2}^\lambda] \leq \sum_{|y| = n} p_1(Y_1, Y_2) p_{Y_3}^{\omega} (\tau_{Y_3} = \infty, \tau_{Y_3}^- = \infty) \left( \frac{1 + A_y B_y}{1 + A_y A_y} \tilde{E}_{Y_1}^{\omega} [\tilde{\tau}_{Y_1} \land \tilde{\tau}_{Y_3}] \right)^\lambda.
\]  

where \( Y_2^* \) is the children of \( Y_2 \) along \([Y_2, Y_3]\). Decompose the sum over \(|y| = n\) by

\[
\sum_{|y| = n} = \sum_{|y| = n, Y_1 = \rho} + \sum_{l = 1}^{(j - 1)} \sum_{|x| = n_0 - 1} \sum_{y: Y_1 = x, |y| = n}.
\]
We get that
\[
E^{\omega,T}(N_{n,2}^\lambda) \leq \sum_{|y| = n, \lambda_Y = \rho} \frac{2S_{\lambda,[\rho,Y]}^T}{p_1(\rho, Y_2)^2} \Theta_\lambda(Y_2, y, Y_3) + \sum_{l=1}^{j-1} \sum_{|x| = l, n_0-1} \sum_{|y| = n, \lambda_Y = x} \frac{2P_\omega^{\omega,T}(\tau_Y^*, \tau_Y^* = \infty, \tau_Y^* = \infty)}{p_1(Y_1, Y_2)^2} \Theta_\lambda(Y_2, y, Y_3),
\]
where
\[
\Theta_\lambda(Y_2, y, Y_3) := \left(1 + A_y B_y\right)^\lambda \left(1 + A_Y^\lambda \left(1 + \tilde{E}_2^{\omega,T}[\tilde{\tau}_Y \wedge \tilde{\tau}_Y]^{\lambda}\right)\right).
\]

Given the GW tree \(T\), note that \(S_{\lambda,[Y_1,Y_2]} \in \sigma\{A_x : Y_1 \leq z \leq Y_2\}, p_1(\rho, Y_2) \in \sigma\{A_u : u \in (T \setminus T_Y) \cup \{Y_2\}\}, P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_Y^* = \infty) \in \sigma\{A_u : u \in T_Y\} \) and \(\Theta_\lambda(Y_2, y, Y_3) \in \sigma\{A_u : Y_2 < u \leq Y_3\}\). Therefore,
\[
E^{T}[N_{n,2}^\lambda] \leq \sum_{|y| = n, \lambda_Y = \rho} E^T \left[\frac{2S_{\lambda,[\rho,Y]}^T}{p_1(\rho, Y_2)^2} \Theta_\lambda(Y_2, y, Y_3) \mathbb{1}_{Z^T(Y_3, n_0) > K_0}\right] + \sum_{l=1}^{j-1} \sum_{|x| = l, n_0-1} \sum_{|y| = n, \lambda_Y = x} E^T \left[\frac{2P_\omega^{\omega,T}(\tau_Y^* < \infty) S_{\lambda,[Y_1,Y_2]}^T}{p_1(Y_1, Y_2)^2} \Theta_\lambda(Y_2, y, Y_3) \mathbb{1}_{Z^T(Y_3, n_0) > K_0}\right].
\]
(32)

Observe that
\[
P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_Y^* = \infty) \geq p_1(Y_3, u) \mathbb{1}_{Y_3 < u \leq |Y_3| + n_0}.
\]

Applying Lemma 12 to the subtree rooted at \(Y_3\) implies that
\[
E^T \left[\frac{\Theta_\lambda(Y_2, y, Y_3) \mathbb{1}_{Z^T(Y_3, n_0) > K_0}}{P_Y^{\omega,T}(\tau_Y^* = \infty, \tau_Y^* = \infty)}\right] \leq c_2 E^T \left[\left(1 + \frac{1}{A_{Y_3}}\right) \Theta_\lambda(Y_2, y, Y_3)\right].
\]

Plugging it into (32) implies that
\[
E^T[N_{n,2}^\lambda] \leq \Delta_1(n) + \Delta_2(n),
\]
where
\[
\Delta_1(n) := 2c_3 \sum_{|y| = n, \lambda_Y = \rho} E^T \left[\frac{S_{\lambda,[\rho,Y]}^T}{p_1(\rho, Y_2)^2}\right] E^T \left[(1 + \frac{1}{A_{Y_3}}) \Theta_\lambda(Y_2, y, Y_3)\right],
\]
(33)
\[
\Delta_2(n) := 2c_3 \sum_{l=1}^{j-1} \sum_{|x| = l, n_0-1} \sum_{|y| = n, \lambda_Y = x} E^T \left[\frac{P_\omega^{\omega,T}(\tau_Y^* < \infty) S_{\lambda,[Y_1,Y_2]}^T}{p_1(Y_1, Y_2)^2}\right] E^T \left[(1 + \frac{1}{A_{Y_3}}) \Theta_\lambda(Y_2, y, Y_3)\right].
\]
(34)
So, \( \mathbb{E}[N_{n,2}^\lambda] \leq \mathbb{E}_\lambda[\Delta_1(n) + \Delta_2(n)] \). (35)

We firstly bound \( \Delta_1(n) \), note that (since \( \lambda \leq 1 \))
\[
\left( \frac{1 + A_y B_y}{1 + A_y A_y} \right)^\lambda \leq \left( 1 + \sum_{z:z = y, z \neq y^*} A_z^\lambda \right) \leq 1 + \sum_{z:z = y, z \neq y^*} \frac{A_z^\lambda}{A_y},
\]
with \( \sum_{z:z = y, z \neq y^*} 1 \leq K_0 \). If \( |Y_2| = mn_0 < n, |Y_3| = (m + k)n_0 > n \), by Markov property and the fact that \( \{A_z, z \in [Y_2, Y_3] := [-1, kn_0 - 1]\} \),
\[
\mathbb{E}^T \left[ (1 + \frac{1}{A_{\lambda}}) \Theta_\lambda(Y_2, y, Y_3) \right] \leq \mathbb{E}^T \left[ \left( 1 + \frac{1}{A^{\lambda}_{\text{mn}_0 - 1}} \right) \left( 1 + \sum_{z:z = y, z \neq y^*} A_z^{\lambda} \right)(1 + A_\lambda^{\lambda}(1 + E_0^\omega(\tilde{\tau}_1 \land \tilde{\tau}_{\text{mn}_0 - 1}^{\lambda})) \right) \leq c_{24} + c_{24} \mathbb{E} \left[ 25(q_1 + \delta)^{-|Y_3| + |Y_2| + 1} \right]. (36)\]

