Differentiability of the speed of biased random walks on Galton-Watson trees

Adam Bowditch and Yuki Tokushige

National University of Singapore,
Singapore 119077,
Singapore.
E-mail address: `matamb@nus.edu.sg`

Kyoto University, RIMS,
Kyoto 606-8502,
Japan.
E-mail address: `tokusige@kurims.kyoto-u.ac.jp`

Abstract. We prove that the speed of a λ-biased random walk on a supercritical Galton-Watson tree is differentiable for λ such that the walk is ballistic and obeys a central limit theorem, and give an expression of the derivative using a certain 2-dimensional Gaussian random variable. The proof heavily uses the renewal structure of Galton-Watson trees that was introduced in Lyons et al. (1996).

1. Introduction

In this paper, we investigate the speed of biased random walks on supercritical Galton-Watson trees. Specifically, we show that the speed is differentiable within a certain range of bias and obtain an expression for the derivative in terms of the covariance of a 2-dimensional Gaussian random variable.

Random walks on GW-trees are a natural setting for studying trapping phenomena as dead-ends, caused by leaves in the trees, trap the walk. Even without leaves, the randomness in the environment slows the walk and several properties that seem obvious turn out to be non-trivial and interesting problems. These models can be used to approach related problems concerning biased random walks on percolation clusters (as studied in Fribergh and Hammond, 2014) and random walk in random environment (see for example Sznitman, 2000) which experience similar phenomena. For a recent review of trapping phenomena we direct the reader to Ben Arous and Černý (2006), Ben Arous and Fribergh (2016) and Fontes and Mathieu (2014)
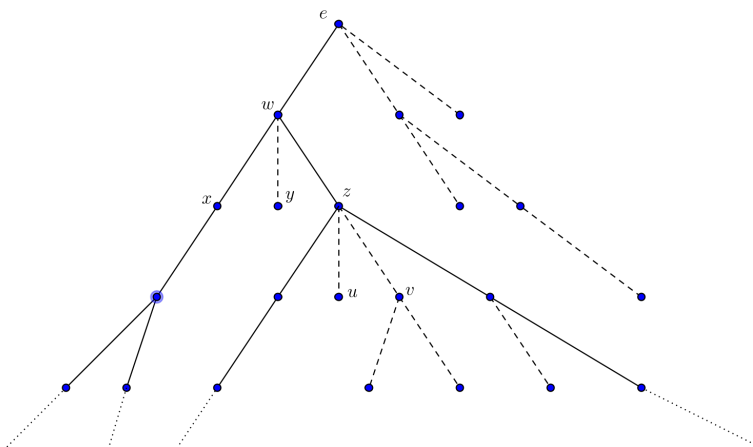

FIGURE 1.1. A sample section of a supercritical GW-tree condi-
tioned to survive $\mathbf{T}$ with solid lines representing the backbone $\mathbf{T}_g$
and dashed lines representing the traps. Here, the root e is the
parent of w (i.e. $e = \pi(w)$) which has children x, y, z where x, z
are on the backbone and y is a bud in the only trap rooted at w.
Similarly, u, v are two of the children of z, both of which are buds
of individual traps rooted at z.

which detail the history of trapping models including their motivation via spin-
glasses and cover recent developments in a range of models of random walks on
underlying graphs including supercritical GW-trees.

We now briefly describe the supercritical GW-tree conditioned on survival via
the Harris decomposition; for more detail see Athreya and Ney (2004); Janson
(2012). Let $\{p_k\}_{k \geq 0}$ denote the offspring distribution of a GW-process W_n with a
single progenitor, mean $\mu > 1$ and probability generating function f. The process
W_n gives rise to a random tree $\mathbf{T}_f$ where individuals are represented by vertices
and edges connect individuals with their offspring. Let q denote the extinction
probability of W_n which is strictly less than 1 since $\mu > 1$ and non-zero only when
$p_0 > 0$. In this case we then define

$$g(s) := \frac{f((1-q)s + q) - q}{1 - q} \qquad \text{and} \qquad h(s) := \frac{f(qs)}{q}$$

which are generating functions of a GW-process without deaths and a subcritical
GW-process respectively (cf. Chapter I.12 of Athreya and Ney, 2004). An f-GW-
tree conditioned on nonextinction $\mathbf{T}$ can be constructed by first generating a g-
GW-tree $\mathbf{T}_g$ and then, to each vertex x of $\mathbf{T}_g$, appending a random number of
independent h-GW-trees (see Figure 1.1). We refer to $\mathbf{T}_g$ as the backbone of $\mathbf{T}$,
the finite trees appended to $\mathbf{T}_g$ as the traps and the vertices in the first generation
of the traps as the buds.

We now introduce the biased random walk on a fixed tree $\mathscr{T}$. We denote by
$e(\mathscr{T})$ the root, which is the vertex representing the unique progenitor. For $x \in \mathscr{T}$,

let $\pi(x)$ denote the parent of x and $\nu(x)$ the number of children of x. A λ-biased random walk on $\mathscr{T}$ is a random walk $(Z_n)_{n\geq 0}$ on the vertices of $\mathscr{T}$ started from $e(\mathscr{T})$ with transition probabilities

$$P_\lambda^{\mathscr{T}}(Z_{n+1} = y | Z_n = x) = A_\lambda(x,y) := \begin{cases} \frac{\lambda}{\lambda+\nu(x)}, & \text{if } y = \pi(x), \\ \frac{1}{\lambda+\nu(x)}, & \text{if } x = \pi(y) \neq e(\mathscr{T}), \\ \frac{1}{\nu(x)}, & \text{if } x = \pi(y) = e(\mathscr{T}), \\ 0, & \text{otherwise.} \end{cases}$$

We use $P_\lambda(\cdot) := \int P_\lambda^{\mathbf{T}}(\cdot)\mathbb{P}(\mathrm{d}\mathbf{T})$ for the *annealed law* obtained by averaging the *quenched law* $P_\lambda^{\mathbf{T}}$ over the law $\mathbb{P}$ on f-GW-trees conditioned to survive. For $x \in \mathbf{T}$, let $d(x)$ denote the distance between x and the root of the tree and write $\lambda_c := f'(q)$ where we note that $\lambda_c = 0$ when $p_0 = 0$.

The behaviour of λ-biased random walks on the GW-tree $\mathbf{T}$ have been extensively studied since Lyons et al. (1996) showed that if $\lambda \in (\lambda_c, \mu)$ then the walk is *ballistic*; that is, $d(Z_n)n^{-1}$ converges P_λ-a.s. to a deterministic constant $\upsilon_\lambda > 0$ called the *speed* of the walk. When $\lambda > \mu$ the walk is recurrent and $d(Z_n)n^{-1}$ converges P_λ-a.s. to 0. When λ is small and $p_0 > 0$, the walk is transient but slowed by having to make long sequences of movements against the drift in order to escape the traps; in particular, if $\lambda \leq \lambda_c$ then the slowing affect is strong enough to cause $d(Z_n)n^{-1}$ to converge P_λ-a.s. to 0. This regime has been studied further in Ben Arous et al. (2012) and Bowditch (2018a) where polynomial scaling results are shown.

The aim of this paper is to study how the value of υ_λ depends on the parameter of bias λ; specifically, our main result is the following.

Theorem 1.1. *Suppose that there exists $\beta > 1$ such that $\sum_{k=1}^\infty p_k \beta^k < \infty$. Then, the function $\lambda \mapsto \upsilon_\lambda$ is differentiable on $(\lambda_c^{1/2}, \mu)$. Moreover, the derivative of the speed υ_λ' can be expressed as the covariance of a two dimensional Gaussian random variable (X, Y). Namely, we have that $\upsilon_\lambda' = E_\lambda[XY]$.*

We remark here that $0 \leq \lambda_c \leq \lambda_c^{1/2} < 1$ since $0 \leq \lambda_c < 1$ and note that the covariance matrix of (X, Y) is given in (3.20).

In the unpublished note Aïdékon (2013), the differentiability of the function $\lambda \mapsto \upsilon_\lambda$ was shown for $0 < \lambda < 1$ in the case $p_0 = 0$, and an expression of the derivative was given which is based on the description of invariant measures for the environment seen from the particle obtained in Aïdékon (2014).

A fluctuation-dissipation theorem FDT (see Dembo and Deuschel, 2010; Kubo, 1966) suggests that the internal fluctuations of a system at equilibrium should be related to the response of the system to an external disturbance. In the context of a random walk, this would suggest that the fluctuations of the walk should be related to the response of imposing a drift. A widely held conjecture is that an FDT should hold in many random walk models (e.g. Gantert et al., 2012; Lebowitz and Rost, 1994); however, it has been shown in Cugliandolo and Kurchan (1993) that this is violated by several mean-field spin glass models at low temperature due to slow dynamics and aging. This is of particular interest due to the connections between spin-glasses and models of random walks in random trapping environments. Some progress towards proving an FDT for a random walk on a supercritical GW-tree without leaves was made in Ben Arous et al. (2013) where it was shown that the diffusivity is equal to the mobility (the derivative of the speed with respect to

the exterior force $\alpha_\lambda = \log(\mu/\lambda))$ at the diffusive point $\lambda = \mu$. Understanding the relation between the diffusivity and the mobility for λ in the ballistic regime remains open.

It has been shown in Bowditch (2018b) and Peres and Zeitouni (2008) that, under the conditions of Theorem 1.1, there exists a constant $\varsigma \in (0, \infty)$ such that, for $\mathbb{P}$-a.e. **T**,

$$B_t^n := \frac{d(Z_{nt}) - nt\upsilon_\lambda}{\varsigma\sqrt{n}}$$

converges in $P_\lambda^{\mathbf{T}}$-distribution to a Brownian motion. In particular, the range of bias $(\lambda_c^{1/2}, \mu)$ is precisely the range in which the walk is ballistic and a central limit theorem holds. We expect that the differentiability should extend to (λ_c, μ) however, our proof relies heavily on second moment bounds of regeneration times which only hold in the smaller range of bias.

The key ingredients of the proof are the *renewal structure*, the *discrete Girsanov formula* and suitable moment bounds on excursion times of random walks in GW-trees.

The renewal structure allows paths of a random walk to be decomposed into *i.i.d.* components. This technique is frequently used to analyse random walks in random environments as well as various other models in probability and statistical mechanics. Lyons et al. (1996) constructed the renewal structure for supercritical GW-trees, which we will heavily utilise in this paper. See Ben Arous et al. (2014); Berger et al. (2019); Dembo et al. (2002) for applications of this method to the analysis of the speed of random walks in random environments. In particular, we refer to the paper Berger et al. (2019), where the authors study the speed of biased random walks on a random conductance model, since our strategy resembles theirs. See Mathieu (2015) also for a study of a similar problem in the context of random walks on word-hyperbolic groups.

We now describe the discrete Girsanov formula which allows us to relate the walk for different values of the bias. Let $\mathscr{T}$ be a rooted infinite tree and $\big(\mathcal{F}_n(\mathscr{T})\big)_{n \geq 0}$ be the filtration on the probability space $(\tilde{\Omega}(\mathscr{T}), \mathcal{F}(\mathscr{T}), P_\lambda^{\mathscr{T}})$ generated by the λ-biased random walk (Z_n) on $\mathscr{T}$. Then for an $\big(\mathcal{F}_n(\mathscr{T})\big)$-stopping time S, an $\mathcal{F}_S(\mathscr{T})$-measurable function $F : \tilde{\Omega}(\mathscr{T}) \to \mathbb{R}$ and $h \geq -\lambda$, we have that

$$E_{\lambda+h}^{\mathscr{T}}\big[F((Z_k)_{k\geq 0})\big] = E_\lambda^{\mathscr{T}}\left[F((Z_k)_{k\geq 0})\prod_{i=1}^S \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right]. \qquad (1.1)$$

We remark here that regeneration times are not stopping times, thus the formula (1.1) does not apply directly to them. Moreover, we will mostly work with the annealed measure conditioned on *non-backtracking* in order to avoid bad behaviour of the first regeneration time. (See Remark 2.7 for details.) The presence of the non-backtracking condition is also an obstacle to apply the Girsanov formula. We will solve this problem using Lemma 3.3.

In order to study the change in υ_λ as we vary the bias λ, we require control on the walk that is uniform in the bias. Specifically, due to the regeneration structure, it will suffice to control the variation of the walk within a single regeneration block. To this end, an important role is played by Proposition 2.5, which gives a moment estimate of regeneration times that is uniform in the bias. Its proof is the main

technical contribution of this paper and Sections 4, 5 and 6 are entirely devoted to the fairly intricate arguments involved in it.

The organisation of this paper is as follows; in Section 2, we first introduce several basic facts on the renewal structure of GW-trees. In Section 3, we will prove Theorem 1.1 using the formula (1.1). We defer the more technical aspects concerning moments of regeneration times to Sections 4, 5 and 6. Specifically, in Section 4 we show that the uniform moment estimates for regenerations times hold for GW-trees without leaves, in Section 5 we prove a moment bound on the generation sizes of GW-trees and, finally, in Section 6 we combine these estimates to prove that the uniform moment estimates for regenerations times extends to the case with leaves.

2. Renewal structure of Galton-Watson trees

In this section, we introduce regeneration times and state their moment estimates, which will be very important for this study.

Definition 2.1. For a rooted tree $\mathscr{T}$ and $x \in \mathscr{T}$, define $P_{\lambda,x}^{\mathscr{T}}(\cdot) := P_{\lambda}^{\mathscr{T}}(\cdot | Z_0 = x)$. (Thus, $P_{\lambda}^{\mathscr{T}} = P_{\lambda,e(\mathscr{T})}^{\mathscr{T}}$.) We will denote the expectation with respect to P_{λ} (resp. $P_{\lambda}^{\mathscr{T}}$) by E_{λ} (resp. $E_{\lambda}^{\mathscr{T}}$).

Definition 2.2. Let $(Z_n)_{n \geq 0}$ be the λ-biased random walk on a rooted tree $\mathscr{T}$.

1: A time $n \in \mathbb{N}$ is called a regeneration time if $d(Z_n) > d(Z_k)$ for all $k < n$ and $d(Z_l) > d(Z_{n-1})$ for all $l > n$.

2: For $x \in \mathscr{T}$, define the first return time σ_x by $\sigma_x := \inf\{n \geq 1 ; Z_n = x\}$.

Remark 2.3. Regeneration times defined above are called *level-regeneration times* in Dembo et al. (2002), and are different from what are defined in Lyons et al. (1996).

Definition 2.4. Let $\mathscr{T}$ be a rooted tree.

1: For $x \in \mathscr{T}$, define $\mathscr{T}(x)$ as the subtree of $\mathscr{T}$ which consists of x and its descendants. The vertex x is naturally regarded as the root of $\mathscr{T}(x)$.

2: We will denote by $\mathscr{T}^*$ a new tree obtained by adding to the graph $\mathscr{T}$ an edge connecting $e(\mathscr{T})$ and a new vertex $e^*(\mathscr{T})$. The vertex $e^*(\mathscr{T})$ is considered as the root of $\mathscr{T}^*$ and the parent of $e(\mathscr{T})$. We often write e and e^* for $e(\mathscr{T})$ and $e^*(\mathscr{T})$ when the tree is clear from context.

The usefulness of renewal structure and regeneration times is that they provide a way to decompose sample paths of random walks into *i.i.d.* pieces. When we deal with random walks on graphs carrying good renewal structures, approximations using regeneration times often enable us to reduce the analysis of the statistical behaviour of random walks to that of *i.i.d.* random variables.

