J. Appl. Prob. **55,** 610–626 (2018)
*doi:*10.1017/jpr.2018.38
© *Applied Probability Trust* 2018

A QUENCHED CENTRAL LIMIT THEOREM FOR BIASED RANDOM WALKS ON SUPERCRITICAL GALTON–WATSON TREES

ADAM BOWDITCH,* *University of Warwick*

Abstract

In this paper we prove a quenched functional central limit theorem for a biased random walk on a supercritical Galton–Watson tree with leaves. This extends a result of Peres and Zeitouni (2008) where the case without leaves was considered. A conjecture of Ben Arous and Fribergh (2016) suggests an upper bound on the bias which we observe to be sharp.

Keywords: Random walk; random environment; Galton–Watson tree; functional central limit theorem; quenched; invariance principle

2010 Mathematics Subject Classification: Primary 60F17
Secondary 60K37; 60F05; 60J80

1. Introduction

We investigate biased random walks on supercritical Galton–Watson (GW) trees with leaves and prove a quenched functional central limit theorem (CLT). This extends the main result of [11] where GW trees without leaves were considered. A GW tree conditioned to survive consists of an infinite backbone with finite trees attached as branches. The backbone is a GW tree whose offspring law does not have deaths. The branches are formed by attaching a random number of independent subcritical GW trees to each vertex of the backbone. This forms dead-ends in the environment which makes it a natural setting for observing trapping as the walk is slowed by taking excursions in the branches of the tree. The number of trees attached to a given vertex on the backbone has a distribution depending on the backbone locally. This means that the branches are not independent and identically distributed (i.i.d.) and, therefore, the trapping incurred by the addition of the leaves demonstrates a significant complication.

The influence of the bias on the trapping is an important feature of the model. As the bias is increased, the local drift away from the root will increase but this does not necessarily speed up the walk. This is because it increases the time trapped in the finite leaves from which the walk cannot escape without taking long sequences of movements against the bias. Lyons *et al.* [10] showed that, for a suitably large bias, the trapping is sufficient to slow the walk to zero speed whereas, for a small bias, the expected trapping time is finite and the walk has a positive limiting speed. Under a further restriction on the bias, the trapping times have finite variance; we use this to prove a quenched invariance principle for the walk.

We next briefly describe the supercritical GW tree conditioned on survival; for more detail see [1, Section I.12]. Let $\{p_k\}$ denote the offspring distribution of a GW process Z_n with a single progenitor, mean $\mu > 1$, and probability generating function f. The process Z_n gives

Received 14 February 2017; revision received 27 February 2018.
* Current address: Department of Mathematics, National University of Singapore, Block S17, 10 Lower Kent Ridge Road, Singapore 119076. Email address: matamb@nus.edu.sg

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

rise to a random tree $\mathcal{T}^f$ where individuals are represented by vertices and edges connect individuals with their offspring. We denote by ρ the root, which is the vertex representing the unique progenitor. Let q denote the extinction probability of Z_n which is nonzero only when $p_0 > 0$. In this case, we then define

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

which are generating functions of a GW process without deaths and a subcritical GW process, respectively. An f-GW tree conditioned on nonextinction $\mathcal{T}$ can be generated by first generating a g-GW tree $\mathcal{T}^g$ and then, to each vertex x of $\mathcal{T}^g$, appending a random number $\mathcal{M}_x$ of independent h-GW trees. We refer to $\mathcal{T}^g$ as the backbone of $\mathcal{T}$, the finite trees appended to $\mathcal{T}^g$ as the traps, and the vertices in the first generation of the traps as the buds; see Figure 1.

We now introduce the biased random walk on a fixed tree. For a fixed tree $\mathfrak{T}$ with $x \in \mathfrak{T}$, let $\overleftarrow{x}$ denote the parent of x and $c(x)$ the set of children of x. A β-biased random walk on $\mathfrak{T}$ is a random walk $(X_n)_{n\geq 0}$ on the vertices of $\mathfrak{T}$ defined by $P_z^{\mathfrak{T}}(X_0 = z) = 1$ and the transition probabilities

$$P_z^{\mathfrak{T}}(X_{n+1} = y \mid X_n = x) = \begin{cases} \dfrac{1}{1+\beta|c(x)|} & \text{if } y = \overleftarrow{x}, \\ \dfrac{\beta}{1+\beta|c(x)|} & \text{if } y \in c(x),\ x \neq \rho, \\ \dfrac{1}{|c(\rho)|} & \text{if } y \in c(x),\ x = \rho, \\ 0 & \text{otherwise.} \end{cases}$$

We use $\mathbb{P}_\rho(\cdot) := \int P_\rho^{\mathcal{T}}(\cdot)\boldsymbol{P}(\mathrm{d}\mathcal{T})$ for the annealed law obtained by averaging the quenched law $P_\rho^{\mathcal{T}}$ over the law $\boldsymbol{P}$ on f-GW trees conditioned to survive. Unless indicated otherwise, we start the walk at ρ.