Applying Cauchy-Schwartz inequality to \( \mathbb{E}^T \left[ \frac{S_{\lambda,1}^2}{p_1(p, Y_2)^2} \right] \) yields
\[
\Delta_1(n) \leq c_{23} \sum_{|y| = n, Y_1 = \rho} 2 \left( \sqrt{\mathbb{E}^T \left[ S_{\lambda,1}^2 \right] \mathbb{E}^T \left[ \frac{1}{p_1(p, Y_2)^2} \right] \mathbb{E}^T \left[ (1 + \frac{1}{A_{\lambda}}) \Theta_\lambda(Y_2, y, Y_3) \right] \right) \right) \leq c_{26} \sum_{|y| = n, Y_1 = \rho} \mathbb{E}^T \left[ \frac{1}{p_1(p, Y_2)^2} \right] \mathbb{E}^T \left[ (1 + \frac{1}{A_{\lambda}}) \Theta_\lambda(Y_2, y, Y_3) \right],
\]
where the last inequality holds because \( \mathbb{E}^T \left[ S_{\lambda,1}^2 \right] \leq c_{27}(n_0) < \infty \). By (36),
\[
\Delta_1(n) \leq c_{28} \sum_{|y| = n, Y_1 = \rho} \mathbb{E}^T \left[ \frac{1}{p_1(p, Y_2)^2} \right] \left( q_1 + \delta \right)^{-|Y_3| + |Y_2| + 1} \leq c_{28} \mathbb{E}^T \left[ \sum_{|y| = n, Y_2 = y} \prod_{|u| = n_0}^{K_0} \frac{1}{p_1(p, u)^2} \sum_{y:|y| = n, Y_2 = y} (q_1 + \delta)^{-|Y_3| + n_0 + 1} \right].
\]

Observe that
\[
\sum_{y:|y| = n, Y_2 = y} (q_1 + \delta)^{-|Y_3| + n_0 + 1} \leq \sum_{z:|z| > n, z \in W(T_u)} (q_1 + \delta)^{-|z| + n_0 + 1}.
\]

Hence,
\[
\Delta_1(n) \leq c_{28} \mathbb{E}^T \left[ \sum_{|y| = n_0} \prod_{|u| = n_0}^{K_0} \frac{1}{p_1(p, u)^2} \sum_{z:|z| > n, z \in W(T_u)} (q_1 + \delta)^{-|z| + n_0 + 1} \right].
\]
Plugging the above inequality and (37) into (35) implies that

\[ \mathbb{E}_Q[\Delta_1(n)] \leq c_{28} \mathbb{E}\left[ \sum_{|z|>n_0} \mathbb{I}_{\sum_{|z|>n_0} z\in \mathcal{W}} \right] \mathbb{E}_Q\left[ \sum_{|z|=n_0} (q_1 + \delta)^{-|z|+1} \right], \]

which by Lemma 12 is bounded by

\[ c_{29} \mathbb{E}_Q\left[ \sum_{|z|>n_0} (q_1 + \delta)^{-|z|+1} \right] = c_{29} \sum_{l>n_0-1} \mathbb{E}_Q\left[ \sum_{|z|=l} (q_1 + \delta)^{-|z|+1} \right]. \]

Recall that \( \mathcal{W} \) is a GW tree of mean \( \mathbb{E}[Z_{n_0}; Z_{n_0} \leq K_0] \leq r^{n_0} \). We can choose \( r \) to be \( q_1 + \delta/2 \) so that

\[ \sum_{l \geq 1} \mathbb{E}_Q\left[ \sum_{|z|=l} (q_1 + \delta)^{-|z|+1} \right] \leq \sum_{l \geq 1} (q_1 + \delta)^{-l-1} r^{l_0} < c_{30} \gamma^0, \]

where \( \gamma := (q_1 + \delta/2)^{n_0} < 1 \) and \( l_0 := \left\lfloor \frac{n}{n_0} \right\rfloor - 1 = j - 1 \). As a result, for any \( n > n_0 \),

\[ \mathbb{E}_Q[\Delta_1(n)] \leq c_{31} \gamma^0 < \infty. \] (37)

Turn to \( \Delta_2(n) \). As \( P_{\omega,T}(\tau_{Y_1} < \infty) \leq P_{\omega,T}(\tau_{\omega,T} < \infty) \), one sees that

\[ \Delta_2(n) \leq 2c_{23} \sum_{l=1}^{j-1} \sum_{|z|=l_0-1} \mathbb{E}_T\left[ \frac{P_{\omega,T}(\tau_{x} < \infty) S_{\lambda}[Y_1,Y_2]}{p_1(Y_1,Y_2)^2} \right] \mathbb{E}_T\left[ (1 + \frac{1}{A_{\gamma}}) \Theta_\lambda(Y_2, y, Y_3) \right], \]

which equals to

\[ \sum_{l=1}^{j-1} \sum_{|z|=l_0-1} \mathbb{P}_T(\tau_{x} < \infty) 2c_{23} \sum_{|y|=n, Y_1=x} \mathbb{E}_T\left[ \frac{S_{\lambda}[Y_1,Y_2]}{p_1(Y_1,Y_2)^2} \right] \mathbb{E}_T\left[ (1 + \frac{1}{A_{\gamma}}) \Theta_\lambda(Y_2, y, Y_3) \right], \]

as \( P_{\omega,T}(\tau_{x} < \infty) \) and \( \frac{S_{\lambda}[Y_1,Y_2]}{p_1(Y_1,Y_2)^2} \) are independent under \( \mathbb{P}_T \).

Note that for all \( z \in T \), \( 2c_{23} \sum_{|y|=n, Y_1=x} \mathbb{E}_T\left[ \frac{S_{\lambda}[Y_1,Y_2]}{p_1(Y_1,Y_2)^2} \right] \mathbb{E}_T\left[ (1 + \frac{1}{A_{\gamma}}) \Theta_\lambda(Y_2, y, Y_3) \right] \) are i.i.d. copies of \( \Delta_1(n - |z|) \). Taking expectation yields that

\[ \mathbb{E}_Q[\Delta_2(n)] \leq \sum_{l=1}^{j-1} \mathbb{E}\left[ \sum_{|z|=l_0-1} \mathbb{I}_{\tau_{x} < \infty}(d(x) - 1) \right] \mathbb{E}_Q[\Delta_1(n - l_0)] \]

\[ \leq bc_{31} \sum_{l=1}^{j-1} \mathbb{E}\left[ \sum_{|z|=l_0-1} \mathbb{I}_{\tau_{x} < \infty} \right] \gamma^{j-l-1}, \]

where the last inequality follows from (37). By Lemma 10 for any \( j \geq 2 \),

\[ \mathbb{E}_Q[\Delta_2(n)] \leq c_{32} \sum_{l=1}^{j-1} \gamma^{j-l-1} \leq c_{33} < \infty. \]

Plugging the above inequality and (37) into (35) implies that

\[ \mathbb{E}[N_{n,2}^\lambda] \leq \mathbb{E}_Q[\Delta_1(n)] + \mathbb{E}_Q[\Delta_2(n)] < \infty. \]