We note that a different sequence of regeneration times (called *super-regeneration times*) have been introduced in Ben Arous et al. (2012) which decouple the event of a regeneration from the structure of the tree. These are particularly useful in decomposing the walk; however, this definition of regeneration times is only suitable when $\lambda < 1$ because it relies on comparison with a biased random walk on $\mathbb{Z}$ with this bias.

An important property is that, by Lemma 3.3 and Proposition 3.4 of Lyons et al. (1996), for $\lambda \in (0, \mu)$ there exist, P_{λ}-a.s., infinitely many regeneration times

$0 =: \tau_0 < \tau_1 < \tau_2 < ....$ and the sequences $\{(\tau_{i+1}-\tau_i, d(Z_{\tau_{i+1}})-d(Z_{\tau_i})\}_{i\geq 0}$ are $i.i.d.$ random vectors under P_λ. A useful fact is that the law of $(\tau_2 - \tau_1, d(Z_{\tau_2}) - d(Z_{\tau_1}))$ under the probability measure P_λ is identical to the law of $(\tau_1, d(Z_{\tau_1}))$ under the probability measure $P_\lambda^{\mathtt{NB}}$, where

$$P_\lambda^{\mathtt{NB}}(A) := \int \mathbb{P}(d\mathbf{T}) P_{\lambda,e(\mathbf{T})}^{\mathbf{T}^*}(A \cap \sigma_{e^*(\mathbf{T})} = \infty) \cdot \left(\int \mathbb{P}(d\mathbf{T}) P_{\lambda,e(\mathbf{T})}^{\mathbf{T}^*}(\sigma_{e^*(\mathbf{T})} = \infty) \right)^{-1}.$$

Therefore, under $P_\lambda^{\mathtt{NB}}$, the distribution of $(\tau_2 - \tau_1, d(Z_{\tau_2}) - d(Z_{\tau_1}))$ coincides with that of $(\tau_1, d(Z_{\tau_1}))$. We will denote by $E_\lambda^{\mathtt{NB}}$ the expectation with respect to $P_\lambda^{\mathtt{NB}}$.

The following moment estimate of regeneration times which is uniform in λ will play an important role in this paper. We note that $\lambda_c < 1$ thus the condition $a < 1$ is to ensure that $\log(\lambda_c)/\log(a) > 0$. Since the proof is quite technical, we postpone it to Sections 4, 5 and 6.

Proposition 2.5. *Suppose $[a,b] \subset (0,\mu)$ with $a < 1$ and that there exists $\beta > 1$ such that $\sum_{k=1}^\infty p_k \beta^k < \infty$. Then, for any $\alpha < \log(\lambda_c)/\log(a)$ we have*

$$\sup_{\lambda\in[a,b]} E_\lambda^{\mathtt{NB}}[\tau_1^\alpha] = \sup_{\lambda\in[a,b]} E_\lambda^{\mathtt{NB}}[(\tau_2 - \tau_1)^\alpha] = \sup_{\lambda\in[a,b]} E_\lambda[(\tau_2 - \tau_1)^\alpha] < \infty.$$

If $a > \lambda_c^{1/2}$ then $\log(\lambda_c)/\log(a) > 2$ therefore we immediately have the following corollary.

Corollary 2.6. *Suppose $[a,b] \subset (\lambda_c^{1/2}, \mu)$ and that there exists $\beta > 1$ such that $\sum_{k=1}^\infty p_k \beta^k < \infty$. Then, for some $\varepsilon > 0$ we have*

$$\sup_{\lambda\in[a,b]} E_\lambda^{\mathtt{NB}}[\tau_1^{2+\varepsilon}] = \sup_{\lambda\in[a,b]} E_\lambda^{\mathtt{NB}}[(\tau_2 - \tau_1)^{2+\varepsilon}] = \sup_{\lambda\in[a,b]} E_\lambda[(\tau_2 - \tau_1)^{2+\varepsilon}] < \infty.$$

Remark 2.7. In general, τ_1 does not satisfy these good moment estimates under the law P_λ which is one of the reasons that we use instead the law $P_\lambda^{\mathtt{NB}}$. This problem arises from the number of excursions of the random walk from the root to itself until the walk escapes. Denote by RE the number of visits of the walk to the root, we have that $\tau_1 \geq \mathsf{RE}$ P_λ-$a.s.$ Moreover, RE under $P_\lambda^{\mathbf{T}}$ is distributed as the geometric random variable whose termination probability is the quenched escape probability $P_\lambda^{\mathbf{T}}(\sigma_{e^*(\mathbf{T})} = \infty)$.

Suppose that $p_1 > 0$ and $\lambda > 1$. We denote by $R^\lambda(\mathbf{T})$ the effective resistance from $e(\mathbf{T})$ to ∞ of $\mathbf{T}$ where $\mathbf{T}$ is regarded as the electric network corresponding to the λ-biased random walk. Let us consider the event where every individual up to n^{th} generation gives birth to only one child, which occurs with probability p_1^n. On this event, $R^\lambda(\mathbf{T})$ is of order λ^n. This observation implies that

$$\mathbb{P}(R^\lambda(\mathbf{T}) > n) \geq c n^{\log p_1 / \log \lambda}$$

for some constant $c > 0$. (Note that $\log p_1 < 0$ since $0 < p_1 < 1$.) In the light of a well-known fact in the theory of electric networks and reversible random walks (see Theorem 2.11 of Barlow, 2017 for instance), the above estimate implies that RE (therefore τ_1 also) does not even satisfy a finite first moment in general.

We also need the following estimates in what follows. Their proofs will also be given in Sections 4, 5 and 6.

Proposition 2.8. *Suppose that there exists $\beta > 1$ such that $\sum_{k=1}^{\infty} p_k \beta^k < \infty$. Then for any $\lambda \in (\lambda_c, \mu)$, we have that*

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\sigma_{e^*(\mathbf{T})} \mathbf{1}_{\{\sigma_{e^*(\mathbf{T})} < \infty\}} \right]\right] < \infty.$$

Proposition 2.9. *The function*

$$\lambda \mapsto \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\sigma_{e^*(\mathbf{T})} = \infty \right)\right]$$

is continuous on (λ_c, μ).

3. Expressions of derivatives of the speed

In this section we prove Theorem 1.1 assuming Proposition 2.5. The following result gives the finite approximation of the derivative. Notice that we do not use expectations $E_\lambda[d(Z_n)]$ and $E_{\lambda+h}[d(Z_n)]$, but $E_\lambda^{\mathrm{NB}}[d(Z_n)]$ and $E_{\lambda+h}^{\mathrm{NB}}[d(Z_n)]$, to approximate the derivative, which is a non-negligible difference from the approach of Berger et al. (2019). This is necessary because of the lack of good moment estimates of the first regeneration time τ_1, which arises from a special role played by the root. See Remark 2.7 for details.

Proposition 3.1. *Suppose $\lambda \in (\lambda_c^{1/2}, \mu)$ and that there exists $\beta > 1$ such that $\sum_{k=1}^{\infty} p_k \beta^k < \infty$. Let h tend to 0 and n tend to ∞ in such a way that $h^2 n$ tends to 1 (i.e. $hn \sim n^{1/2}$). Then*

$$\frac{\upsilon_{\lambda+h} - \upsilon_\lambda}{h} - \frac{E_{\lambda+h}^{\mathrm{NB}}[d(Z_n)] - E_\lambda^{\mathrm{NB}}[d(Z_n)]}{hn}$$

tends to 0.

Proof: Define $\eta_n := \inf\{k : \tau_k \geq n\}$ for $n \in \mathbb{N}$, then η_n is a stopping time with respect to the filtration generated by random variables τ_1, and $\{\tau_{i+1} - \tau_i\}_{i \geq 1}$. By the definition of η_n, we have

$$n \leq \tau_{\eta_n} \leq n + \max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i), \ P_\lambda^{\mathrm{NB}}\text{-a.s.} \tag{3.1}$$

Combining this with Wald's identity we then have

$$n \leq E_\lambda^{\mathrm{NB}}[\tau_{\eta_n}] = E_\lambda^{\mathrm{NB}}[\eta_n] E_\lambda^{\mathrm{NB}}[\tau_1] \leq n + E_\lambda^{\mathrm{NB}}\left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right]. \tag{3.2}$$

Using that $0 \leq |d(Z_n) - d(Z_{\tau_{\eta_n}})| \leq \tau_{\eta_n} - n$, (3.1) then implies that

$$|E_\lambda^{\mathrm{NB}}[d(Z_n)] - E_\lambda^{\mathrm{NB}}[d(Z_{\tau_{\eta_n}})]| \leq E_\lambda^{\mathrm{NB}}\left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right].$$

Wald's identity then gives

$$E_\lambda^{\mathrm{NB}}[d(Z_{\tau_{\eta_n}})] = E_\lambda^{\mathrm{NB}}[\eta_n] E_\lambda^{\mathrm{NB}}[d(Z_{\tau_1})].$$

Hence, we get

$$\left| E_\lambda^{\mathrm{NB}}[d(Z_n)] - n\upsilon_\lambda \right| \leq \left| E_\lambda^{\mathrm{NB}}[d(Z_n)] - E_\lambda^{\mathrm{NB}}[d(Z_{\tau_{\eta_n}})] \right| + \left| E_\lambda^{\mathrm{NB}}[d(Z_{\tau_{\eta_n}})] - n\upsilon_\lambda \right|$$

$$\leq E_\lambda^{\mathrm{NB}}\left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right] + \left| E_\lambda^{\mathrm{NB}}[\eta_n] E_\lambda^{\mathrm{NB}}[d(Z_{\tau_1})] - n\upsilon_\lambda \right| \quad (3.3)$$

By (6.4) in Lyons et al. (1996) we have that

$$v_\lambda = \frac{E_\lambda[d(Z_{\tau_2}) - d(Z_{\tau_1})]}{E_\lambda[\tau_2 - \tau_1]} = \frac{E_\lambda^{\mathrm{NB}}[d(Z_{\tau_1})]}{E_\lambda^{\mathrm{NB}}[\tau_1]}, \tag{3.4}$$

therefore, using (3.2) and that $v_\lambda \leq 1$, we have

$$\left| E_\lambda^{\mathrm{NB}}[\eta_n] E_\lambda^{\mathrm{NB}}[d(Z_{\tau_1})] - nv_\lambda \right| \leq \left| E_\lambda^{\mathrm{NB}}[\eta_n] E_\lambda^{\mathrm{NB}}[\tau_1] - n \right| \leq E_\lambda^{\mathrm{NB}} \left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right]. \tag{3.5}$$

By combining (3.3) and (3.5), we get

$$\left| E_\lambda^{\mathrm{NB}}[d(Z_n)] - nv_\lambda \right| \leq 2 E_\lambda^{\mathrm{NB}} \left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right]$$

In order to complete the proof of Proposition 3.1, it suffices to show that there exist constants $0 < t_\lambda < \min\{\mu - \lambda, \lambda - \lambda_c\}$, $0 < \kappa < 1/2$ and $c_\lambda > 0$ such that

$$E_{\lambda'}^{\mathrm{NB}} \left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right] \leq c_\lambda n^\kappa \tag{3.6}$$

for any $\lambda' \in (\lambda - t_\lambda, \lambda + t_\lambda)$. The estimate (3.6) can be proved as follows: for $\kappa > 0$, we have

$$E_{\lambda'}^{\mathrm{NB}} \left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right] \leq n^\kappa + \sum_{k \geq [n^\kappa]} P_{\lambda'}^{\mathrm{NB}} \left(\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \geq k \right).$$

By Corollary 2.6, there exists a constant $c_\lambda > 0$ such that for $\lambda' \in (\lambda - t_\lambda, \lambda + t_\lambda)$ and sufficiently large k, we have

$$P_{\lambda'}^{\mathrm{NB}} \left(\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \geq k \right) = 1 - \{1 - P_{\lambda'} (\tau_2 - \tau_1 \geq k)\}^n$$

$$\leq 1 - (1 - c_\lambda k^{-(2+\varepsilon)})^n \leq 2 c_\lambda n k^{-(2+\varepsilon)}.$$

Thus, for sufficiently large n, we get

$$E_{\lambda'}^{\mathrm{NB}} \left[\max_{0 \leq i \leq n} (\tau_{i+1} - \tau_i) \right] \leq n^\kappa + 4 c_\lambda n^{-\kappa(1+\varepsilon)+1}.$$

Since $\max\{\kappa, -\kappa(1 + \varepsilon) + 1\} \leq \frac{1}{2+\varepsilon}$, we obtain the estimate (3.6). Therefore, we have shown that

$$\frac{E_{\lambda+h}^{\mathrm{NB}}[d(Z_n) - nv_{\lambda+h}] - E_\lambda^{\mathrm{NB}}[d(Z_n) - nv_\lambda]}{hn}$$

tends to 0 when h tends to 0 and n tends to ∞ in such a way that $h^2 n$ tends to 1. □

By Proposition 3.1, in order to show the differentiability of the function $\lambda \mapsto v_\lambda$, it suffices to prove the existence of the limit

$$\lim_{h,n} \frac{E_{\lambda+h}^{\mathrm{NB}}[d(Z_n)] - E_\lambda^{\mathrm{NB}}[d(Z_n)]}{hn},$$

where h tends to 0 and n tends to ∞ in such a way that $h^2 n$ tends to 1. We will need the following estimates.

Lemma 3.2. *Suppose that $\lambda \in (\lambda_c^{1/2}, \mu)$ and that there exists $\beta > 1$ such that $\sum_{k=1}^{\infty} p_k \beta^k < \infty$. Then, we have*

$$\sup_n \frac{1}{n} E_\lambda^{\text{NB}}[(d(Z_n) - nv_\lambda)^2] < \infty \tag{3.7}$$

and

$$\sup_n \frac{1}{n} \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(d(Z_n) - nv_\lambda)^2 \mathbf{1}_{\{\sigma_{e^*}(\mathbf{T}) > n\}}\right]\right] < \infty. \tag{3.8}$$

Proof: We first prove (3.7). Using (3.1) and the arguments of Proposition 3.1 we have that

$$|d(Z_n) - nv_\lambda| \le |d(Z_{\tau_{\eta_n}}) - nv_\lambda| + \max_{0 \le i \le n} (\tau_{i+1} - \tau_i) \quad P_\lambda^{\text{NB}}\text{-a.s.},$$

and, P_λ^{NB}-a.s.,

$$d(Z_{\tau_{\eta_n}}) - nv_\lambda = \sum_{i=0}^{\eta_n - 1} \left(d(Z_{\tau_{i+1}}) - d(Z_{\tau_i}) - E_\lambda^{\text{NB}}[d(Z_{\tau_1})]\right) + \left(\eta_n \cdot E_\lambda^{\text{NB}}[d(Z_{\tau_1})] - nv_\lambda\right)$$

Hence, we get

$$E_\lambda^{\text{NB}}[(d(Z_n) - nv_\lambda)^2]$$

$$\le 4\Bigg\{ E_\lambda^{\text{NB}}\left[\left(\max_{0 \le i \le n}(\tau_{i+1} - \tau_i)\right)^2\right] + E_\lambda^{\text{NB}}\left[\left(\eta_n \cdot E_\lambda^{\text{NB}}[d(Z_{\tau_1})] - nv_\lambda\right)^2\right]$$

$$+ E_\lambda^{\text{NB}}\left[\left\{\sum_{i=0}^{\eta_n - 1}\left(d(Z_{\tau_{i+1}}) - d(Z_{\tau_i}) - E_\lambda^{\text{NB}}[d(Z_{\tau_1})]\right)\right\}^2\right]\Bigg\}.$$

Recall that $d(Z_{\tau_{i+1}}) - d(Z_{\tau_i})$ are *i.i.d.* then Wald's second identity implies

$$E_\lambda^{\text{NB}}\left[\left(\sum_{i=0}^{\eta_n - 1}\left(d(Z_{\tau_{i+1}}) - d(Z_{\tau_i}) - E_\lambda^{\text{NB}}[d(Z_{\tau_1})]\right)\right)^2\right]$$

$$= E_\lambda^{\text{NB}}\left[\left(d(Z_{\tau_2}) - d(Z_{\tau_1}) - E_\lambda^{\text{NB}}[d(Z_{\tau_1})]\right)^2\right] E_\lambda[\eta_n]$$

thus by Corollary 2.6 and the estimate (3.2) along with (3.6), we have

$$\sup_n \frac{1}{n} E_\lambda^{\text{NB}}\left[\left(\sum_{i=0}^{\eta_n - 1}\left(d(Z_{\tau_{i+1}}) - d(Z_{\tau_i}) - E_\lambda^{\text{NB}}[d(Z_{\tau_1})]\right)\right)^2\right] < \infty.$$

It is not difficult to see that Corollary 2.6 implies $n^{-1} E_\lambda^{\text{NB}}[(\max_{0 \le i \le n}(\tau_{i+1} - \tau_i))^2]$ is also bounded in n. Hence, we get the conclusion if we show

$$\sup_{n \ge 1} \frac{1}{n} E_\lambda^{\text{NB}}\left[\left(\eta_n \cdot E_\lambda^{\text{NB}}[d(Z_{\tau_1})] - nv_\lambda\right)^2\right] < \infty.$$

It is shown in Chapter 4 of Cox (1962) that

$$E_\lambda^{\text{NB}}[\eta_n^2] = E_\lambda^{\text{NB}}[\eta_n]^2 + O(n) = \frac{n^2}{E_\lambda^{\text{NB}}[\tau_1]^2} + O(n).$$

By using the formula (3.4) and the estimate (3.2), we get

$$E_\lambda^{\text{NB}}\left[\left(\eta_n \cdot E_\lambda^{\text{NB}}[d(Z_{\tau_1})] - nv_\lambda\right)^2\right]$$

$$= E_\lambda^{\text{NB}}[\eta_n]^2 E_\lambda^{\text{NB}}[d(Z_{\tau_1})]^2 - \frac{E_\lambda^{\text{NB}}[d(Z_{\tau_1})]^2}{E_\lambda^{\text{NB}}[\tau_1]^2} \cdot n^2 + O(n)$$

which is at most order n therefore this implies (3.7).

We next prove (3.8). In order to deduce (3.8) from (3.7), it suffice to show that

$$\sup_n \frac{1}{n} \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(d(Z_n) - nv_\lambda)^2 \mathbf{1}_{\{n < \sigma_{e^*(\mathbf{T})} < \infty\}} \right] \right] < \infty.$$

This immediately follows from Proposition 2.8 and an obvious bound $|d(Z_n) - nv_\lambda| \le (1 + v_\lambda)n$. □

Lemma 3.2 implies uniform integrability of the sequence $\{(d(Z_n) - nv_\lambda)/\sqrt{n}\}_{n \ge 1}$ under the conditioned annealed measure P_λ^{NB} when $\lambda \in (\lambda_c^{1/2}, \mu)$ and the offspring distribution $\{p_k\}_{k \ge 0}$ has finite exponential moment. On the other hand, by Corollary 2.6 and a standard argument in the renewal theory, we get that the sequence $\{(d(Z_n) - nv_\lambda)/\sqrt{n}\}_{n \ge 1}$ satisfies the annealed CLT under the same assumptions. Hence, we have

$$\lim_{n \to \infty} \frac{E_\lambda^{\text{NB}}[d(Z_n) - nv_\lambda]}{\sqrt{n}} = 0,$$

for any $\lambda \in (\lambda_c^{1/2}, \mu)$. Thus, in order to prove the differentiability of the speed, we only need to show the existence of the limit

$$\frac{1}{hn} E_{\lambda+h}^{\text{NB}}[d(Z_n) - nv_\lambda], \tag{3.9}$$

for any sequence h and n such that $h \to 0$ and $n \to \infty$ in such a way that $h^2 n \to 1$. To do so, we wish to relate the measures E_λ^{NB} and $E_{\lambda+h}^{\text{NB}}$ by using the Girsanov formula (1.1). However, the formula (1.1) does not directly apply to $E_{\lambda+h}^{\text{NB}}[d(Z_n) - nv_\lambda]$ because of the presence of the non-backtracking condition $\{\sigma_{e^*(\mathbf{T})} = \infty\}$. In order to overcome this problem, we need the following lemma.

Lemma 3.3. *Suppose that* $\lambda \in (\lambda_c^{1/2}, \mu)$ *and that there exists* $\beta > 1$ *such that* $\sum_{k=1}^\infty p_k \beta^k < \infty$. *Then, we have that*

$$\frac{1}{\sqrt{n}} \left(\mathbb{E}\left[E_{\lambda+h}^{\mathbf{T}^*}\left[(d(Z_n) - nv_\lambda) \mathbf{1}_{\{\sigma_{e^*(\mathbf{T})} = \infty\}} \right] \right] \right.$$

$$\left. - \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(d(Z_n) - nv_\lambda) \prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \mathbf{1}_{\{\sigma_{e^*(\mathbf{T})} = \infty\}} \right] \right] \right) \tag{3.10}$$

converges to 0 *as* h *tends to* 0 *and* n *tends to* ∞ *in such a way that* $h^2 n \to 1$.

Since the proof of Lemma 3.3 requires a careful analysis of the Girsanov weight, we will defer it until the end of the next subsection.

3.1. *The discrete Girsanov formula.* In this subsection, we will analyse the Girsanov weight $\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}$. By the Taylor expansion, there exists $s = s(x, y) \in [0, 1]$ such that

$$\log \frac{A_{\lambda+h}(x, y)}{A_\lambda(x, y)} = h B_\lambda(x, y) + \frac{h^2}{2} C_\lambda(x, y) + \frac{h^3}{6} D_{\lambda+sh}(x, y),$$

where

$$B_\lambda(x,y) = \frac{\mathrm{d}}{\mathrm{d}\lambda} \log A_\lambda(x,y) = \begin{cases} 0 & \text{when } x = e, \\ \frac{1}{\lambda} - \frac{1}{\lambda+\nu(x)} & \text{when } y = \pi(x), \\ -\frac{1}{\lambda+\nu(x)} & \text{when } x = \pi(y), \end{cases}$$

$$C_\lambda(x,y) = \frac{\mathrm{d}}{\mathrm{d}\lambda} B_\lambda(x,y) = \begin{cases} 0 & \text{when } x = e, \\ -\frac{1}{\lambda^2} + \frac{1}{(\lambda+\nu(x))^2} & \text{when } y = \pi(x), \\ \frac{1}{(\lambda+\nu(x))^2} & \text{when } x = \pi(y), \end{cases}$$

and

$$D_\lambda(x,y) = \frac{\mathrm{d}}{\mathrm{d}\lambda} C_\lambda(x,y) = \begin{cases} 0 & \text{when } x = e, \\ \frac{2}{\lambda^3} - \frac{2}{(\lambda+\nu(x))^3} & \text{when } y = \pi(x), \\ -\frac{2}{(\lambda+\nu(x))^3} & \text{when } x = \pi(y). \end{cases}$$

By using these expressions, we have

$$\prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} = \exp(hP_n - h^2 Q_n + R_{n,h}) \quad P_\lambda\text{-a.s.,} \tag{3.11}$$

where

$$P_n := \sum_{j=0}^{n-1} B_\lambda(Z_j, Z_{j+1}),$$

$$Q_n := \sum_{j=0}^{n-1} \frac{1}{2} B_\lambda^2(Z_j, Z_{j+1}),$$

$$R_{n,h} := \sum_{j=0}^{n-1} \left\{ h^2 \left(\frac{1}{2} B_\lambda^2(Z_j, Z_{j+1}) + \frac{1}{2} C_\lambda(Z_j, Z_{j+1}) \right) + \frac{h^3}{6} D_{\lambda+sh}(Z_j, Z_{j+1}) \right\}.$$

Since

$$|B_\lambda(x,y)| \leq \frac{1}{\lambda} + 1, \ |C_\lambda(x,y)| \leq \frac{1}{\lambda^2} + 1, \ |D_\lambda(x,y)| \leq \frac{2}{\lambda^3} + 2, \tag{3.12}$$

we get

$$1 = \sum_y A_{\lambda+h}(x,y) = \sum_y A_\lambda(x,y) \exp\left(hB_\lambda(x,y) + \frac{h^2}{2} C_\lambda(x,y) + \frac{h^3}{6} D_{\lambda+sh}(x,y) \right)$$

$$= \sum_y A_\lambda(x,y) \left(1 + hB_\lambda(x,y) + \frac{h^2}{2} B_\lambda^2(x,y) + \frac{h^2}{2} C_\lambda(x,y) + O(h^3) \right).$$

This implies that for any $x \in \mathbf{T}$,

$$\sum_y A_\lambda(x,y) B_\lambda(x,y) = 0, \quad P_\lambda\text{-a.s.,} \tag{3.13}$$

$$\sum_y A_\lambda(x,y) \left(B_\lambda^2(x,y) + C_\lambda(x,y) \right) = 0, \quad P_\lambda\text{-a.s.}$$

By using the Markov property and the equality (3.13), we obtain

$$E_\lambda^{\mathbf{T}}[B_\lambda(Z_j, Z_{j+1}) \mid Z_j] = \sum_y B(Z_j, y) A_\lambda(Z_j, y) = 0, \quad P_\lambda^{\mathbf{T}}\text{-a.s.}$$

This implies

$$E_\lambda^{\mathbf{T}}[B_\lambda(Z_j, Z_{j+1})] = E_\lambda[B_\lambda(Z_j, Z_{j+1})] = 0, \quad P_\lambda\text{-}a.s. \tag{3.14}$$

Similarly, we have

$$E_\lambda^{\mathbf{T}}[B_\lambda^2(Z_j, Z_{j+1}) + C_\lambda(Z_j, Z_{j+1})] = 0, \quad P_\lambda\text{-}a.s. \tag{3.15}$$

We now let h tend to 0 and n tend to ∞ in such a way that $h^2 n$ tends to 1. We show that the limits of hP_n and $h^2 Q_n$ are described by a CLT and a LLN respectively and the limit of $R_{n,h}$ is negligible.

1) The CLT for P_n: By the renewal structure of GW-trees, we know that the collection $\{\sum_{j=\tau_i}^{\tau_{i+1}-1} B_\lambda(Z_j, Z_{j+1})\}_{i\geq 1}$ are *i.i.d.* random variables under P_λ, and are distributed as $\sum_{i=0}^{\tau_1-1} B_\lambda(Z_j, Z_{j+1})$ under P_λ^{NB}. Recall that by (3.14) we have that

$$E_\lambda[P_n] = 0. \tag{3.16}$$

On the other hand, noticing that τ_1 is finite a.s. by the renewal structure we have the following LLN:

$$\lim_{n\to\infty} n^{-1} P_n = \frac{E_\lambda\left[\sum_{j=\tau_1}^{\tau_2-1} B_\lambda(B_j, B_{j+1})\right]}{E_\lambda[\tau_2 - \tau_1]} = \frac{E_\lambda^{\mathrm{NB}}\left[\sum_{j=0}^{\tau_1-1} B_\lambda(B_j, B_{j+1})\right]}{E_\lambda^{\mathrm{NB}}[\tau_1]} \quad P_\lambda\text{-}a.s. \tag{3.17}$$

Noticing that $n^{-1} P_n \leq 1 + \lambda^{-1}$ by (3.12), the dominated convergence theorem implies that the same convergence as (3.17) holds in $L^1(P_\lambda)$. This fact together with (3.16) implies that

$$E_\lambda\left[\sum_{j=\tau_1}^{\tau_2-1} B_\lambda(Z_j, Z_{j+1})\right] = E_\lambda^{\mathrm{NB}}\left[\sum_{j=0}^{\tau_1-1} B_\lambda(Z_j, Z_{j+1})\right] = 0. \tag{3.18}$$

By (3.12) and Corollary 2.6 we also have that

$$E_\lambda\left[\left(\sum_{j=\tau_1}^{\tau_2-1} B_\lambda(Z_j, Z_{j+1})\right)^2\right] \leq E_\lambda^{\mathrm{NB}}\left[\left(\sum_{j=0}^{\tau_1-1} B_\lambda(Z_j, Z_{j+1})\right)^2\right]$$

$$\leq \left(1 + \frac{1}{\lambda}\right)^2 E_\lambda^{\mathrm{NB}}\left[\tau_1^2\right] < \infty.$$

Moreover, we have that

$$n^{-1/2}\left(\sum_{j=0}^{\tau_{\eta_n}-1} B_\lambda(Z_j, Z_{j+1}) - P_n\right) \leq n^{-1/2}\left(1 + \frac{1}{\lambda}\right) \max_{0\leq i\leq n}(\tau_{i+1} - \tau_i),$$

which converges to 0 in probability by (3.6). It therefore follows that $n^{-1/2} P_n$ converges in distribution to a centred Gaussian.

2) The LLN for Q_n: Recalling that τ_1 is P_λ-a.s. finite and from (3.12) that $B_\lambda^2(Z_j, Z_{j+1})$ is bounded above, by the law of large numbers we know that, P_λ-a.s.

$$\lim_{n\to\infty} \frac{1}{n} Q_n = \frac{1}{2E_\lambda[\tau_2 - \tau_1]} E_\lambda\left[\sum_{j=\tau_1}^{\tau_2-1} B_\lambda^2(Z_j, Z_{j+1})\right]$$

$$= \frac{1}{2E_\lambda^{\mathrm{NB}}[\tau_1]} E_\lambda^{\mathrm{NB}} \left[\sum_{j=0}^{\tau_1 - 1} B_\lambda^2(Z_j, Z_{j+1}) \right].$$

3) The estimate for $R_{n,h}$: For some constant $c < \infty$, we have $n \leq ch^{-2}$. Using this and (3.12), we have

$$\left| \sum_{j=0}^{n-1} \frac{h^3}{6} D_{\lambda + sh}(Z_j, Z_{j+1}) \right| \leq \frac{ch}{3} \left(\frac{1}{\lambda^3} + 1 \right).$$

By (3.15) we have that

$$E_\lambda^{\mathbf{T}} \left[\sum_{j=0}^{n-1} \left(B_\lambda^2(Z_j, Z_{j+1}) + C_\lambda(Z_j, Z_{j+1}) \right) \right] = 0, \ P_\lambda\text{-}a.s.$$

hence, by using the similar argument to the above one, we see that

$$\lim_{n \to \infty} R_{n,h} = 0, \ P_\lambda\text{-}a.s.$$

Note also that $R_{n,h}$ satisfies the following uniform estimate for h sufficiently small.

$$|R_{n,h}| \leq h^2 n \left(\frac{1}{\lambda} + 1 + \frac{1}{2} \left(\frac{1}{\lambda^2} + 1 \right) \right) + \frac{ch}{3} \left(\frac{1}{\lambda^3} + 1 \right)$$

$$\leq 2c \left(\frac{1}{\lambda} + 1 + \frac{1}{2} \left(\frac{1}{\lambda^2} + 1 \right) \right) + \frac{1}{\lambda^3} + 1. \tag{3.19}$$

4) The joint CLT for $\left(n^{-1/2}(d(Z_n) - nv_\lambda), n^{-1/2}P_n \right)_{n \geq 1}$**:** We have given a proof of the annealed CLT for the sequences of random variables $\{n^{-1/2}(d(Z_n) - nv_\lambda)\}_{n \geq 1}$ and $\{n^{-1/2}P_n\}_{n \geq 1}$, but in what follows, we need the joint CLT for the sequence of random vectors $\left(n^{-1/2}(d(Z_n) - nv_\lambda), n^{-1/2}P_n \right)$. Note that for any $\lambda \in (\lambda_c^{1/2}, \mu)$,

$$\left(d(Z_{\tau_{l+1}}) - d(Z_{\tau_l}) - v_\lambda(\tau_{l+1} - \tau_l), \ \sum_{j=\tau_l}^{\tau_{l+1}-1} B_\lambda(Z_j, Z_{j+1}) \right)_{l \geq 1}$$

are *i.i.d.* $\mathbb{R}^2$-valued random variables under P_λ.