The estimate of \( \mathbb{E}[(N_{n,2}^\lambda)^4] \) follows from similar arguments. We feel free to omit it. \( \square \)
A Proofs of one dimensional results

Proof of Lemma 2. For any \( i \geq 1 \), let \( S_i = -\sum_{j=1}^{i} \log(A_j A_{j-1}) \) and define \( S_0 = 0 \). As \( i \mapsto \tilde{P}_i^\omega(\tilde{\tau}_i - 1 > \tilde{\tau}_n) \) is the solution to the Dirichlet problem

\[
\begin{cases}
\varphi(-1) = 0, \quad \varphi(n) = 1 \\
E_i^\omega(\varphi(\tilde{\eta}_i)) = \varphi(i)
\end{cases}
i \in [0, n - 1].
\]

It follows that

\[
\tilde{P}_i^\omega(\tilde{\tau}_- > \tilde{\tau}_n) = \frac{\sum_{j=0}^{i} \exp(S_j)}{\sum_{j=0}^{n} \exp(S_j)}.
\]

As a consequence, for any \( 0 \leq l \leq n \),

\[
\tilde{P}_0^\omega(\tilde{\tau}_1 < \tilde{\tau}_- - 1) = \frac{1}{\sum_{j=0}^{l} \exp(S_j)} \geq \frac{\exp(-\max_{0 \leq j \leq l} S_j)}{l + 1}
\]

\[
\tilde{P}_{l+1}^\omega(\tilde{\tau}_- > \tilde{\tau}_n) = \frac{\exp(S_{l+1})}{\sum_{j=l+1}^{n} \exp(S_j)} \leq \exp(-\max_{0 \leq j \leq l} (S_j - S_{l+1}))
\]

\[
\tilde{P}_{l-1}^\omega(\tilde{\tau}_- < \tilde{\tau}_n) = \frac{\exp(S_l)}{\sum_{j=0}^{l} \exp(S_j)} \leq \exp(-\max_{0 \leq j \leq l} (S_j - S_l)).
\]

We only need to consider \( n \) large, take \( l = \lfloor z_1 n \rfloor \), note that

\[
\tilde{P}_l^\omega(\tilde{\tau}_- > \tilde{\tau}_- - 1 \wedge \tilde{\tau}_n) = p(l, l + 1) \tilde{P}_{l+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_n) + p(l, l - 1) \tilde{P}_{l-1}^\omega(\tilde{\tau}_- < \tilde{\tau}_n)
\]

\[
\leq \max(\tilde{P}_{l+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_n), \tilde{P}_{l-1}^\omega(\tilde{\tau}_- < \tilde{\tau}_n)).
\]

Therefore,

\[
\tilde{P}_0^\omega(\tilde{\tau}_- \wedge \tilde{\tau}_- > m) \geq \tilde{P}_0^\omega(\tilde{\tau}_1 < \tilde{\tau}_- - 1) \tilde{P}_1^\omega(\tilde{\tau}_1^* < \tilde{\tau}_- - 1 \wedge \tilde{\tau}_n)^m
\]

\[
\geq \frac{\exp(-\max_{0 \leq j \leq l} S_j)}{l + 1} \left(1 - \tilde{P}_1^\omega(\tilde{\tau}_1^* \geq \tilde{\tau}_- - 1 \wedge \tilde{\tau}_n)\right)^m
\]

\[
\geq \frac{\exp(-\max_{0 \leq j \leq l} S_j)}{l + 1} \left(1 - \exp(-\max_{l+1 \leq k \leq n} (S_k - S_{l+1}) \wedge \max_{0 \leq k \leq l} (S_k - S_l))\right)^m
\]

\[
\geq \frac{1}{l + 1} \max_{0 \leq k \leq l} (S_k - S_{l+1}) \geq z_1\frac{1}{l + 1} \max_{0 \leq k \leq l} (S_k - S_l) \geq z_1.
\]

As \( m \approx e^{zn} \), we have \( (1 - e^{-zn})^m = O(1) \), taking expectation under \( P(\cdot | A_0 \in [a, \frac{1}{a}]) \) yields

\[
\tilde{P}_0(\tilde{\tau}_- \wedge \tilde{\tau}_- > m | A_0 \in [a, \frac{1}{a}])
\]

\[
\geq \frac{c}{n} P(\max_{0 \leq k \leq l} S_k \leq 0, \max_{0 \leq k \leq l} (S_k - S_l) \geq zn | A_0 \in [a, \frac{1}{a}]) P(\max_{l+1 \leq k \leq n} (S_k - S_{l+1}) \geq zn)
\]

\[
\geq \frac{c}{n} P(\max_{0 \leq k \leq l} S_k \leq 0, S_l \leq -zn | A_0 \in [a, \frac{1}{a}]) P((S_n - S_{l+1}) \geq zn).
\]

For \( k \geq 1 \), write \( \mathcal{S}_k = -\sum_{i=1}^{k} \log A_i \), then as \( S_k = -\log A_0 + \mathcal{S}_{k-1} + \mathcal{S}_k \),

\[
P(\max_{0 \leq k \leq l} S_k \leq 0, S_l \leq -zn | A_0 \in [a, \frac{1}{a}])
\]

\[
\geq P(A_0 \geq 1, A_l \geq 1, \max_{1 \leq k \leq l-1} \mathcal{S}_k \leq 0, \mathcal{S}_l \leq -\frac{zn}{2} | A_0 \in [a, \frac{1}{a}])
\]

\[
= P(A_0 \geq 1 | A_0 \in [a, \frac{1}{a}]) P(A_l \geq 1 | \mathcal{S}_l \leq -\frac{zn}{2})
\]

\[
= P(A_0 \geq 1 | A_0 \in [a, \frac{1}{a}]) P(\max_{1 \leq k \leq l-1} \mathcal{S}_k \leq 0, \mathcal{S}_l \leq -\frac{zn}{2})
\]
note that
\[
\mathbb{P}(\max_{1 \leq k \leq l-1} \mathcal{I}_k \leq 0, \mathcal{I}_{l-1} \leq -\frac{zn}{2}) \geq \frac{1}{l} \mathbb{P}(\mathcal{I}_{l-1} \leq -\frac{zn}{2})
\]
and
\[
S_n - S_{l+1} = -\log A_{l+1} - \log A_n - 2 \sum_{k=l+2}^{n-1} \log A_k.
\]