This fact together with the moment estimate of regeneration times immediately implies the following result. Note that for $\sigma_{10}(\lambda)$, we use $E_\lambda \left[\sum_{j=\tau_1}^{\tau_2-1} B_\lambda(Z_j, Z_{j+1}) \right] = 0$ from (3.18) and that $\sigma_{00}(\lambda)$ coincides with the diffusion constant in the central limit theorems proved in Bowditch (2018b).

Proposition 3.4. *Suppose $\lambda \in (\lambda_c^{1/2}, \mu)$ and that there exists $\beta > 1$ such that $\sum_{k=1}^{\infty} p_k \beta^k < \infty$. Then, the sequence $\left\{ (n^{-1/2}(d(Z_n) - nv_\lambda), n^{-1/2}P_n) \right\}_{n \geq 1}$ under P_λ converges weakly to the two dimensional Gaussian random variable (X, Y) with the covariance matrix $\Sigma_\lambda := (\sigma_{ij}(\lambda))_{0 \leq i,j \leq 1}$ given by*

$$\sigma_{00}(\lambda) := \frac{1}{E_\lambda[\tau_2 - \tau_1]} E_\lambda \left[\left(\left(d(Z_{\tau_2}) - d(Z_{\tau_1}) \right) - E_\lambda[d(Z_{\tau_2}) - d(Z_{\tau_1})] \right)^2 \right],$$

$$\sigma_{11}(\lambda) := \frac{1}{E_\lambda[\tau_2 - \tau_1]} E_\lambda \left[\sum_{j=\tau_1}^{\tau_2-1} B_\lambda^2(Z_j, Z_{j+1}) \right], \tag{3.20}$$

$$\sigma_{10}(\lambda) = \sigma_{01}(\lambda) := \frac{1}{E_\lambda[\tau_2 - \tau_1]} E_\lambda \left[\left(d(Z_{\tau_2}) - d(Z_{\tau_1}) \right) \sum_{j=\tau_1}^{\tau_2 - 1} B_\lambda(Z_j, Z_{j+1}) \right].$$

Moreover, under P_λ^{NB} the sequence $\left\{ (n^{-1/2}(d(Z_n) - nv_\lambda), n^{-1/2} P_n) \right\}_{n \geq 1}$ converges weakly to the same two dimensional Gaussian random variable (X, Y).

Proof: We have already proved the first claim. The second claim is immediate from the fact that the distribution of

$$\left(d(Z_{\tau_2}) - d(Z_1) - v_\lambda(\tau_2 - \tau_1), \sum_{j=\tau_1}^{\tau_2 - 1} B_\lambda(Z_j, Z_{j+1}) \right) \quad \text{under } P_\lambda$$

is same as that of

$$\left(d(Z_1) - v_\lambda \tau_1, \sum_{j=0}^{\tau_1} B_\lambda(Z_j, Z_{j+1}) \right) \quad \text{under } P_\lambda^{\mathrm{NB}}.$$

$\square$

We now prove Lemma 3.3 by using discussions given in this subsection.

Proof of Lemma 3.3. By the Markov property, we have that

$$\mathbb{E}\left[E_{\lambda+h}^{\mathbf{T}^*} \left[(d(Z_n) - nv_\lambda) \mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right]$$

$$= \mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[(d(Z_n) - nv_\lambda) \prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \mathbf{1}_{\{\sigma_{e^*}>n\}} \cdot E_{\lambda+h, Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right] \right]$$

and

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[(d(Z_n) - nv_\lambda) \prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \cdot \mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right]$$

$$= \mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[(d(Z_n) - nv_\lambda) \prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \mathbf{1}_{\{\sigma_{e^*}>n\}} \cdot E_{\lambda, Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right] \right].$$

Thus, by Hölder's inequality and Jensen's inequality we have that (3.10) is equal to

$$\left| \mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[\left(\frac{d(Z_n) - nv_\lambda}{\sqrt{n}} \right) \mathbf{1}_{\{\sigma_{e^*}>n\}} \cdot \prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right. \right. \right.$$

$$\left. \left. \left. \cdot \left(E_{\lambda+h, Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] - E_{\lambda, Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right) \right] \right] \right|$$

which is bounded above by $\mathcal{E}_{h,n}^1 \cdot \mathcal{E}_{h,n}^2 \cdot \mathcal{E}_{h,n}^3$ where

$$\mathcal{E}_n^1 := \mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[\left(\frac{d(Z_n) - nv_\lambda}{\sqrt{n}} \right)^2 \mathbf{1}_{\{\sigma_{e^*}>n\}} \right] \right]^{1/2},$$

$$\mathcal{E}_{h,n}^2 := \mathbb{E}\left[E_\lambda^{\mathbf{T}^*} \left[\left(\prod_{i=1}^{n} \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right)^4 \right] \right]^{1/4},$$

$$\mathcal{E}_{h,n}^3 := E_\lambda \left[\left(E_{\lambda+h,Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] - E_{\lambda,Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] \right)^4 \right]^{1/4}.$$

It suffices to show that

$$\sup_n \mathcal{E}_n^1 < \infty, \tag{3.21}$$

$$\sup_{h,n:h^2n\sim 1} \mathcal{E}_{h,n}^2 < \infty, \text{ and} \tag{3.22}$$

$$\lim_{h,n:h^2n\sim 1} \mathcal{E}_{h,n}^3 = 0. \tag{3.23}$$

The estimate (3.21) follows from (3.8).

We now prove the estimate (3.22). Notice that

$$\left(\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right)^4 = \exp(4hP_n - 4h^2Q_n + 4R_{n,h})$$

and

$$\prod_{i=1}^n \frac{A_{\lambda+4h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} = \exp(4hP_n - 16h^2Q_n + R_{n,4h}).$$

By the estimate (3.19), there exists a constant $C_\lambda > 0$ such that $|R_{n,h}| \leq C_\lambda$ and $|R_{n,3h}| \leq C_\lambda$. Since $h^2 \sim n^{-1}$, there exists a constant C_λ' such that $|h^2Q_n| \leq C_\lambda'$. Thus, there exists a constant $C_\lambda'' > 0$ such that

$$\left(\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right)^4 \leq C_\lambda'' \left(\prod_{i=1}^n \frac{A_{\lambda+4h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right).$$

Noticing that

$$E_\lambda^{\mathbf{T}^*} \left[\prod_{i=1}^n \frac{A_{\lambda+4h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)} \right] = E_{\lambda+4h}^{\mathbf{T}^*}[1] = 1,$$

we get the conclusion.

Finally, we show (3.23). Note that

$$Q(\lambda) := E_{\lambda,Z_n}^{\mathbf{T}^*} \left[\mathbf{1}_{\{\sigma_{e^*}=\infty\}} \right] = P_{\lambda,Z_n}^{\mathbf{T}^*} (\sigma_{e^*} = \infty)$$

is bounded above by 1 and monotonically decreasing in λ. Furthermore, since $d(Z_n)$ converges P_λ-a.s. to ∞ as $n \to \infty$, we have that $Q(\tilde\lambda)$ converges P_λ-a.s. to 1 for any $\tilde\lambda \in (\lambda_c, \mu)$. Fix $t > 0$ such that $[\lambda - t, \lambda + t] \subset (\lambda_c, \mu)$ then for $|h| \leq t$, by bounded convergence theorem we then have that

$$\mathcal{E}_{h,n}^3 \leq \sqrt{2} E_\lambda \left[P_{\lambda-t,Z_n}^{\mathbf{T}^*} (\sigma_{e^*} = \infty) - P_{\lambda+t,Z_n}^{\mathbf{T}^*} (\sigma_{e^*} = \infty) \right]^{1/4}$$

which converges to 0 as $n \to \infty$. $\square$

3.2. *The proof of the differentiability of the speed.* In this subsection, we will prove Theorem 1.1.

Proof of Theorem 1.1: By Proposition 3.4 it is now sufficient to prove that $v'_\lambda = E_\lambda^{\text{NB}}[XY]$. By (3.11), we have that

$$E_\lambda^{\text{NB}}\left[(d(Z_n) - nv_\lambda)\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right]$$
$$= E_\lambda^{\text{NB}}\left[(d(Z_n) - nv_\lambda)\exp(hP_n - h^2 Q_n + R_{n,h})\right]. \tag{3.24}$$

Therefore, once we justify that we can pass to the limit in (3.24), by using Lemma 3.3 and Proposition 3.4 we will get

$$\frac{1}{hn}E_{\lambda+h}^{\text{NB}}\left[d(Z_n) - nv_\lambda\right]$$
$$\to E_\lambda^{\text{NB}}\left[X\exp\left(Y - \frac{1}{2E_\lambda[\tau_2 - \tau_1]}E_\lambda\left[\sum_{j=\tau_1}^{\tau_2-1}B_\lambda^2(Z_j, Z_{j+1})\right]\right)\right] \tag{3.25}$$

where (X, Y) is the two dimensional Gaussian random variable with the covariance matrix Σ_λ. Notice that we have shown the continuity of the escape probability in Lemma 2.9. Since it is shown in (3.20) that

$$\text{Var}(Y) = \frac{1}{E_\lambda^{\text{NB}}[\tau_1]}E_\lambda^{\text{NB}}\left[\sum_{j=0}^{\tau_1-1}B_\lambda^2(Z_j, Z_{j+1})\right],$$

the above convergence and the integration by parts formula for Gaussian laws implies

$$\frac{v_{\lambda+h} - v_\lambda}{h} \to E_\lambda^{\text{NB}}\left[X\exp\left(Y - \frac{1}{2}\text{Var}(Y)\right)\right] = E_\lambda^{\text{NB}}[XY] = E_\lambda[XY].$$

In order to justify the step (3.25), it suffices to show the uniform integrability of

$$\left\{\frac{1}{hn}\left(d(Z_n) - nv_\lambda\right)\cdot\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right\}_{n\geq 1}.$$

under P_λ^{NB}. By Hölder's inequality, we have

$$E_\lambda^{\text{NB}}\left[\left(\frac{1}{hn}\left(d(Z_n) - nv_\lambda\right)\cdot\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right)^{6/5}\right]$$
$$\leq E_\lambda^{\text{NB}}\left[\frac{1}{(hn)^2}\left(d(Z_n) - nv_\lambda\right)^2\right]^{3/5}E_\lambda^{\text{NB}}\left[\left(\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right)^3\right]^{2/5}$$

In Lemma 3.2, we have already seen that $E_\lambda^{\text{NB}}\left[\frac{1}{(hn)^2}\left(d(Z_n) - nv_\lambda\right)^2\right]$ is bounded in n. That $E_\lambda^{\text{NB}}\left[\left(\prod_{i=1}^n \frac{A_{\lambda+h}(Z_{i-1}, Z_i)}{A_\lambda(Z_{i-1}, Z_i)}\right)^3\right]$ is also bounded in n follows from the estimate (3.22). □

4. Uniform moment bounds on regeneration times

In this section we study regeneration and return times for biased random walks on supercritical GW-trees whose offspring law has exponential moments and no deaths (i.e. $p_0 = 0$). First, we prove that for any $u \in \mathbb{N}$ and $[a, b] \subset (0, \mu)$ we have

$$\sup_{\lambda \in [a,b]} E_\lambda^{\mathrm{NB}} [\tau_1^u] = \sup_{\lambda \in [a,b]} E_\lambda \left[(\tau_2 - \tau_1)^u \right] < \infty. \tag{4.1}$$

This will be used in the proof of Proposition 2.5 where we consider the case with leaves. Following this, we show that the escape probability is continuous in λ thus proving Proposition 2.9.

Towards proving (4.1), we note that, since the interval $[a, b]$ is compact, it suffices to show that for any $\lambda \in (0, \mu)$ and $u \in \mathbb{N}$ there exists $\varepsilon > 0$ such that

$$\sup_{|h| \le \varepsilon} E_\lambda^{\mathrm{NB}} [\tau_1^u] = \sup_{|h| \le \varepsilon} E_{\lambda+h} \left[(\tau_2 - \tau_1)^u \right] < \infty.$$

For $\lambda < 1$ this follows trivially by choosing $\varepsilon < 1 - \lambda$ and comparing with a biased random walk on $\mathbb{Z}$ (e.g. Lemma 5.1 of Dembo et al., 1996). We consider the case $\lambda \ge 1$ and proceed similarly to Proposition 3 in Peres and Zeitouni (2008) in which it is shown that $E_\lambda^{\mathrm{NB}}[\tau_1^u] = E_\lambda \left[(\tau_2 - \tau_1)^u \right] < \infty$ for any $\lambda \in (0, \mu)$ and $u \in \mathbb{N}$.

Our main contribution here is that we show that this bound is uniform in the bias λ in compact intervals for which Remark 4.1 will play an important role.

Remark 4.1. By Rayleigh's monotonicity principle we have that for any infinite tree $\mathscr{T}$ and any $v \in \mathscr{T}$,

$$P_{\lambda,v}^{\mathscr{T}}(\sigma_e = \infty)$$

is monotonically decreasing in λ. This follows using the relationship between electrical networks and reversible Markov chains (see Lyons and Peres, 2016 for further detail).

We now show that the speed v_λ is bounded away from 0 uniformly in λ in compact subsets of $[1, \mu)$.

Lemma 4.2. *Suppose $p_0 = 0$. For any $b \in [1, \mu)$ there exists a constant $c_b > 0$ such that*

$$\inf_{\lambda \in [1,b]} v_\lambda \ge c_b.$$

Proof: By Theorem 3.1 of Lyons et al. (1996), for $\lambda \in (1, \mu)$ we have that $v_\lambda \ge (1 - \lambda^{-1})^3 (1 - q_\lambda)^2 / 12$ where q_λ is the smallest non-negative solution to $f(1 - \lambda^{-1}(1 - q_\lambda)) = q_\lambda$. It is immediate from this that for any $a > 1$ there exists $c_{a,b} > 0$ such that

$$\inf_{\lambda \in [a,b]} v_\lambda \ge c_{a,b}.$$

It therefore remains to consider λ arbitrarily close to 1.

Let ξ be a random variable with the offspring distribution. By Theorem 1.1 of Aïdékon (2014) we have that

$$v_\lambda = \mathbb{E}\left[\frac{(\xi - \lambda)\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^\xi \tilde{p}_\lambda^{(i)}} \right] \bigg/ \mathbb{E}\left[\frac{(\xi + \lambda)\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^\xi \tilde{p}_\lambda^{(i)}} \right]$$

where $\tilde{p}_\lambda^{(i)}$ are independent copies of $P_\lambda^{\mathbf{T}}(\sigma_e = \infty)$ (which are also independent of ξ).

Since $\tilde{p}_\lambda^{(i)}$ are independent of ξ we have that, for $\lambda \in [1, 3/2]$,

$$
\begin{aligned}
\mathbb{E} &\left[\frac{(\xi - \lambda)\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{\xi} \tilde{p}_\lambda^{(i)}} \right] \\
&= \sum_{k=1}^{\infty} \mathbb{P}(\xi = k) \mathbb{E}\left[\frac{(k - \lambda)\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right] \qquad (4.2) \\
&\geq p_1(1 - \lambda)\mathbb{E}\left[\frac{\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{1} \tilde{p}_\lambda^{(i)}} \right] + \frac{1}{4} \sum_{k=2}^{\infty} p_k \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right]
\end{aligned}
$$

since $\xi - \lambda \geq \xi/4$ for $\xi \geq 2$ and $\lambda \leq 3/2$. Similarly,

$$
\mathbb{E}\left[\frac{(\xi + \lambda)\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{\xi} \tilde{p}_\lambda^{(i)}} \right] \leq 2 \sum_{k=1}^{\infty} p_k \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right]. \qquad (4.3)
$$

By Remark 4.1, for any tree $\tilde{p}_\lambda^{(i)}$ is decreasing in λ. Moreover, $\mathbb{P}(\tilde{p}_{1+\varepsilon}^{(i)} > 0) > 0$ for any $\varepsilon \in (0, \mu - 1)$. It follows that there exists $c > 0$ such that for any $k \geq 1$ and $\lambda \in [1, 1 + \varepsilon]$ for $\varepsilon > 0$ suitably small we have that

$$
\begin{aligned}
\frac{k}{k + 1} &\geq \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right] \\
&= \frac{k}{k + 1}\left(1 - \mathbb{E}\left[\frac{\lambda - 1}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right] \right) \\
&\geq \frac{k}{k + 1}\left(1 - \mathbb{E}\left[\frac{\varepsilon}{\varepsilon + \sum_{i=0}^{k} \tilde{p}_{1+\varepsilon}^{(i)}} \right] \right) \\
&\geq \frac{ck}{k + 1}. \qquad (4.4)
\end{aligned}
$$

In particular, we can choose $\varepsilon > 0$ sufficiently small such that

$$
p_1(\lambda - 1)\mathbb{E}\left[\frac{\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{1} \tilde{p}_\lambda^{(i)}} \right] \leq \frac{1}{8} \sum_{k=2}^{\infty} p_k \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right]
$$

uniformly over $\lambda \in [1, 1 + \varepsilon]$. Combining this with (4.2) and (4.3) we have

$$
\upsilon_\lambda \geq \frac{1}{16} \sum_{k=2}^{\infty} p_k \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right] \Big/ \sum_{k=1}^{\infty} p_k \mathbb{E}\left[\frac{k\tilde{p}_\lambda^{(0)}}{\lambda - 1 + \sum_{i=0}^{k} \tilde{p}_\lambda^{(i)}} \right]
$$

which is bounded below for $\lambda \in [1, 1 + \varepsilon]$ for $\varepsilon > 0$ suitably small using (4.4)   □

We now use the Girsanov formula (1.1) to obtain a useful bound relating the laws for different values of λ. Let $\Delta_n := \inf\{m \geq 0 : d(Z_m) = n\}$ be the first time the walk reaches distance n from the root.

Lemma 4.3. *For any tree $\mathcal{T}$ of height at least n, $\lambda \in (0, 1]$ and $h \in (0, \lambda)$ we have that*

$$
P_{\lambda-h}^{\mathcal{T}}(\Delta_n > m, \sigma_e > m) \leq e^{nh} P_\lambda^{\mathcal{T}}(\Delta_n > m, \sigma_e > m).
$$

Proof: First note that the function $F((Z_k)_{k\geq 0}) = \mathbf{1}_{\{\Delta_n > m, \sigma_e > m\}}$ is measurable with respect to $\mathcal{F}_m(\mathscr{T})$ therefore, by the Girsanov formula (1.1) we have that

$$P_{\lambda-h}^{\mathscr{T}}(\Delta_n > m, \sigma_e > m) = E_\lambda^{\mathscr{T}}\left[\mathbf{1}_{\{\Delta_n>m,\sigma_e>m\}}\prod_{i=1}^m \frac{A_{\lambda-h}(Z_{i-1},Z_i)}{A_\lambda(Z_{i-1},Z_i)}\right]. \qquad (4.5)$$

For a walk started from the root, every time the walk takes a step back towards the root it crosses an edge that has previously been crossed. In particular, there is a most recent time that edge was crossed and, due to the tree structure, it must have been crossed directed away from the root. It follows that, for any path $(z_k)_{k=0}^m$ in $\mathscr{T}$, every pair (z_{i-1}, z_i) either corresponds to a unique pair (z_{j-1}, z_j) using this coupling or belongs to the unique self avoiding path starting from the root and ending at z_m. Denote by γ this unique path of length $d(z_m)$.

For a neighbouring pair of vertices $x, y \in \mathscr{T}$ it is straightforward to show that

$$\frac{A_{\lambda-h}(x,y)A_{\lambda-h}(y,x)}{A_\lambda(x,y)A_\lambda(y,x)} \leq 1$$

for $\lambda \in (0,1]$ and $h \in (0,\lambda)$. It follows that,

$$\prod_{i=1}^{d(z_m)} \frac{A_{\lambda-h}(z_{i-1},z_i)}{A_\lambda(z_{i-1},z_i)} = \prod_{x\in\gamma\setminus\{z_0,z_m\}} \frac{\lambda+\nu(x)}{\lambda-h+\nu(x)} \leq e^{h(d(z_m)-1)}. \qquad (4.6)$$

Noting that $\{\Delta_n > m\} \subset \{d(Z_m) < n\}$, combining (4.5) and (4.6) completes the proof. $\qquad\square$

An important result that we will use in the following proof is that the distance between regenerations have exponential moments. That is, by Lemma 4.2 of Dembo et al. (2002) we have that for any $\lambda \in (0,\mu)$ there exists $\theta(\lambda) =: \theta > 0$ such that $E_\lambda[e^{\theta d(Z_{\tau_1})}] < \infty$. In fact, we require the stronger uniform moment bound Lemma 4.4, whose proof is a straightforward extension of that of Lemma 4.2 in Dembo et al. (2002) using Remark 4.1 which we omit.

Lemma 4.4. *Suppose $p_0 = 0$. For any $[a,b] \subset (0,\mu)$ there exists $\theta(a,b) =: \theta > 0$ such that*

$$\sup_{\lambda\in[a,b]} E_\lambda[e^{\theta d(Z_{\tau_1})}] < \infty. \qquad (4.7)$$

We now proceed to the main result of this section. This follows similarly to Proposition 3 in Peres and Zeitouni (2008) however, we include the proof since the extension to uniformity over λ is delicate.

Proposition 4.5. *Suppose $p_0 = 0$, $b \in [1,\mu)$ and that there exists $\beta > 1$ such that $\sum_{k\geq 1} p_k\beta^k < \infty$. For all $u \in \mathbb{N}$ and $\lambda \in [1,\mu)$ there exists $\varepsilon > 0$ such that*

$$\sup_{\lambda\in[1-\varepsilon,b]} E_\lambda\left[(\tau_2 - \tau_1)^u\right] < \infty.$$

Proof: First note that

$$E_\lambda[(\tau_2 - \tau_1)^u] = E_\lambda^{\mathrm{NB}}[\tau_1^u] = \frac{\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}[\tau_1^u\mathbf{1}_{\{\sigma_{e^*}=\infty\}}]\right]}{\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\sigma_{e^*} = \infty)\right]}.$$

Since the denominator $\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\sigma_{e^*} = \infty)\right]$ is monotonic in λ by Remark 4.1, it suffices to consider $\sup_{\lambda\in[1-\varepsilon,b]} \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\tau_1^u\mathbf{1}_{\{\sigma_{e^*}=\infty\}}\right]\right]$. Using the uniform exponential

moment bound (4.7), the Cauchy-Schwarz inequality and integration by parts we have

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\tau_1^u \mathbf{1}_{\{\sigma_{e^*}=\infty\}}\right]\right] \tag{4.8}$$

$$= \sum_{n=1}^\infty \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\tau_1^u; \sigma_{e^*}=\infty, d(Z_{\tau_1})=n\right]\right]$$

$$= \sum_{n=1}^\infty \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\Delta_n^u; \sigma_{e^*}=\infty, d(Z_{\tau_1})=n\right]\right]$$

$$\leq \sum_{n=1}^\infty \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\Delta_n^{2u}; \sigma_{e^*}=\infty\right]\right]^{1/2} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(d(Z_{\tau_1})=n)\right]^{1/2}$$

$$\leq \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[e^{\theta d(Z_{\tau_1})}\right]\right] \sum_{n=1}^\infty e^{-\theta n}\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\Delta_n^{2u}; \sigma_{e^*}=\infty\right]\right]^{1/2}$$

$$\leq \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[e^{\theta d(Z_{\tau_1})}\right]\right] \sum_{n=1}^\infty e^{-\theta n} n^{10u}\left(\sum_{k=0}^\infty (k+1)^{2u}\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_{e^*}=\infty)\right]\right)^{1/2}.$$

By Lemma 4.3 we have that for $\varepsilon > 0$ suitably small

$$\sup_{h\in(0,\varepsilon)} \mathbb{E}\left[P_{1-h}^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_{e^*}=\infty)\right] \leq \sup_{h\in(0,\varepsilon)} \mathbb{E}\left[P_{1-h}^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_{e^*}>kn^{10})\right]$$

$$\leq e^{\varepsilon n}\mathbb{E}\left[P_1^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_{e^*}>kn^{10})\right].$$

Choosing $\varepsilon < \theta/2$ and using (4.7), it suffices to show that

$$\sup_{\lambda\in[1,b]} \sum_{n=1}^\infty e^{-\theta n/2} n^{10u}\left(\sum_{k=0}^\infty (k+1)^{2u}\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_{e^*}>kn^{10})\right]\right)^{1/2} < \infty. \tag{4.9}$$

For $k \geq 1$, let

$$\mathcal{A}_{1,k,n} := \bigcup_{m\leq kn^{10}} \{|\nu(Z_m)| \geq \log(kn^{10})^2\}$$

be the event that the walk visits a vertex with at least $\log(kn^{10})^2$ offspring by time kn^{10}. By the exponential moments assumption we have that for all n large

$$\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\mathcal{A}_{1,k,n})\right] \leq kn^{10}\mathbb{P}(|\nu(e)| \geq \log(kn^{10})^2) \leq e^{-c\log(n^{10})^2}e^{-c\log(k)^2}$$

for some constant c depending only on β.

Let $N_{k,n} := |\{m \leq kn^{10} : Z_l \neq Z_m \forall l < m\}|$ be the number of distinct vertices visited by time kn^{10}. Set

$$\mathcal{A}_{2,k,n} := \left\{N_{k,n} < \sqrt{kn^{10}}\right\} \cap \left\{\sigma_e > kn^{10}\right\}$$

to be the event that, up to time kn^{10}, the walk visits at most $(kn^{10})^{1/2}$ distinct vertices and does not return to the root e^*. On the event $\mathcal{A}_{2,k,n} \cap \mathcal{A}_{1,k,n}^c$ there is a time $m \leq kn^{10}$ and a vertex v with degree at most $\log(kn^{10})^2$ such that $Z_m = v$ and v is subsequently visited at least $(kn^{10})^{1/2}$ times without a visit to the root. By the Gambler's ruin, for a walk started at v of distance at most n from the root, the probability that the walk returns to v before reaching the root is at most $1 - 1/(2n\log(kn^{10})^2)$ uniformly in k, m, v and $\lambda \geq 1$. It follows that the probability

that v is visited by the the walk $(kn^{10})^{1/2}$ times without a visit to the root is at most

$$\left(1 - \frac{1}{2n\log(kn^{10})^2}\right)^{\sqrt{kn^{10}}}.$$

It follows that for n suitably large (independently of $k \geq 1$)

$$\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\mathcal{A}_{2,k,n})\right] \leq \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\mathcal{A}_{1,k,n})\right] + kn^{10}\left(1 - \frac{1}{2n\log(kn^{10})^2}\right)^{\sqrt{kn^{10}}}$$

$$\leq 2e^{-c\log(n^{10})^2}e^{-c\log(k)^2}.$$

On the event $\mathcal{A}_{2,k,n}^c \cap \{\sigma_e > kn^{10}\}$ there are at least $k^{1/2}n^3$ vertices which are visited by the walk before time kn^{10} with at least time n^2 between the first hitting times. Write $\psi_1 := \min\{m > 0 : Z_l \neq Z_m \forall l < m\}$ and, for $i \geq 2$,

$$\psi_i := \min\{m > \psi_{i-1} + n^2 : Z_l \neq Z_m \forall l < m\}.$$

Then, let

$$\mathcal{G}_j = \bigcap_{i=1}^j \left\{\max_{m \leq n^2} |d(Z_{\psi_i}) - d(Z_{\psi_i+m})| < n\right\}.$$

We have that

$$\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\Delta_n > kn^{10}, \sigma_e > kn^{10}, \mathcal{A}_{2,k,n}^c)\right] \tag{4.10}$$

$$\leq \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\bigcap_{i=1}^{k^{1/2}n^3}\left\{\max_{m \leq n^2} |d(Z_{\psi_i}) - d(Z_{\psi_i+m})| < n\right\}\right)\right]$$

$$= \prod_{i=1}^{k^{1/2}n^3} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\max_{m \leq n^2} |d(Z_{\psi_i}) - d(Z_{\psi_i+m})| < n \Big| \mathcal{G}_{i-1}\right)\right]$$

$$= \prod_{i=1}^{k^{1/2}n^3} \left(1 - \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\max_{m \leq n^2} |d(Z_{\psi_i}) - d(Z_{\psi_i+m})| \geq n \Big| \mathcal{G}_{i-1}\right)\right]\right).$$

If the walk regenerates at time ψ_i then $(Z_m)_{m \geq \psi_i}$ is independent of $\mathcal{G}_{i-1}$ (conditionally on Z_{ψ_i}) therefore (4.10) is bounded above by

$$\prod_{i=1}^{k^{1/2}n^3} \left(1 - \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\max_{m \leq n^2} |d(Z_{\psi_i}) - d(Z_{\psi_i+m})| \geq n, d(Z_m) \geq d(Z_{\psi_i}) \forall m \geq \psi_i\right)\right]\right)$$

$$= \prod_{i=1}^{k^{1/2}n^3} \left(1 - \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(\Delta_n < n^2, \sigma_{e^*} = \infty\right)\right]\right)$$

$$= \prod_{i=1}^{k^{1/2}n^3} \left(1 - \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\sigma_e = \infty)\right] P_\lambda^{\mathrm{NB}}\left(\Delta_n < n^2\right)\right).$$

We have seen that $P_\lambda(\sigma_e = \infty)$ is bounded away from 0 for $\lambda \in [1, b]$ therefore it remains to show that, for n large, $P_\lambda^{\mathrm{NB}}\left(\Delta_n < n^2\right)$ is bounded away from 0 uniformly in $\lambda \in [1, b]$. By Markov's inequality

$$P_\lambda^{\mathrm{NB}}\left(\Delta_n \geq n^2\right) \leq \frac{E_\lambda^{\mathrm{NB}}[\Delta_n]}{n^2} \leq \frac{E_\lambda^{\mathrm{NB}}[\tau_1]}{n} \leq \frac{E_\lambda^{\mathrm{NB}}[d(Z_{\tau_1})]}{v_\lambda n}$$

where we have used that there are at most n regenerations up to level n and the formula of the speed (3.4). By Lemmas 4.2 and 4.4 we then have that this converges to 0 (uniformly in λ) as $n \to \infty$ which completes the proof. $\square$

We now prove the following lemma which claims stronger estimates than Proposition 2.8 under the assumption that $p_0 = 0$.