Therefore,
\[
\tilde{P}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \geq \frac{c}{n^2} \mathbb{P}(\mathcal{I}_{l-1} \leq -\frac{zn}{2}) \mathbb{P}(S_n - S_{l+1} \geq zn)
\]
\[
\geq \frac{c}{n^2} \mathbb{P}(\mathcal{I}_{l-1} \leq -\frac{zn}{2}) \mathbb{P}(A_{l+1} \leq 1) \mathbb{P}(A_n \leq 1) \mathbb{P}(\sum_{k=l+2}^{n-1} \log A_k \geq \frac{zn}{2})
\]
\[
\geq \frac{c}{n^2} \mathbb{P}(\sum_{k=1}^{l-1} \log A_k \geq \frac{zn}{2}) \mathbb{P}(\sum_{k=l+2}^{n-1} \log A_k \leq -\frac{zn}{2})
\]

Applying Cramér’s theorem to sums of i.i.d. random variables \(\log A_k\), we have
\[
\tilde{P}_0(\tilde{\tau}_n \wedge \tilde{\tau}_{-1} > m | A_0 \in [a, \frac{1}{a}]) \gtrsim n \exp(-n \left(z_1 I\left(\frac{z}{2z_1}\right) + (1-z_1) I\left(\frac{-z}{2(1-z_1)}\right)\right))
\]
where \(I(x) = \sup_{t \in \mathbb{R}} \{tx - \log \mathbb{E}(A^t)\}\) is the associated rate function.

\[\square\]

\textbf{Proof of Lemma 3} Replace \(I\left(\frac{z}{2(1-z_1)}\right)\) using
\[
I(-x) = \sup_{t \in \mathbb{R}} \{-tx - \log \mathbb{E}(A^t)\} = \sup_{t \in \mathbb{R}} \{-tx - \log \mathbb{E}(A^{l-t})\}
\]
\[
= \sup_{s \in \mathbb{R}} \{-(1-s)x - \log \mathbb{E}(A^s)\} = I(x) - x.
\]

For fixed \(z\), by convexity of the rate function \(I\), the supremum of \(-z_1 I\left(\frac{z}{2z_1}\right) - (1-z_1) I\left(\frac{-z}{2(1-z_1)}\right)\) is obtained when \(z_1 = \frac{1}{2}\), we are left to compute
\[
\sup_{0 < s} \left\{ \frac{\log q_s - I(z)}{z} + \frac{1}{2} \right\},
\]
clearly \(\frac{\log q_s - I(z)}{z} \leq -t^*\), when \(z\) is such that \((t \mapsto \log \mathbb{E}(A^t))'(t^*) = z > 0\), the maximum is obtained.

\[\square\]

\textbf{Proof of Lemma 4}. Observe that
\[
\tilde{P}_{Y_1}(\tilde{\tau}_y < \tilde{\tau}_{Y_1}) G_{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}}(y, y) = \tilde{P}_{Y_1}(\tilde{\tau}_y < \tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}) \tilde{E}_y^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}} \left[ \sum_{k=0}^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}} 1_{\{\tilde{\tau}_k = y\}} \right]
\]
\[
\leq \tilde{P}_{Y_1}(\tilde{\tau}_y < \tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}) \tilde{E}_y^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}} \left[ \sum_{k=0}^{\tilde{\tau}_{Y_1} \wedge \tilde{\tau}_{Y_3}} 1_{\{\tilde{\tau}_k = y\}} \right] = \tilde{E}_y^{\tilde{\tau}_{Y_1}} \left[ \sum_{k=0}^{\tilde{\tau}_{Y_1}} 1_{\{\tilde{\tau}_k = y\}} \right].
\]
Obviously,\[
\widetilde{E}^\omega_t \left[ \sum_{k=0}^{\tau_{Y_1} \wedge \tau_{Y_3}} 1_{\{\tilde{q}_k = y\}} \right] \leq \widetilde{E}^\omega_t [\tau_{Y_1} \wedge \tau_{Y_3}] \].
\]
This gives us \((11)\).
Moreover, to get \((12)\), we only need to show that for any \(0 \leq p < m\), we have
\[
\tilde{E}^\omega_p [\tau_{p-1} \wedge \tau_{m}] \leq 1 + A_p A_{p+1} + A_p A_{p+1} \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}] \].
\]
In fact, since \(0 \leq \lambda \leq 1\), \((39)\) implies that
\[
\tilde{E}^\omega_p [\tau_{p-1} \wedge \tau_{m}] \leq 1 + (A_p A_{p+1})^\lambda + (A_p A_{p+1})^\lambda \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}]^\lambda.
\]
Applying this inequality a few times along the interval \([Y_1, Y_3]\), we obtain \((12)\). It remains to show \((39)\). Observe that
\[
E^\omega_t [\tau_{p-1} \wedge \tau_{m}] = \tilde{E}^\omega_t (p, p-1) + \tilde{E}^\omega_t (p, p+1) (1 + \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}])
\]
\[
= 1 + \tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}]
\]
\[
= 1 + \tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}] + \tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}] + \tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}].
\]
It follows that
\[
\tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}] = \frac{1 + \tilde{E}^\omega_t (p, p+1) \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}]}{1 - \tilde{E}^\omega_t (p, p+1) e^\omega_{p+1} [\tau_{p} < \tau_{m}]} \leq \frac{1 + \tilde{E}^\omega_t (p, p+1) \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}]}{\tilde{E}^\omega_t (p, p-1)}.
\]
Therefore,
\[
\tilde{E}^\omega_t [\tau_{p-1} \wedge \tau_{m}] \leq (1 + A_p A_{p+1}) + A_p A_{p+1} \tilde{E}^\omega_{p+1} [\tau_{p-1} \wedge \tau_{m}] \.
\]
\[
\text{Proof of Lemma } 5 \text{ Recall that } E[A^t] < \infty \text{ for any } t \in \mathbb{R}. \text{ By Hölder’s inequality, it suffices to show that there exists some } \delta' > 0 \text{ such that for all } n \text{ large enough,}
\]
\[
E \left[ \left( \tilde{E}^\omega_0 [\tau_{t-1} \wedge \tau_{n}] \right)^{\lambda (1 + \delta')} \right] \leq (q_1 + \delta)^{-n}.
\]
\[
\text{It remains to prove } (40). \text{ In fact, we only need to show that for } 1 > \lambda' = \lambda (1 + \delta) > 0,
\]
\[
\limsup_{n \to \infty} \frac{\log E \left[ \left( \tilde{E}^\omega_0 [\tau_{t-1} \wedge \tau_{n}] \right)^{\lambda'} \right]}{n} \leq \psi (\lambda' + 1/2)
\]
where \(\psi(t) = \log E(A^t)\). One therefore sees that if \(t^* - 1/2 > \lambda'\), then \(\psi (\lambda' + 1/2) < \psi (t^*) = - \log q_1\). To show \((41)\), recall that for any \(0 \leq i \leq n - 1\),
\[
\tilde{G}^{\tau_{t-1} \wedge \tau_{n}} (i, i) = \tilde{E}^\omega_t \left[ \left( \sum_{k=0}^{\tau_{t-1} \wedge \tau_{n}} 1_{\eta = i} \right) \right]
\]
\[
= \frac{1}{1 - \tilde{\omega}(i, i - 1) \tilde{P}_{i-1} (\tau_i < \tau_{t-1}) - \tilde{\omega}(i, i + 1) \tilde{P}_{i+1} (\tau_i < \tau_{n})}.
\]
Then, $\tilde{E}_0^\omega[\tilde{\tau}_- \land \tilde{\tau}_n] = 1 + \sum_{i=0}^{n-1} \tilde{P}_0^\omega(\tilde{\tau}_i < \tilde{\tau}_-)$ implies that

$$
\tilde{E}_0^\omega[\tilde{\tau}_- \land \tilde{\tau}_n] = 1 + \sum_{i=0}^{n-1} \tilde{P}_0^\omega(\tilde{\tau}_i < \tilde{\tau}_-)
$$

Recall that by (38), if $S_i := \sum_{j=1}^{i} - \log(A_{j-1}A_j)$ for $i \geq 1$ and $S_0 = 0$, then

$$
\tilde{P}_0^\omega(\tilde{\tau}_i < \tilde{\tau}_-) = \frac{1}{\sum_{k=0}^{i} e^{S_k}} \quad \quad \tilde{P}_{i-1}^\omega(\tilde{\tau}_{i-1} < \tilde{\tau}_i) = \frac{1}{\sum_{k=0}^{i} e^{S_k}}
$$

$$
\tilde{P}_{i+1}^\omega(\tilde{\tau}_n < \tilde{\tau}_i) = \frac{1}{\sum_{k=i+1}^{n} e^{S_k-S_{i+1}}}
$$

It is immediate that

$$
\tilde{P}_0^\omega(\tilde{\tau}_i < \tilde{\tau}_-) = \frac{1}{\sum_{k=0}^{i} e^{S_k}} \quad \quad \frac{1}{\sum_{k=i+1}^{n} e^{S_k-S_{i+1}}}
$$

Let $X_k = - \log A_k$. For any $0 \leq i \leq n$, define

$$
H_i(-X) := \max_{0 \leq j \leq i} (-X_j - X_{j+1} - \cdots - X_{i-1}),
$$

$$
H_{n-i-1}(X) := \max_{i+2 \leq j \leq n} (X_{i+2} + \cdots + X_j).
$$

Note that

$$
S_k - S_i \leq 2H_i(-X) + (-X_i)_+, \forall 0 \leq k \leq i,
$$

and that

$$
S_k - S_{i+1} \leq 2H_{n-i-1}(X) + (X_{i+1})_+, \forall i+1 \leq k \leq n.
$$

Then,

$$
\frac{1}{1 + A_i A_{i+1}} \sum_{k=0}^{i} e^{S_k} \geq \frac{1}{1 + A_i A_{i+1}} \frac{1}{(1 + i) e^{2H_i(-X) + (-X_i)_+}} \geq \frac{1}{n(A_i + 1)(1 + A_i A_{i+1})} e^{-2H_i(-X)}.
$$

Similarly,

$$
\frac{A_i A_{i+1}}{1 + A_i A_{i+1}} \sum_{k=i+1}^{n} e^{S_k-S_{i+1}} \geq \frac{(A_{i+1} \land 1) A_i A_{i+1}}{n(1 + A_i A_{i+1})} e^{-2H_{n-i-1}(X)}.
$$

So,

$$
\frac{1}{1 + A_i A_{i+1}} \sum_{k=0}^{i} e^{S_k} + \frac{A_i A_{i+1}}{1 + A_i A_{i+1}} \sum_{k=i+1}^{n} e^{S_k-S_{i+1}} \geq \frac{1}{n(A_i + 1)(1 + A_i A_{i+1})} e^{-2H_i(-X)} + \frac{(A_{i+1} \land 1) A_i A_{i+1}}{n(1 + A_i A_{i+1})} e^{-2H_{n-i-1}(X)}.
$$

$$
\geq \frac{1}{n(A_i \lor 1)(1 + A_i A_{i+1})} \land \frac{(A_{i+1} \land 1) A_i A_{i+1}}{1 + A_i A_{i+1}} e^{-2H_i(-X)} \lor e^{-2H_{n-i-1}(X)}.
$$
This implies that
\[
\tilde{P}^\omega_0(\tilde{\tau}_i < \tilde{\tau}_{-1})
\leq n \left( (A_i \lor 1)(1 + A_i A_{i+1}) + \frac{1 + A_i A_{i+1}}{(A_{i+1} \land 1) A_i A_{i+1}} \right) e^{2H_i(\omega) \land H_{n-i-1}(X)}.
\]
Thus, for any \( \lambda \leq 1, n \geq 2, \)
\[
\tilde{E}^\omega_0[\tilde{\tau}_{-1} \land \tilde{\tau}_n] \leq n + n^2 \sum_{i=0}^{n-1} \left( (A_i \lor 1)(1 + A_i A_{i+1}) + \frac{1 + A_i A_{i+1}}{(A_{i+1} \land 1) A_i A_{i+1}} \right) e^{2\lambda H_i(\omega) \land H_{n-i-1}(X)}.
\]
By independence,
\[
\mathbb{E} \tilde{E}^\omega_0[\tilde{\tau}_{-1} \land \tilde{\tau}_n] \leq n + n^2 \max_{0 \leq i \leq n-1} \mathbb{E}[e^{2\lambda H_i(\omega) \land H_{n-i-1}(X)}] \tag{42}
\]
Recall that \( \psi(\lambda) = \log \mathbb{E}[A^\lambda] \) and \( \mathcal{F}_k = -\sum_{i=1}^{k} \log A_i \). Let \( t > 0 \), for \( i \geq 1, x > 0, \)
\[
\mathbb{P}(H_i(\omega) \geq x i) \leq \mathbb{P}(\max_{0 \leq k \leq i} \left[ t \mathcal{F}_k - \psi(t) k \right] \geq x i - \psi(t) i)
\leq \mathbb{P}(\max_{0 \leq k \leq i} e^{-t \mathcal{F}_k - \psi(t) k} \geq e^{(x t - \psi(t)) i})
\leq e^{-(x t - \psi(t)) i}, \tag{43}
\]
where the last inequality stems from Doob’s maximal inequality and the fact that \( e^{-t \mathcal{F}_j - \psi(t) j} \) is a martingale. Since \( x \geq \mathbb{E}[-\log A], I(x) = \sup_{t>0} \left\{ tx - \psi(t) \right\}, \) we have
\[
\mathbb{P}(H_i(\omega) \geq x i) \leq e^{-I(x) i}. \tag{44}
\]
Similarly, for any \( j \geq 1 \) and \( x > \mathbb{E}[-\log A] \).
\[
\mathbb{P}(H_j(\omega) \geq x j) \leq \mathbb{P}(\max_{0 \leq k \leq j} \left[ t \mathcal{F}_k - \psi(t) k \right] \geq x j - \psi(-t) j)
\leq \mathbb{P}(\max_{0 \leq k \leq j} e^{t \mathcal{F}_k - \psi(t) k} \geq e^{(x t - \psi(-t)) j})
\leq e^{-(x t - \psi(-t)) j}, \tag{45}
\]
which implies that
\[
\mathbb{P}(H_j(\omega) \geq x j) \leq e^{-I(-x) j}. \tag{46}
\]
Further, for \( 0 < x < \mathbb{E}[-\log A], \) one sees that by Cramér’s theorem,
\[
\mathbb{P}(H_j(\omega) \leq x j) \leq \mathbb{P}(X_1 + \cdots + X_j \leq x j) = \mathbb{P}(-X_1 - \cdots - X_j \geq -x j) \leq e^{-I(-x) j}. \tag{47}
\]
Take \( \eta > 0. \) In \( \tag{42}, \) we can replace \( H_i(\omega) \land H_{n-i-1}(X) \) by \( H_i(\omega) \land H_{n-i-1}(X) \land K \eta n \) with some \( K \geq 1 \) large enough. In fact,
\[
\mathbb{E}[e^{2\lambda H_i(\omega) \land H_{n-i-1}(X)}] \leq \mathbb{E}[e^{2\lambda H_i(\omega) \land H_{n-i-1}(X)}; H_i(\omega) \lor H_{n-i-1}(X) \leq K \eta n]
\leq \mathbb{E}[e^{\xi_{K}(t)}]
+ \mathbb{E}[e^{2\lambda H_i(\omega) \land H_{n-i-1}(X)}; H_i(\omega) \lor H_{n-i-1}(X) \geq K \eta n]. \tag{45}
\]