Lemma 4.6. *Suppose that $p_0 = 0$ and that $\sum_{k \geq 1} p_k \beta^k < \infty$ for some $\beta > 1$. Then for any $\lambda \in (0, \mu)$ and any $\in \mathbb{N}$, we have*

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(\sigma_{e^*})^l \mathbf{1}_{\{\sigma_{e^*} < \infty\}}\right]\right] < \infty.$$

Proof: On the event $\{\sigma_{e^*} < \infty\}$, we obviously have that

$$\sigma_{e^*} \leq \tau_1 \quad P_\lambda\text{-}a.s. \tag{4.11}$$

The estimate (4.11) together with Lemma 5.1 in Dembo et al. (1996) implies the result for $0 < \lambda < 1$. The case $\lambda = 1$ can be shown by using (4.11) and Theorem 2 in Piau (1998).

We will show the claim for $1 < \lambda < \mu$. Notice that $\sigma_{e^*} = \sigma_{e^*} \wedge \tau_1$ almost surely on the event $\{\sigma_{e^*} < \infty\}$. Therefore similarly to (4.8), we obtain that

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(\sigma_{e^*})^l \mathbf{1}_{\{\sigma_{e^*} < \infty\}}\right]\right]$$

$$= \sum_{n=1}^\infty \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[(\sigma_{e^*} \wedge \tau_1)^l; \ \sigma_{e^*} < \infty, d(Z_{\tau_1}) = n\right]\right]$$

$$\leq \mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[e^{\theta d(Z_{\tau_1})}\right]\right]$$

$$\cdot \sum_{n=1}^\infty e^{-\theta n} n^{10k} \left(\sum_{k=0}^\infty (k+1)^{2l} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}\left(kn^{10} < \sigma_{e^*} < \infty, \ kn^{10} < \Delta_n\right)\right]\right)^{1/2}.$$

This follows similarly to (4.9). $\square$

We conclude this section by proving Proposition 2.9. For this, we first show the following lemma.

Lemma 4.7. *Suppose that $p_0 = 0$. For any $[a, b] \subset (0, \mu)$*

$$\lim_{n \to \infty} \sup_{\lambda \in [a,b]} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(n < \sigma_{e^*} < \infty)\right] = 0.$$

Proof: Let $s_n \to \infty$ be an increasing sequence that we shall specify later and $\theta > 0$ be as in Lemma 4.4. We now split into the cases where $d(Z_{\tau_1}) < s_n$ and $d(Z_{\tau_1}) \geq s_n$. First, for $d(Z_{\tau_1}) \geq s_n$, by Markov's inequality

$$\lim_{n \to \infty} \sup_{\lambda \in [a,b]} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(n < \sigma_{e^*} < \infty, d(Z_{\tau_1}) \geq s_n)\right]$$

$$\leq \lim_{n \to \infty} \sup_{\lambda \in [a,b]} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(d(Z_{\tau_1}) \geq s_n)\right]$$

$$= \lim_{n \to \infty} \sup_{\lambda \in [a,b]} P_\lambda(d(Z_{\tau_1}) \geq s_n)$$

$$\leq \lim_{n \to \infty} \sup_{\lambda \in [a,b]} E_\lambda[e^{\theta d(Z_{\tau_1})}] e^{-\theta s_n}$$

which converges to 0 as $n \to \infty$ by Lemma 4.4.

For $d(Z_{\tau_1}) < s_n$ we note that

$$\{n < \sigma_{e^*} < \infty, \, d(Z_{\tau_1}) < s_n\} \subset \{\max_{m \leq n} d(Z_m) < s_n\}$$

since once the walk reaches level s_n it cannot return to e^* on the event $\{d(Z_{\tau_1}) < s_n\}$. With a slight abuse of notation, let $\mathbb{Z}^+$ denote the tree which is isomorphic to $\mathbb{Z}^+$. By comparison with a simple random walk on $\mathbb{Z}^+$ we have that

$$\limsup_{n \to \infty} \sup_{\lambda \in [a,b]} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(n < \sigma_{e^*} < \infty, d(Z_{\tau_1}) < s_n) \right]$$

$$\leq \limsup_{n \to \infty} \sup_{\lambda \in [a,b]} P_\lambda \left(\max_{m \leq n} d(Z_m) < s_n \right)$$

$$\leq \lim_{n \to \infty} P_\mu^{\mathbb{Z}^+} \left(\max_{m \leq n} d(Z_m) < s_n \right)$$

$$= \lim_{n \to \infty} P_\mu^{\mathbb{Z}^+} \left(\kappa(s_n) > n \right)$$

where $\kappa(l) := \inf\{n \geq 0 : d(Z_n) = l\}$ is the first hitting time of level l.

A simple calculation shows that $E_\mu^{\mathbb{Z}^+}[\kappa(l)] \leq C_\mu \mu^l$. Therefore, choosing $s_n = \log(n)/\log(\mu^2)$ and using Markov's inequality we have that $P_\mu^{\mathbb{Z}^+}(\kappa(s_n) > n) \leq C_\mu n^{-1/2}$ and therefore

$$\limsup_{n \to \infty} \sup_{\lambda \in [a,b]} \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(n < \sigma_{e^*} < \infty, d(Z_{\tau_1}) < s_n) \right] = 0.$$

$\square$

We are now ready to prove that the escape probability $\mathbb{E}[P_\lambda^{\mathbf{T}^*}(\sigma_{e^*(\mathbf{T})} < \infty)]$ is a continuous function of the bias.

Proof of Proposition 2.9: First note that we can assume $p_0 = 0$ without loss of generality since $P_\lambda^{\mathbf{T}^*}\left(\sigma_{e^*(\mathbf{T})} = \infty \right) = P_\lambda^{\mathbf{T}_g^*}\left(\sigma_{e^*(\mathbf{T}_g)} = \infty \right)$ $\mathbb{P}$-a.s. where we recall that $\mathbf{T}_g^*$ is the backbone of $\mathbf{T}^*$.

Note that for any tree $\mathscr{T}$ and any $n \in \mathbb{N}$, the function $\lambda \mapsto P_\lambda^{\mathscr{T}}(\sigma_{e^*} < n)$ is continuous since $P_\lambda^{\mathscr{T}}(\sigma_{e^*} < n)$ only depends on the first n steps of (Z_n). We now show that for any $n \in \mathbb{N}$,

$$\text{the function } \lambda \mapsto \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\sigma_{e^*} < n) \right] \text{ is continuous.} \tag{4.12}$$

For a tree $\mathscr{T}$ and $n \in \mathbb{N}$ we write $\mathscr{T}[n]$ for the truncated tree up to n^{th} generation, and define

$$\mathcal{B}_{n,m} := \bigcap_{k=1}^{n} \{\mathscr{T}; \text{ every vertex of } \mathscr{T}[n] \text{ in } k^{\text{th}} \text{ generation has at most } m \text{ children}\}.$$

Noticing that $P_\lambda^{\mathscr{T}}(\sigma_{e^*} < n)$ only depends on $\mathscr{T}[n]$ and that $\{\mathscr{T}[n]; \, \mathscr{T} \in \mathcal{B}_{n,m}\}$ is a finite set, we obtain that the function $\lambda \mapsto \mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(\sigma_{e^*} < n) \mathbf{1}_{\mathcal{B}_{n,m}} \right]$ is a continuous function. Now the claim (4.12) follows since $\mathbb{P}(\mathcal{B}_{n,m}^c)$ is independent of λ and converges to 0 as $m \to \infty$ for any n.

In order to deduce the conclusion from (4.12), it suffices to prove that

$$\mathbb{E}\left[P_\lambda^{\mathbf{T}^*}(n < \sigma_{e^*} < \infty) \right]$$

is uniformly convergent to 0 in (λ_c, μ) as $n \to \infty$. This immediately follows from Lemma 4.7. $\qquad\square$

5. Moments of generation sizes of Galton-Watson trees

In this section we prove several technical estimates for subcritical GW-trees which we will require later when showing moment bounds for the time between regenerations of the walk. For this section, we take W_n to be a GW-process with mean number of offspring $\mu := \mathbb{E}[W_1] < 1$ which will typically be applied as λ_c in Section 6.

The following lemma gives a bound on the moments of generation sizes of GW-processes. The main purpose of this lemma is to prove Lemma 5.2.

Lemma 5.1. *Suppose W_n is a GW-process with mean number of offspring $\mu := \mathbb{E}[W_1] < 1$ and which satisfies $\mathbb{E}[\beta^{W_1}] < \infty$ for some $\beta > 1$. Then, for any $m \in \mathbb{N}$ there exists $C_m < \infty$ such that $\mathbb{E}[W_n^m] \le C_m \mu^n$.*

Proof: We prove this inductively in m. The case $m = 1$ holds with $\mathbb{E}[W_n] = \mu^n$ (cf. Chapter I.2 of Athreya and Ney, 2004). Suppose that for some $m \ge 2$ there exist $C_j < \infty$ for $j = 1, ..., m-1$ such that $\mathbb{E}[W_n^j] \le C_j \mu^n$.

Let $W_n^{(1)}, W_n^{(2)}, ...$ be independent copies of W_n then, using the branching property,

$$\mathbb{E}[W_{n+1}^m] = \mathbb{E}[\mathbb{E}[W_{n+1}^m | W_1]] = \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{k=1}^{W_1} W_n^{(k)}\right)^m \Big| W_1\right]\right].$$

For $l, m, N \in \mathbb{N}$ let $\mathcal{I}_l^m(N) := \{\mathbf{k} = (k_1, ..., k_m) \in \{1, ..., N\}^m : \sum_{j=1}^N \mathbf{1}_{\bigcup_{i=1}^m \{k_i = j\}} = l\}$ be the m-tuples of positive integers at most N with exactly l distinct values. Expanding the term in the above expression and using that $W_n^{(k)}$ are independent of W_1 we have that

$$\mathbb{E}[W_{n+1}^m] = \mathbb{E}\left[\mathbb{E}\left[\sum_{\mathbf{k} \in \{1, ..., W_1\}^m} \prod_{i=1}^m W_n^{(k_i)} | W_1\right]\right]$$

$$= \sum_{l=1}^m \mathbb{E}\left[\sum_{\mathbf{k} \in \mathcal{I}_l^m(W_1)} \mathbb{E}\left[\prod_{i=1}^m W_n^{(k_i)}\right]\right]. \qquad (5.1)$$

If $\mathbf{k} \in \mathcal{I}_1^m(W_1)$ then $k_i = k_1$ for all i and, since there are W_1 choices of k_1, we have

$$\mathbb{E}\left[\sum_{\mathbf{k} \in \mathcal{I}_1^m(W_1)} \mathbb{E}\left[\prod_{i=1}^m W_n^{(k_i)}\right]\right] = \mathbb{E}\left[\sum_{\mathbf{k} \in \mathcal{I}_1^m(W_1)} \mathbb{E}[(W_n^{(k_1)})^m]\right] = \mathbb{E}[W_1]\mathbb{E}[W_n^m]. \quad (5.2)$$

Otherwise, using independence of $W_n^{(1)}, W_n^{(2)}, ...$ and our induction hypothesis that $\mathbb{E}[W_n^j] \le C_j \mu^n$ for $j \le m-1$, for $k \in \mathcal{I}_l^m(W_1)$ we have

$$\mathbb{E}\left[\prod_{i=1}^m W_n^{(k_i)}\right] \le \mu^{nl}\left(\max_{j \le m-1} C_j\right)^l.$$

There are $\prod_{j=0}^{l-1}(W_1 - j)$ choices for the l distinct values in $\{1, ..., W_1\}$ then l^{m-l} choices for the remaining $m - l$ duplicates and at most $m!$ orderings of the indices. In particular, for $l \geq 2$,

$$\mathbb{E}\left[\sum_{\mathbf{k}\in\mathcal{I}_l^m(W_1)} \mathbb{E}\left[\prod_{i=1}^{m} W_n^{(k_i)}\right]\right] \leq \mu^{nl}(\max_{j\leq m-1} C_j)^l l^{m-l} m! \mathbb{E}\left[\prod_{j=0}^{l-1}(W_1 - j)\right]. \quad (5.3)$$

By the exponential moment assumption we have that $\mathbb{E}[\prod_{j=0}^{l-1}(W_1-j)] \leq \mathbb{E}[W_1^l] < \infty$. Combining (5.1), (5.2) and (5.3), we can choose constants M_l^m such that

$$\mathbb{E}[W_{n+1}^m] \leq \mu\mathbb{E}[W_n^m] + \sum_{l=2}^{m} M_l^m \mu^{nl}$$

$$\leq \mu^2 \mathbb{E}[W_{n-1}^m] + \mu \sum_{l=2}^{m} M_l^m \mu^{(n-1)l} + \sum_{l=2}^{m} M_l^m \mu^{nl}$$

$$= \mu^2 \mathbb{E}[W_{n-1}^m] + \sum_{l=2}^{m} M_l^m \mu^{nl}(1 + \mu^{-(l-1)}).$$

Iterating and using the geometric sum formula yields

$$\mathbb{E}[W_{n+1}^m] \leq \mu^n \mathbb{E}[W_1^m] + \sum_{l=2}^{m} M_l^m \mu^{nl} \sum_{k=0}^{n-1} \mu^{-k(l-1)}$$

$$\leq \mu^n \mathbb{E}[W_1^m] + \sum_{l=2}^{m} M_l^m \mu^{nl} \frac{\mu^{-n(l-1)} - 1}{\mu^{-(l-1)} - 1}$$

which is bounded above by $C_m \mu^{n+1}$ as required since $\mathbb{E}[W_1^m] < \infty$ by the exponential moments assumption. $\qquad\square$

The corresponding lower bound holds trivially by noting that $\mathbb{P}(W_n \geq 1) \leq \mathbb{E}[W_n^m]$ for any $m \in \mathbb{N}$ and using that $\mathbb{P}(W_n \geq 1)\mu^{-n}$ is decreasing and converges (e.g. Theorem B in Lyons et al., 1995). This shows that, up to constants, this is the best possible bound.

The following result extends Lemma 5.1 to the expectation of products of the generation sizes at varying times. This is an extension of Lemma 2.4.1 in Bowditch (2017) which proves this for $m \leq 3$.