Observe that
\[ \Xi^+_K(i) \leq \mathbb{E}[e^{2\lambda H_i(-X)}; H_i(-X) \geq K\eta n] + \mathbb{E}[e^{2\lambda H_{n-1}(X)}; H_{n-1}(X) \geq K\eta n] \]
\[ =: \Xi_1 + \Xi_2 \]

Let us bound \( \Xi_1 \),
\[ \Xi_1 = \mathbb{E}\int_{-\infty}^{K\eta n} 2\lambda e^{2\lambda x} 1_{H_i(-X) \geq K\eta n} dx = \int_{-\infty}^{K\eta n} 2\lambda e^{2\lambda x} P(H_i(-X) \geq K\eta \vee x) dx \]
\[ = \int_{-\infty}^{K\eta n} 2\lambda e^{2\lambda x} P(H_i(-X) \geq K\eta n) + \int_{K\eta n}^{\infty} 2\lambda e^{2\lambda x} P(H_i(-X) \geq x) dx \]
\[ = e^{2\lambda K\eta n} P(H_i(-X) \geq K\eta n) + \int_{K}^{\infty} 2\lambda \eta n e^{2\lambda \eta n} P(H_i(-X) \geq t\eta n) dt \]

By applying (43), one sees that for any \( 0 \leq i \leq n-1 \) and \( \mu = 3 > 2\lambda \),
\[ \Xi_1 \leq e^{2\lambda K\eta n} e^{-\mu \eta n + \psi(\mu) i} + \int_{K}^{\infty} 2\lambda \eta n e^{2\lambda \eta n} e^{-\mu \eta n + \psi(\mu) i} dt \]
\[ \leq e^{-K\eta n + \psi(3)n} + 2\lambda \eta n \int_{K}^{\infty} \eta n e^{-\mu \eta n} dt \]
\[ \leq 3e^{-K\eta n + \psi(3)n}, \]
which is less than 1 when we choose \( K \) large enough. Similarly, we can show that for any \( i \leq n-1 \),
\[ \Xi_2 \leq 1, \]
for \( K \) large enough. Consequently, (42) becomes that
\[ \mathbb{E}\tilde{E}_0^{\omega} [\tilde{\tau}_1 \wedge \tilde{\tau}_n] \leq n 3n^3 + n^3 \max_{0 \leq i \leq n-1} \Xi^+_K(i). \quad (48) \]

It remains to bound \( \Xi^-_K(i) \). Take sufficiently small \( \varepsilon > 0 \) and let \( L = \lfloor \frac{1}{\varepsilon} \rfloor \). For any \( i \) such that \( l_1 \varepsilon n \leq i < (l_1 + 1) \varepsilon n \) and \( l_2 \varepsilon n \leq n - i - 1 < (l_2 + 1) \varepsilon n \) with \( 0 \leq l_1, l_2 \leq L \), we have
\[ \Xi^-_K(i) \leq \sum_{0 \leq k_1, k_2 \leq K} e^{2\lambda k_1 \wedge k_2 \eta n + 2\lambda \eta n} P(k_1 \eta n \leq H_i(-X) < (k_1 + 1) \eta n) P(k_2 \eta n \leq H_{n-1}(X) < (k_2 + 1) \eta n) \]
\[ \leq \sum_{0 \leq k_1, k_2 \leq K} e^{2\lambda k_1 \wedge k_2 \eta n + 2\lambda \eta n} P(H_i(-X) \geq k_1 \eta n) P(k_2 \eta n \leq H_{n-1}(X) < (k_2 + 1) \eta n). \]

By (44), we have
\[ P(H_i(-X) \geq k_1 \eta n) \leq e^{-I(x_1) i} \]
where \( x_1 \) is the point in \( \left[ \frac{k_1 \eta n}{(l_1 + 1) \varepsilon n}, \frac{k_1 \eta n}{l_1 \varepsilon n} \right] \) where \( I \) reaches the minimum in this interval. By large deviation estimates (46) (47), we have
\[ P(k_2 \eta n \leq H_{n-1}(X) < (k_2 + 1) \eta n) \leq e^{-I(x_2)(n-i)} \]
where \( x_2 \) is the point in \( \left[ \frac{k_2 \eta n}{(l_2 + 1) \varepsilon n}, \frac{(k_2 + 1) \eta n}{l_2 \varepsilon n} \right] \) where \( I \) reaches the minimum in this interval. Therefore,
\[ \Xi^-_K(i) \leq \sum_{0 \leq k_1, k_2 \leq K} e^{2\lambda k_1 \wedge k_2 \eta n + 2\lambda \eta n} e^{-I(x_1) l_1 \varepsilon n} e^{-I(x_2) l_2 \varepsilon n} \]

32
Taking maximum over all $l_1, l_2, k_1, k_2$ yields that
\[
\mathbb{E} \tilde{E}_0^{\omega} [\tilde{\tau}_{\lambda} \wedge \tilde{\tau}_n] \leq 3n^2 + n^2 K^2 \max_{l_1, l_2, k_1, k_2} \exp\{2\lambda k_1 \wedge k_2 \eta n + 2\lambda \eta n - I(x_1)l_1[\epsilon n] - I(-x_2)l_2[\epsilon n]\}.
\] (49)

Observe that
\[
2\lambda k_1 \wedge k_2 \eta n + 2\lambda \eta n - I(x_1)l_1[\epsilon n] - I(-x_2)l_2[\epsilon n] \\
\leq 2\lambda (x_1 l_1 + x_2 l_2)[\epsilon n] - I(x_1)l_1[\epsilon n] - I(-x_2)l_2[\epsilon n] + 3\lambda \eta n.
\]

Define
\[
L(\lambda) := \sup_{\mathcal{D}} \{x_1 z_1 \wedge x_2 z_2 \lambda - I(x_1)z_1 - I(-x_2)z_2\},
\]
where $\mathcal{D} := \{x_1, x_2, z_1, z_2 \geq 0, z_1 + z_2 \leq 1\}$.