Lemma 5.2. *Suppose W_n is a GW-process with mean number of offspring $\mu :=$ $\mathbb{E}[W_1] < 1$ and which satisfies $\mathbb{E}[\beta^{W_1}] < \infty$ for some $\beta > 1$. Then, for any $m \in \mathbb{N}$ there exists $\tilde{C}_m < \infty$ such that for any $(n_i)_{i=1}^m \in \mathbb{N}^m$ we have*

$$\mathbb{E}\left[\prod_{i=1}^{m} W_{n_i}\right] \leq \tilde{C}_m \mu^{\max_{l\leq m} n_l}.$$

Proof: Let $W_n^{(k)}$ be independent GW-processes for $k \geq 1$. Using the branching property of GW-processes and convexity of polynomials of degree $l \in \mathbb{N}$ we have

$$\mathbb{E}[W_n^l|W_0 = j] = \mathbb{E}\left[\left(\sum_{k=1}^{j} W_n^{(k)}\right)^l\right] \leq j^l \mathbb{E}\left[\sum_{k=1}^{j} \frac{(W_n^{(k)})^l}{j}\right] = j^l \mathbb{E}[W_n^l]. \quad (5.4)$$

Without loss of generality let $n_1 \leq n_2 \leq ... \leq n_m$ be ordered. Noting that W_n is a Markov process, by (5.4) we have

$$\mathbb{E}\left[\prod_{i=1}^{m} W_{n_i}\right] = \mathbb{E}\left[\mathbb{E}\left[W_{n_m} | W_{n_{m-1}}\right] \prod_{i=1}^{m-1} W_{n_i}\right] \leq \mathbb{E}[W_{n_m - n_{m-1}}]\mathbb{E}\left[W_{n_{m-1}}^2 \prod_{i=1}^{m-2} W_{n_i}\right].$$

Iterating and applying Lemma 5.1 then gives

$$\mathbb{E}\left[\prod_{i=1}^{m} W_{n_i}\right] \leq \prod_{i=1}^{m} \mathbb{E}[W_{n_i - n_{i-1}}^{m+1-i}] \leq \prod_{i=1}^{m} C_{m+1-i}\mu^{n_k - n_{i-1}} \leq \tilde{C}_m \mu^{n_m}$$

where $\tilde{C}_m = (\max_{l \leq m} C_l)^m < \infty$.   □

6. The proof of Proposition 2.5

The main aim of this section is to prove Corollary 2.6 which states that for any closed ball B contained within $(\lambda_c^{1/2}, \mu)$ there exists $\varepsilon > 0$ such that the time between regenerations has finite $(2 + \varepsilon)^{\text{th}}$ moments uniformly over $\lambda \in B$. We deduce this from the more general result Proposition 2.5.

We first state the following lemma which gives a useful bound for the αth moments of a geometric random variable. This will be used repeatedly throughout this section.

Lemma 6.1. *For any $\alpha > 0$ there exists $C_\alpha < \infty$ such that for any $p \in (0,1)$ we have*

$$\sum_{k=1}^{\infty} k^\alpha p^k (1-p) \leq C_\alpha p(1-p)^{-\alpha}.$$

Proof: Note that if $f : \mathbb{R} \to \mathbb{R}^+$ is increasing and $g : \mathbb{R} \to \mathbb{R}^+$ is decreasing then for $x \in [k, k+1)$ we have that $f(x-1) \leq f(k) \leq f(x)$ and $g(x) \leq g(k) \leq g(x-1)$. Therefore,

$$\sum_{k=1}^{\infty} f(k)g(k) = \sum_{k=1}^{\infty} \int_{k}^{k+1} f(k)g(k)\mathrm{d}x$$
$$\leq \sum_{k=1}^{\infty} \int_{k}^{k+1} f(x)g(x-1)\mathrm{d}x$$
$$= \int_{1}^{\infty} f(x)g(x-1)\mathrm{d}x.$$

Take the specific case that $f(x) = x^\alpha$ (which is increasing since $\alpha > 0$) and $g(x) = p^x$ (which is decreasing for $p \in (0,1)$). Then, for $p \in [1/2, 1)$, we have that

$$p^{-1}(1-p)^{1+\alpha} \sum_{k=1}^{\infty} k^\alpha p^k \leq p^{-2}(1-p)^{1+\alpha} \int_{1}^{\infty} x^\alpha p^x \mathrm{d}x$$
$$= p^{-2}\left(\frac{1-p}{\log(p^{-1})}\right)^{1+\alpha} \int_{\log(p^{-1})}^{\infty} x^\alpha e^{-x}\mathrm{d}x$$
$$\leq 4\Gamma(1+\alpha)$$

since $\left(\frac{1-p}{\log(p^{-1})}\right)^{1+\alpha} \leq 1$.

For $p \in (0, 1/2]$ we have that

$$p^{-1}(1-p)^{1+\alpha} \sum_{k=1}^{\infty} k^{\alpha} p^k \leq \sum_{k=1}^{\infty} k^{\alpha} 2^{-(k-1)}$$

which converges. □

We now introduce some notation concerning hitting and regeneration times. Recall that $\sigma_x := \inf\{n \geq 1 : Z_n = x\}$ is the first return time to $x \in \mathbf{T}$. Let $S(0) := 0$, $S(n) := \inf\{k > S(n-1) : Z_k, Z_{k-1} \in \mathbf{T}_g\}$ for $n \geq 1$ and $Y_n := Z_{S(n)}$, then Y_n is a λ-biased random walk on $\mathbf{T}_g$ coupled to Z_n. Write $\zeta_0 := 0$ and for $m = 1, 2, \ldots$ let

$$\zeta_m := \inf\{k > \zeta_{m-1} : d(Y_j) < d(Y_k), d(Y_l) > d(Y_{k-1}) \text{ for all } j < k \leq l\}$$

be regeneration times for the walk Y. We then have that $\tau_k = \inf\{m \geq 0 : Z_m = Y_{\zeta_k}\}$ are the corresponding regeneration times for Z and we define $\varrho_k := Z_{\tau_k} = Y_{\zeta_k}$ to be the regeneration points. By Proposition 3.4 of Lyons et al. (1996) we have that there exists, P_λ-a.s., an infinite sequence of regeneration times $(\tau_k)_{k \geq 1}$ and

$$\{(\tau_{k+1} - \tau_k), (d(\varrho_{k+1}) - d(\varrho_k))\}_{k \geq 1}$$

are $i.i.d.$ (as are the corresponding variables for Y).

Let ξ_f, ξ_g, ξ_h be random variables with probability generating functions f, g and h respectively then let ξ be equal in distribution to the number of vertices in the first generation of $\mathbf{T}$. Throughout we will assume that ξ_f has some exponential moments.

Remark 6.2. Since the generation sizes of $\mathbf{T}_g$ are dominated by those of $\mathbf{T}$ we have that ξ_g is stochastically dominated by ξ. Using Bayes' law we have that $\mathbf{P}(\xi = k) = p_k(1-q^k)(1-q)^{-1} \leq cp_k$ therefore both ξ and ξ_g inherit the exponential moment bounds of ξ_f. Furthermore $\mathbf{P}(\xi_h = k) = p_k q^k$ therefore ξ_h automatically has exponential moments.

We now show that the duration of an excursion in a single trap has finite α^{th} moments (uniformly for the bias in a small ball). If $p_0 = 0$ then traps are trivial therefore assume that $p_0 > 0$. Denote by $\mathbf{T}_h$ a GW tree with this law and $\mathbf{T}_h^*$ the tree $\mathbf{T}_h$ where we append an additional vertex $e^*(\mathbf{T}_h)$ as the root in the usual way (for convenience we write e^* when there is no confusion). Let $W_k^{\mathbf{T}_h^*}$ denote the kth generation size of the tree $\mathbf{T}_h^*$. We denote by $P_{\lambda,x}^{\mathbf{T}_h^*}$ the quenched law of the walk with bias λ started from x.

Lemma 6.3. *Suppose $p_0 > 0$ and $a < 1$ then, for any $\alpha < \log(\lambda_c)/\log(a)$,*

$$\sup_{\lambda \geq a} \mathbb{E}\left[E_\lambda^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha]\right] < \infty.$$

Proof: Write $\underline{\alpha} := \max\{k \in \mathbb{Z} : k < \alpha\}$. Throughout we will use that for $N \in \mathbb{N}$ and $x_n \in \mathbb{R}_+$ for $n = 1, \ldots, N$ we have

$$\left(\sum_{n=1}^{N} x_n\right)^\alpha \leq N^{\underline{\alpha}} \sum_{n=1}^{N} x_n^\alpha \tag{6.1}$$

which follows from convexity for $\alpha \geq 1$ and the bound $||\cdot||_{1/\alpha} \leq ||\cdot||_1$ for l^p norms with $\alpha < 1$.

We can write

$$\sigma_{e^*} = \sum_{x \in \mathbf{T}_h^*} v_x \qquad \text{where} \qquad v_x = \sum_{k=0}^{\sigma_{e^*}-1} \mathbf{1}_{\{Z_k=x\}}$$

is the number of visits to x before returning to e^*. By (6.1) it then follows that

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha] \right] = \mathbb{E}\left[E_\lambda^{\mathbf{T}_h^*}\left[\left(\sum_{x \in \mathbf{T}_h^*} v_x \right)^\alpha \right] \right] \le \mathbb{E}\left[d(\mathbf{T}_h^*)^\alpha \sum_{x \in \mathbf{T}_h^*} E_\lambda^{\mathbf{T}_h^*}[v_x^\alpha] \right] \quad (6.2)$$

where, using a decomposition up to the first hitting time of x we have that

$$E_{\lambda,e^*}^{\mathbf{T}_h^*}[v_x^\alpha] = P_{\lambda,e^*}^{\mathbf{T}_h^*}(\sigma_x < \sigma_{e^*}) \, E_{\lambda,x}^{\mathbf{T}_h^*}[v_x^\alpha] \le E_{\lambda,x}^{\mathbf{T}_h^*}[v_x^\alpha].$$

Started from x, for the walk to reach to e^* before returning to x, the walk must initially move to $\pi(x)$. It follows that the number of visits to x before reaching e^* is geometrically distributed with termination probability

$$P_{\lambda,x}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x) = \frac{\lambda}{\lambda + \nu(x)} \cdot P_{\lambda,\pi(x)}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x) \qquad\qquad (6.3)$$

where $P_{\lambda,\pi(x)}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x)$ depends only on λ and the distance between e^* and x. By Lemma 6.1 we have that, for some constant C_α,

$$E_{\lambda,x}^{\mathbf{T}_h^*}[v_x^\alpha] \le C_\alpha P_{\lambda,x}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x)^{-\alpha} P_{\lambda,x}^{\mathbf{T}_h^*}(\sigma_x < \sigma_{e^*}) \le C_\alpha P_{\lambda,x}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x)^{-\alpha}. \quad (6.4)$$

For $r \in \mathbb{N}$ write

$$\mathcal{R}(\lambda,\alpha,r) = \begin{cases} r^\alpha & \text{if } \lambda = 1, \\ \lambda^{-r\alpha} & \text{if } \lambda < 1, \\ 1 & \text{if } \lambda > 1, \end{cases}$$

then, by the Gambler's ruin and (6.3), we have that

$$P_{\lambda,x}^{\mathbf{T}_h^*}(\sigma_{e^*} < \sigma_x)^{-\alpha} \le (1 + \lambda^{-1}\nu(x))^\alpha \mathcal{R}(\lambda,\alpha,d^*(x))$$

where $d^*(x)$ denotes the distance between $x \in \mathbf{T}_h^*$ and the root e^*. Substituting this with (6.4) into (6.2) and using (6.1) we have that

$$\mathbb{E}\left[E_{\lambda,e^*}^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha] \right] \le C_\alpha \mathbb{E}\left[d(\mathbf{T}_h^*)^\alpha \sum_{x \in \mathbf{T}_h^*} (1 + \lambda^{-1}\nu(x))^\alpha \mathcal{R}(\lambda,\alpha,d^*(x)) \right]$$

$$\le \tilde{C}_\alpha \mathbb{E}\left[d(\mathbf{T}_h^*)^\alpha \sum_{x \in \mathbf{T}_h^*} (1 + \lambda^{-\alpha}\nu(x)^\alpha) \mathcal{R}(\lambda,\alpha,d^*(x)) \right]$$

$$\le \tilde{C}_\alpha \mathbb{E}\left[d(\mathbf{T}_h^*)^\alpha \sum_{k=0}^\infty W_k^{\mathbf{T}_h^*}\left(1 + \lambda^{-\alpha}\left(W_{k+1}^{\mathbf{T}_h^*} \right)^{\alpha+1} \right) \mathcal{R}(\lambda,\alpha,k) \right]$$

where, for the final inequality, we have replaced the sum over vertices in the tree with a sum over the generations and bounded the number of children of a vertex in generation k with the total number of vertices in generation $k+1$.

Since $W_1^{\mathbf{T}_h^*} = 1 = W_0^{\mathbf{T}_h}$ and $W_{k+1}^{\mathbf{T}_h^*} = W_k^{\mathbf{T}_h}$ for $k \ge 1$, we have that $1 + \lambda^{-\alpha}\left(W_{k+1}^{\mathbf{T}_h^*} \right)^{\alpha+1} \le C \sum_{j=0}^\infty \left(W_j^{\mathbf{T}_h} \right)^{\alpha+1}$ for any $k \ge 1$ and a constant $C \le 1 + a^{-\alpha}$.

The process $W_n^{\mathbf{T}_h}$ is a GW-process with offspring distribution ξ_h which has mean λ_c and exponential moments. It therefore follows from Lemma 5.2 that

$$\mathbb{E}\left[E_{\lambda,e^*}^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha]\right] \leq C_\alpha \mathbb{E}\left[\sum_{k_1=0}^\infty \cdots \sum_{k_{2\underline{\alpha}+2}=0}^\infty \mathcal{R}(\lambda,\alpha,k_1)\prod_{j=1}^{2\underline{\alpha}+2} W_{k_j}^{\mathbf{T}_h^*}\right]$$

$$\leq C_\alpha \sum_{k_1=0}^\infty \cdots \sum_{k_{2\underline{\alpha}+2}=0}^\infty \mathcal{R}(\lambda,\alpha,k_1)\mathbb{E}\left[\prod_{j=1}^{2\underline{\alpha}+2} W_{k_j}^{\mathbf{T}_h^*}\right]$$

$$\leq \tilde{C}_\alpha \sum_{k_1=0}^\infty \cdots \sum_{k_{2\underline{\alpha}+2}=0}^\infty \mathcal{R}(\lambda,\alpha,k_1)\lambda_c^{\max_{j\leq 2\underline{\alpha}+2} k_j}. \qquad (6.5)$$

Taking first those terms in (6.5) where $k_1 \geq k_j$ for all j, we have

$$\sum_{k_1=0}^\infty \cdots \sum_{k_{2\underline{\alpha}+2}=0}^\infty \mathbf{1}_{\{k_1=\max_{j\leq 2\underline{\alpha}+2} k_j\}}\mathcal{R}(\lambda,\alpha,k_1)\lambda_c^{k_1}$$

$$\leq (2\underline{\alpha}+1)\sum_{k_1=0}^\infty (k_1+1)\mathcal{R}(\lambda,\alpha,k_1)\lambda_c^{k_1}$$

which is bounded above uniformly over $\lambda \geq a$ since $a^{-\alpha}\lambda_c < 1$ by our choice of α.

Next, writing $m = \max_{j=2,\dots,2\underline{\alpha}+2} k_j$, taking the remaining terms in (6.5) and noting that

$$\sum_{k_1=0}^{m-1}\mathcal{R}(\lambda,\alpha,k_1) \leq m^2 a^{-m}$$

we have

$$\sum_{k_1=0}^\infty \cdots \sum_{k_{2\underline{\alpha}+2}=0}^\infty \mathbf{1}_{\{k_1<m\}}\mathcal{R}(\lambda,\alpha,k_1)\lambda_c^m \leq (2\underline{\alpha}+1)\sum_{m=1}^\infty m^3 a^{-m\alpha}\lambda_c^m$$

which is finite by our choice of α. $\qquad\qquad\square$

Let $\chi_k := S(k+1) - S(k)$ denote the total time taken between Z_n making the kth and $(k+1)^{\text{th}}$ transition along the backbone. This time consists of

$$N_k := \sum_{n=S(k)+1}^{S(k+1)} \mathbf{1}_{\{Z_n=Y_k\}}$$

excursions into the finite trees appended to the backbone at this vertex and one additional step to the next backbone vertex. Write $\vartheta_k^{(0)} := S(k)$ and $\vartheta_k^{(j)} := \inf\{n > \vartheta_k^{(j-1)} : Z_n = Y_k\}$ for $j \geq 1$ to be the hitting times of the backbone after time $S(k)$. We can then write

$$\chi_k := 1 + \sum_{j=1}^{N_k} \gamma_{k,j} \qquad \text{where} \qquad \gamma_{k,j} := \vartheta_k^{(j)} - \vartheta_k^{(j-1)} \qquad (6.6)$$

is the duration of the jth such excursion.

Proof of Proposition 2.5: Recall that $\underline{\alpha} := \max\{k \in \mathbb{Z} : k < \alpha\}$ and write $\overline{\alpha} := \min\{k \in \mathbb{Z} : k \geq \alpha\}$. We therefore have that $E_\lambda^{\text{NB}}[\tau_1^\alpha]$ can be written as

$$E_\lambda^{\text{NB}}[\tau_1^\alpha] = E_\lambda^{\text{NB}}\left[\left(\sum_{k=1}^{\zeta_1} \chi_k\right)^\alpha\right]$$

$$\leq E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{k=1}^{\zeta_1} \chi_k^\alpha\right]$$

by (6.1). Using (6.1) again with the decomposition (6.6) we can write this as

$$E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{k=1}^{\zeta_1}\left(1 + \sum_{j=1}^{N_k} \gamma_{k,j}\right)^\alpha\right]$$

$$\leq E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{k=1}^{\zeta_1} (N_k + 1)^{\overline{\alpha}}\left(1 + \sum_{j=1}^{N_k} \gamma_{k,j}^\alpha\right)\right].$$

The excursion times $\gamma_{k,j}$ are distributed as the first return time to e^* for a walk started from e^* on $\mathbf{T}_h^*$. Moreover, under P_λ^{NB}, they are independent of the backbone, the buds and the walk on the backbone and buds. In particular, they are independent of the regeneration times of Y and the number of excursions therefore the above expectation can be bounded above by

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha]\right] E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{k=1}^{\zeta_1} (N_k + 1)^{\overline{\alpha}}\right].$$

Where, by Lemma 6.3, we have that $\sup_{\lambda \in [a,b]} \mathbb{E}\left[E_\lambda^{\mathbf{T}_h^*}[\sigma_{e^*}^\alpha]\right] < \infty$.

Let $(z_j)_{j=0}^\infty$ denote the ordered distinct vertices visited by Y and

$$\mathcal{L}(z, j) := \sum_{k=0}^{j} \mathbf{1}_{\{Y_k = z\}}, \quad \mathcal{L}(z) := \mathcal{L}(z, \infty)$$

the local times of the vertex z. Write

$$M_{z,l} := \sum_{j=0}^\infty \mathbf{1}_{\{Z_j = z, \, Z_{j+1} \notin \mathbf{T}_g, \, \mathcal{L}(z,j)=l\}}$$

to be the number of excursions from z (by Z) on the lth visit to z (by Y) for $l = 1, ..., \mathcal{L}(z)$ and $\mathcal{J} := |\{Y_j\}_{j=1}^{\zeta_1 - 1}|$ the number of distinct vertices visited by Y between time 1 and time $\zeta_1 - 1$. Each $k \leq \zeta_1$ corresponds to a unique pair (z_j, l) with $j \leq \mathcal{J}$ and $l \leq \mathcal{L}(z_j)$ with $M_{z_j,l} = N_k$ therefore

$$E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{k=1}^{\zeta_1} (N_k + 1)^{\overline{\alpha}}\right]$$

$$= E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \sum_{j=1}^{\mathcal{J}} \sum_{l=1}^{\mathcal{L}(z_j)} (M_{z_j,l} + 1)^{\overline{\alpha}}\right]$$

$$= \sum_{j=1}^\infty \sum_{l=1}^\infty E_\lambda^{\text{NB}}\left[\zeta_1^{\overline{\alpha}} \mathbf{1}_{\{j \leq \mathcal{J}, \, l \leq \mathcal{L}(z_j)\}} (M_{z_j,l} + 1)^{\overline{\alpha}}\right]$$

$$\leq \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \left(E_\lambda^{\text{NB}} \left[\zeta_1^{2\alpha} \mathbf{1}_{\{j \leq \mathcal{J}, \, l \leq \mathcal{L}(z_j)\}} \right] E_\lambda^{\text{NB}} \left[(M_{z_j,l} + 1)^{2\overline{\alpha}} \right] \right)^{1/2} \qquad (6.7)$$

by the Cauchy-Schwarz inequality. For all $1 \leq j \leq \mathcal{J}$ we have that $\mathcal{L}(z_j) \leq \zeta_1$; moreover, $\mathcal{J} \leq \zeta_1$ therefore

$$\mathbf{1}_{\{j \leq \mathcal{J}, \, l \leq \mathcal{L}(z_j)\}} \leq \mathbf{1}_{\{j, l \leq \zeta_1\}}.$$

Due to the independence structure of the GW-tree, for any fixed j the distribution of the number of children of z_j is equal to the distribution of the number of children of the root. Since the root does not have a parent, we have that the walk is more likely to take an excursion into one of the neighbouring traps when at the root than from a vertex with the same number of children. We can, therefore, stochastically dominate the number of excursions from a backbone vertex by the number of excursions from the root to see that $E_\lambda^{\text{NB}} \left[(M_{z_j,l} + 1)^{2\overline{\alpha}} \right] \leq E_\lambda^{\text{NB}} \left[(M_{z_0,1} + 1)^{2\overline{\alpha}} \right]$.

Using this and the Cauchy-Schwarz inequality, the expression (6.7) is bounded above by

$$E_\lambda^{\text{NB}} \left[\zeta_1^{4\alpha} \right]^{1/4} E_\lambda^{\text{NB}} \left[(M_{z_0,1} + 1)^{2\overline{\alpha}} \right]^{1/2} \sum_{j,l=1}^{\infty} P_\lambda^{\text{NB}} \left(j, l \leq \zeta_1 \right)^{1/4}.$$

By Remark 6.2 the offspring distribution ξ_g has exponential moments, we therefore have that the time between regenerations of Y has finite 4α moments uniformly over $\lambda \in [a, b]$ by Proposition 4.5. That is, $\sup_{\lambda \in [a,b]} E_\lambda^{\text{NB}} \left[\zeta_1^{4\overline{\alpha}} \right] < \infty$.

Write W_n and W_n^g to be the GW-processes associated with $\mathbf{T}$ and $\mathbf{T}_g$. The number of excursions from the root is geometrically distributed with termination probability $1 - p_{ex}$ where

$$p_{ex} := \frac{W_1 - W_1^g}{W_1}.$$

Using Lemma 6.1 we therefore have that, for a constant C independent of λ,

$$E_\lambda^{\text{NB}} \left[(M_{z_0,1} + 1)^{2\overline{\alpha}} \right] \leq C \mathbb{E}[(1 - p_{ex})^{-2\overline{\alpha}}] \leq C \mathbb{E}[W_1^{2\overline{\alpha}}] < \infty$$

since $W_1 \overset{\text{d}}{=} \xi$ which has exponential moments.

It remains to show that

$$\sum_{j=1}^{\infty} \sum_{l=1}^{\infty} P_\lambda^{\text{NB}} \left(j, l \leq \zeta_1 \right)^{1/4} \qquad (6.8)$$

is finite. Note that $P_\lambda^{\text{NB}} \left(j, l \leq \zeta_1 \right) = P_\lambda^{\text{NB}} \left(\zeta_1 \geq l \right)$ whenever $l \geq j$. Using Chebyshev's inequality we can then bound (6.8) above by

$$2 \sum_{j=1}^{\infty} \sum_{l=j}^{\infty} P_\lambda^{\text{NB}} \left(\zeta_1 \geq l \right)^{1/4} \leq 2 \sum_{j=1}^{\infty} \sum_{l=j}^{\infty} \left(\frac{E_\lambda^{\text{NB}} \left[\zeta_1^u \right]}{l^u} \right)^{1/4}$$

for any integer u. In particular, we have that $\sup_{\lambda \in [a,b]} E_\lambda^{\text{NB}} \left[\zeta_1^u \right]$ is finite for any integer u by Proposition 4.5. Choosing $u > 8$ we then have that this sum is finite which completes the proof of the moment estimate of τ_1. $\square$

Finally, we complete the proof of Proposition 2.8.

Proof of Proposition 2.8: We first observe that

$$\mathbb{E}\left[E_\lambda^{\mathbf{T}^*}\left[\sigma_{e^*(\mathbf{T})}\mathbf{1}_{\{\sigma_{e^*(\mathbf{T})}<\infty\}}\right]\right] \le E_\lambda\left[1+\sum_{k=1}^{\sigma_{e^*(\mathbf{T}_g)}^Y}\chi_k \; ; \; \sigma_{e^*(\mathbf{T})}<\infty\right], \qquad (6.9)$$

where $\sigma_{e^*(\mathbf{T}_g)}^Y$ is the first time that (Y_n) returns to $e^*(\mathbf{T}_g)$. The reason why (6.9) is an inequality is that the walk (Z_n) may enter traps attached to $e(\mathbf{T}) = e(\mathbf{T}_g)$ and return to $e^*(\mathbf{T}) = e^*(\mathbf{T}_g)$ without any transitions on the backbone $\mathbf{T}_g$. It is straightforward to check (6.9) by using Lemma 4.6, Lemma 6.3 and arguments in the proof of Proposition 2.8. □

Acknowledgements

References

Aïdékon, E. Note on the monotonicity of the speed of the biased random walk on a Galton-Watson tree (2013). Unpublished note.

Aïdékon, E. Speed of the biased random walk on a Galton-Watson tree. *Probab. Theory Related Fields*, **159** (3-4), 597–617 (2014). MR3230003.

Athreya, K. B. and Ney, P. E. *Branching processes.* Dover Publications, Inc., Mineola, NY (2004). ISBN 0-486-43474-5. MR2047480.

Barlow, M. T. *Random walks and heat kernels on graphs*, volume 438 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge (2017). ISBN 978-1-107-67442-4. MR3616731.

Ben Arous, G. and Černý, J. Dynamics of trap models. In *Mathematical statistical physics*, pp. 331–394. Elsevier B. V., Amsterdam (2006). MR2581889.

Ben Arous, G. and Fribergh, A. Biased random walks on random graphs. In *Probability and statistical physics in St. Petersburg*, volume 91 of *Proc. Sympos. Pure Math.*, pp. 99–153. Amer. Math. Soc., Providence, RI (2016). MR3526827.

Ben Arous, G., Fribergh, A., Gantert, N., and Hammond, A. Biased random walks on Galton-Watson trees with leaves. *Ann. Probab.*, **40** (1), 280–338 (2012). MR2917774.

Ben Arous, G., Fribergh, A., and Sidoravicius, V. Lyons-Pemantle-Peres monotonicity problem for high biases. *Comm. Pure Appl. Math.*, **67** (4), 519–530 (2014). MR3168120.

Ben Arous, G., Hu, Y., Olla, S., and Zeitouni, O. Einstein relation for biased random walk on Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, **49** (3), 698–721 (2013). MR3112431.

Berger, N., Gantert, N., and Nagel, J. The speed of biased random walk among random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.*, **55** (2), 862–881 (2019). MR3949956.

Bowditch, A. *Biased randomly trapped random walks and applications to random walks on Galton-Watson trees.* Ph.D. thesis, University of Warwick (2017).

Bowditch, A. Escape regimes of biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, **170** (3-4), 685–768 (2018a). MR3773798.

Bowditch, A. A quenched central limit theorem for biased random walks on supercritical Galton-Watson trees. *J. Appl. Probab.*, **55** (2), 610–626 (2018b). MR3832907.

Cox, D. R. *Renewal theory.* Methuen & Co. Ltd., London; John Wiley & Sons, Inc., New York (1962). MR0153061.

Cugliandolo, L. F. and Kurchan, J. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Phys. Rev. Lett.*, **71**, 173–176 (1993). DOI: 10.1103/PhysRevLett.71.173.

Dembo, A. and Deuschel, J.-D. Markovian perturbation, response and fluctuation dissipation theorem. *Ann. Inst. Henri Poincaré Probab. Stat.*, **46** (3), 822–852 (2010). MR2682268.

Dembo, A., Gantert, N., Peres, Y., and Zeitouni, O. Large deviations for random walks on Galton-Watson trees: averaging and uncertainty. *Probab. Theory Related Fields*, **122** (2), 241–288 (2002). MR1894069.

Dembo, A., Peres, Y., and Zeitouni, O. Tail estimates for one-dimensional random walk in random environment. *Comm. Math. Phys.*, **181** (3), 667–683 (1996). MR1414305.

Fontes, L. R. G. and Mathieu, P. On the dynamics of trap models in $\mathbb{Z}^d$. *Proc. Lond. Math. Soc. (3)*, **108** (6), 1562–1592 (2014). MR3218319.

Fribergh, A. and Hammond, A. Phase transition for the speed of the biased random walk on the supercritical percolation cluster. *Comm. Pure Appl. Math.*, **67** (2), 173–245 (2014). MR3149843.

Gantert, N., Mathieu, P., and Piatnitski, A. Einstein relation for reversible diffusions in a random environment. *Comm. Pure Appl. Math.*, **65** (2), 187–228 (2012). MR2855544.

Janson, S. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, **9**, 103–252 (2012). MR2908619.

Kubo, R. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, **29** (1), 255–284 (1966). DOI: 10.1088/0034-4885/29/1/306.

Lebowitz, J. L. and Rost, H. The Einstein relation for the displacement of a test particle in a random environment. *Stochastic Process. Appl.*, **54** (2), 183–196 (1994). MR1307334.

Lyons, R., Pemantle, R., and Peres, Y. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.*, **23** (3), 1125–1138 (1995). MR1349164.

Lyons, R., Pemantle, R., and Peres, Y. Biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, **106** (2), 249–264 (1996). MR1410689.

Lyons, R. and Peres, Y. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York (2016). ISBN 978-1-107-16015-6. MR3616205.

Mathieu, P. Differentiating the entropy of random walks on hyperbolic groups. *Ann. Probab.*, **43** (1), 166–187 (2015). MR3298471.

Peres, Y. and Zeitouni, O. A central limit theorem for biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, **140** (3-4), 595–629 (2008).

MR2365486.

Piau, D. Théorème central limite fonctionnel pour une marche au hasard en environnement aléatoire. *Ann. Probab.*, **26** (3), 1016–1040 (1998). MR1634413.

Sznitman, A.-S. Slowdown estimates and central limit theorem for random walks in random environment. *J. Eur. Math. Soc. (JEMS)*, **2** (2), 93–143 (2000). MR1763302.