By Lemma 8.1 in [1], one concludes that
\[
\limsup_{n \to \infty} \frac{\log \mathbb{E} \tilde{E}_0^{\omega} [\tilde{\tau}_{\lambda} \wedge \tilde{\tau}_n]}{n} \leq L(2\lambda) = \psi \left( \frac{1 + 2\lambda}{2} \right).
\]

\[
\square
\]

B Some observations on random walks on random trees

Proof of Lemma 10 As $\beta(x)$ is identically distributed under $\mathbb{P}$,
\[
\mathbb{E}_\rho \left( \sum_{|x|=n} 1_{\tau_x < \infty} \right) \mathbb{E}(\beta) = \mathbb{E} \left[ \sum_{|x|=n} P^{\omega, T}_\rho (\tau_x < \infty) \right] \mathbb{E}(\beta) = \mathbb{E} \left( \sum_{|x|=n} \mathbb{E}^T (P^{\omega, T}_\rho (\tau_x < \infty)) \mathbb{E}^T (\beta(x)) \right).
\]

Here we used the fact that $\mathbb{E}^T P^{\omega, T}_\rho (\tau_x < \infty)$ and $\mathbb{E}^T (\beta(x))$ are independent. Now $P^{\omega, T}_\rho (\tau_x < \infty)$ is an increasing function of $A_x$ since
\[
P^{\omega, T}_\rho (\tau_x < \infty) = P^{\omega, T}_\rho (\tau_x \leq \infty) \left( \sum_{k \geq 0} P^{\omega, T}_\rho (\tau_x \leq \infty < \min(\tau_x, \infty) \right) p(\tau_x, x)
\]
\[
= \frac{P^{\omega, T}_\rho (\tau_x \leq \infty)}{1 - P^{\omega, T}_\rho (\tau_x \leq \min(\tau_x, \infty))} A_x B_x.
\]

recall that $\beta(x)$ is also an increasing function of $A_x$, moreover, conditionally on $A_x$, $P^{\omega, T}_\rho (\tau_x < \infty)$ and $\beta(x)$ are independent, thus by FKG inequality,
\[
\mathbb{E}^T (P^{\omega, T}_\rho (\tau_x < \infty)) \beta(x)) = \mathbb{E}^T (\mathbb{E}^T (P^{\omega, T}_\rho (\tau_x < \infty)) \beta(x) | A_x))
\]
\[
= \mathbb{E}^T (\mathbb{E}^T (P^{\omega, T}_\rho (\tau_x < \infty) | A_x) \mathbb{E}^T (\beta(x) | A_x))
\]
\[
\geq \mathbb{E}^T (P^{\omega, T}_\rho (\tau_x < \infty)) \mathbb{E}^T (\beta(x)).
\]
Therefore,

\[
\mathbb{E}\left(\sum_{|x|=n} \mathbb{E}^T(P^{\omega,T}_\rho(\tau_x < \infty))\mathbb{E}^T(\beta(x))\right) \leq \mathbb{E}\left(\sum_{|x|=n} \mathbb{E}^T(P^{\omega,T}_\rho(\tau_x < \infty))\beta(x)\right) \\
= \mathbb{E}\left(\sum_{|x|=n} P^{\omega,T}_\rho(\tau_x < \infty))\beta(x)\right).
\]

For any GW tree and any trajectory on the tree, there is at most one regeneration time at the n-th generation, therefore,

\[
\sum_{|x|=n} \mathbb{1}_{\tau_x < \infty, \eta_k \neq x, \forall k \geq \tau_x} \leq 1.
\]

By taking expectation w.r.t. \(E^{\omega,T}_\rho\) and using the Markov property at \(\tau_x\),

\[
\sum_{|x|=n} P^{\omega,T}_\rho(\tau_x < \infty))\beta(x) \leq 1.
\]

Whence

\[
\mathbb{E}(\sum_{|x|=n} \mathbb{1}_{\tau_x < \infty})\mathbb{E}(\beta) \leq 1.
\]

By transient assumption it suffices to take \(c_{11} = \frac{1}{\mathbb{E}[\beta]} < \infty\).

\[\square\]

Proof of Lemma 11 and Corollary 3. Let \(T_i, i \geq 1\) be independent copies of GW tree with offspring distribution \((q)\), each endowed with independent environment \((\omega_x, x \in T_i)\). Let \(\rho^{(i)}\) be the root of \(T_i\). In such setting, \(\beta(\rho^{(i)}), i \geq 1\) are i.i.d. sequence with common distribution \(\beta\).

For each \(T_i\), take the left most infinite ray, denoted \(v_0^{(i)} = \rho^{(i)}, v_1^{(i)}, \ldots, v_n^{(i)}, \ldots\) Let \(\Omega(x) = \{y \neq x; x = y\}\) be the set of all brothers of \(x\). Fix some constant \(C\), define

\[
R_i = \inf\{n \geq 1; \exists z \in \Omega(v_n^{(i)}), \frac{1}{A_z}\beta(z) \leq C\}.
\]

By Equation (15),

\[
\frac{1}{\beta(v_{R_i-1}^{(i)})} \leq 1 + \frac{1}{A_{v_{R_i-1}^{(i)}}A_z\beta(z)} \leq 1 + \frac{C}{A_{v_{R_i-1}^{(i)}}}.
\]

Also \(R_i\) and \(\{A_{v_n^{(i)}}, n \geq 0\}\) are independent under \(Q\). By iteration,

\[
\frac{1}{\beta(\rho^{(i)})} \leq 1 + \frac{1}{A_{v_0^{(i)}}A_{v_1^{(i)}}\beta(v_1^{(i)})} \leq 1 + \frac{1}{A_{v_0^{(i)}}A_{v_1^{(i)}}}(1 + \frac{1}{A_{v_1^{(i)}}A_{v_2^{(i)}}\beta(v_2^{(i)})}) \\
\leq \ldots \\
\leq 1 + \sum_{k=1}^{R_i-1} \frac{1}{A_{v_0^{(i)}}A_{v_k^{(i)}}} \prod_{j=1}^{k-1} A_{v_j^{(i)}}^{-2} + \frac{C}{A_{v_0^{(i)}}} \prod_{l=1}^{R_i-1} A_{v_l^{(i)}}^{-2}.
\]

For any \(n \geq 0\), denote

\[
C(n) = 1 + \sum_{k=1}^{n} \frac{1}{A_{v_0^{(i)}}A_{v_k^{(i)}}} \prod_{j=1}^{k-1} A_{v_j^{(i)}}^{-2} + \frac{C}{A_{v_0^{(i)}}} \prod_{l=1}^{n} A_{v_l^{(i)}}^{-2}.
\]

(50)
Thus $\frac{1}{\beta(q)} \leq C(R_i - 1)$, note also that, since $\xi_2 = E(A^{-2}) = 1 + \frac{3}{2c} + \frac{3}{c}$, $E(C(n)) \leq c_34\xi_2^{n+1}$. Therefore, for any $K \geq 1$,

$$\frac{1}{\sum_{i=1}^{K} \beta(q)} \leq C(\min_{1 \leq i \leq K} R_i - 1).$$

Taking expectation under $P$ yields (as $R_i$ i.i.d. let $R$ be a r.v. with the common distribution)

$$E(\frac{1}{\sum_{i=1}^{K} \beta(q)}) \leq E(E(\min_{1 \leq i \leq K} R_i - 1 \mid R_i; 1 \leq i \leq K))$$

$$\leq c_{34}E(\xi_2^{\min_{1 \leq i \leq K} R_i}) \leq c_{34} \sum_{n=0}^{\infty} \xi_2^{n+1}P(R \geq n + 1)^K$$

$$\leq c_{34} \sum_{n=0}^{\infty} \xi_2^{n+1}E(\delta_{C}^{\sum_{k=0}^{n-1}(d(v_k) - 2)})^K$$

where $\delta_C = P(\delta_{C} > C)$. Let $f(s) = \sum_{k \geq 1} q_k s^k$, as $f(s)/s \downarrow q_1$ as $s \downarrow 0$, for any $\varepsilon > 0$, we can take $C$ large enough to ensure $\frac{f(\delta_{C})}{\delta_{C}} \leq q_1(1 + \varepsilon)$, thus

$$E(\frac{1}{\sum_{i=1}^{K} \beta(q)}) \leq c_{34} \sum_{n=0}^{\infty} \xi_2^{n+1}(\frac{f(\delta_{C})}{\delta_{C}})^nK \leq c_{34} \sum_{n=0}^{\infty} \xi_2^{n+1}(q_1(1 + \varepsilon))^{nK}.$$

Now take $\varepsilon$ such that $q_1(1 + \varepsilon) < 1$, then take $K$ large enough such that $\xi_2(q_1(1 + \varepsilon))^K < 1$ leads to

$$E(\frac{1}{\sum_{i=1}^{K} \beta(q)}) < c_{12} < \infty.$$

Similarly, the following also holds

$$E(\frac{1}{\sum_{i=1}^{K} A_{\beta(q)} \beta(q)}) < c_{12} < \infty.$$

In particular, if $q_1 \xi_2 < 1$, we can take $K = 1$ in $\xi_2(q_1(1 + \varepsilon))^K < 1$. Further, it follows from (50) and Cauchy-Schwartz inequality that

$$C(n)^2 \leq (n + 2) \left(1 + \sum_{k=1}^{n} \frac{1}{A_{v_{0}^{(i)} v_{0}^{(i)}}} \prod_{j=1}^{k-1} A_{v_{j}^{(i)}}^{-1} \prod_{l=1}^{n} A_{v_{l}^{(i)} v_{l}^{(i)}}^{-1} \right).$$

Thus,

$$E[C^2(n)] \leq c_{35}(n + 2)\xi_4^{n+1}.$$

As soon as $\xi_4 < \infty$, the previous argument works again to conclude that for $K$ large enough,

$$E(\frac{1}{\sum_{i=1}^{K} \beta(q)}) + E(\frac{1}{\sum_{i=1}^{K_0} A_{\beta(q)} \beta(q)}) < c_{13} < \infty.$$

**Acknowledgments:** We would like to thank an anonymous referee for carefully reading the paper and providing corrections.
References

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

[2] Omer Angel, Nicholas Crawford, and Gady Kozma. Localization for linearly edge reinforced random walks. *Duke Mathematical Journal*, 163(5):889–921, 2014.

[3] Anne-Laure Basdevant, Arvind Singh, et al. Continuous-time vertex reinforced jump processes on Galton–Watson trees. *The Annals of Applied Probability*, 22(4):1728–1743, 2012.

[4] John D Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, pages 25–37, 1977.

[5] Andrea Collevecchio. Limit theorems for vertex-reinforced jump processes on regular trees. *Electron. J. Probab*, 14(66):1936–1962, 2009.

[6] Burgess Davis and Stanislav Volkov. Continuous time vertex-reinforced jump processes. *Probability theory and related fields*, 123(2):281–300, 2002.

[7] Burgess Davis and Stanislav Volkov. Vertex-reinforced jump processes on trees and finite graphs. *Probability theory and related fields*, 128(1):42–62, 2004.

[8] Margherita Disertori and Tom Spencer. Anderson localization for a supersymmetric sigma model. *Communications in Mathematical Physics*, 300(3):659–671, 2010.

[9] Margherita Disertori, Tom Spencer, and Martin R Zirnbauer. Quasi-diffusion in a 3D supersymmetric hyperbolic sigma model. *Communications in Mathematical Physics*, 300(2):435–486, 2010.

[10] Yueyun Hu and Zhan Shi. A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. *Probability theory and related fields*, 138(3-4):521–549, 2007.

[11] Yueyun Hu, Zhan Shi, et al. Slow movement of random walk in random environment on a regular tree. *The Annals of Probability*, 35(5):1978–1997, 2007.

[12] Russell Lyons. A simple path to Biggins’ martingale convergence for branching random walk. In *Classical and modern branching processes*, pages 217–221. Springer, 1997.

[13] Russell Lyons and Robin Pemantle. Random walk in a random environment and first-passage percolation on trees. *The Annals of Probability*, pages 125–136, 1992.

[14] Russell Lyons and Yuval Peres. Probability on trees and networks. Book in progress, 2005.

[15] C. Sabot, P. Tarrès, and X. Zeng. The Vertex Reinforced Jump Process and a Random Schrödinger operator on finite graphs. *ArXiv e-prints*, July 2015.

[16] Christophe Sabot and Pierre Tarres. Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. *arXiv preprint arXiv:1111.3991*, 2011.

[17] Christophe Sabot and Pierre Tarres. Ray-Knight Theorem: a short proof. *arXiv preprint arXiv:1311.6622*, 2013.