For $x \in \mathcal{T}$, let $|x| := d(\rho, x)$ denote the distance between x and the root of the tree. Lyons *et al.* [10] showed that if $\beta \in (\mu^{-1}, f'(q)^{-1})$ then $|X_n|n^{-1}$ converges $\mathbb{P}$-a.s. ($\mathbb{P}$-almost surely)

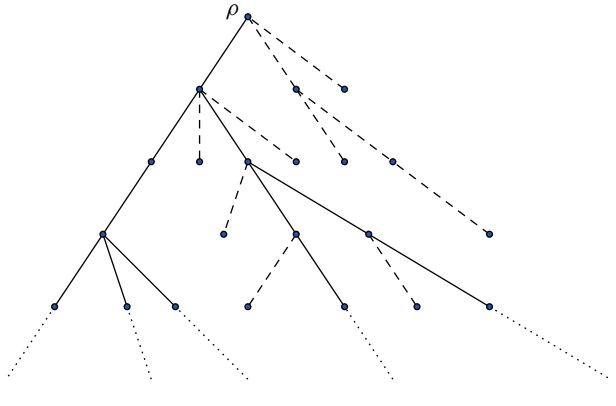


FIGURE 1: A section of a tree $\mathcal{T}$ rooted at ρ with backbone $\mathcal{T}^g$ represented by solid edges and branches depicted by dashed lines.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

to a deterministic constant $v > 0$ called the speed of the walk. When $\beta < \mu^{-1}$, the walk is recurrent and $|X_n|n^{-1}$ converges $\mathbb{P}$-a.s. to 0. When the bias is large, the walk is transient but slowed by having to make long sequences of movements against the bias in order to escape the traps; in particular, if $\beta \geq f'(q)^{-1}$ then the slowing affect is strong enough to cause $|X_n|n^{-1}$ to converge $\mathbb{P}$-a.s. to 0. This regime has received further attention in [3] and [6], where polynomial scaling results can be found.

For $\sigma, t > 0$ and $n = 1, 2, \ldots$, define

$$B_t^n := \frac{|X_{\lfloor nt \rfloor}| - ntv}{\sigma\sqrt{n}}.$$

Our main result, Theorem 1.1, is a quenched invariance principle for B_t^n. We denote by $D([0, \infty), \mathbb{R})$ the space of càdlàg functions mapping $[0, \infty)$ to $\mathbb{R}$ which we equip with the Skorokhod J_1-topology.

Theorem 1.1. *Suppose that $p_0 > 0$, $\mu > 1$, $\beta \in (\mu^{-1}, f'(q)^{-1/2})$, and there exists $\lambda > 1$ such that*

$$\sum_{k \geq 0} \lambda^k p_k < \infty. \tag{1.1}$$

There exists $\sigma > 0$ such that, for $\mathbf{P}$-a.e. ($\mathbf{P}$-almost every) $\mathcal{T}$, the process $(B_t^n)_{t \geq 0}$ converges in $P^{\mathcal{T}}$-distribution on $D([0, \infty), \mathbb{R})$ to a standard Brownian motion.

The condition $p_0 > 0$ ensures that the tree has leaves; the case without leaves ($p_0 = 0$) was considered in [11]. This no longer requires the condition $\beta < f'(q)^{-1/2}$, which is due to the trapping in the dead-ends caused by the leaves. Indeed, when $p_0 = 0$, we have $q = 0$ and $f'(0) = 0$, therefore, the condition becomes irrelevant. This regime was studied further by Peres and Zeitouni [11] in the case where $\beta = \mu^{-1}$ and $p_0 = 0$; in this setting $v = 0$ and it was shown that B_t^n converges in distribution to the absolute value of a Brownian motion. This result was extended by Dembo and Sun [7] to random walks on multi-type GW trees with leaves. Although the dead-ends in this model trap the walk, the bias is small and therefore the slowing is weak.

We allow the tree to have leaves; this creates traps in the environment which slow the walk. This trapping becomes more significant as the bias is increased to the extent that when $\beta > f'(q)^{-1/2}$ the trapping times in the finite branches have infinite variance and a CLT does not hold.

For our main result, we assume that the exponential moments condition (1.1) holds. This is a purely technical assumption which we expect could be relaxed to a sufficiently large moment condition, however, the main focus of this paper is to obtain the optimal upper bound on the bias.

In [5], the second moment condition for trapping times in finite trees was used to prove an annealed invariance principle and quenched CLT for a biased random walk on a subcritical GW tree conditioned to survive. In that model the backbone is one-dimensional but the branches have a similar distribution to those in the supercritical trees we consider here. Due to the one-dimensional backbone, the fluctuations are significantly influenced by the specific instance of the environment. This results in an environment-dependent centring for the walk in the quenched CLT. In the supercritical case, the walk will randomly choose one of infinitely many escape routes; this creates a mixing of the environment which yields a deterministic centring in the quenched result Theorem 1.1.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

We begin, in Section 2, by proving an annealed functional CLT for the walk by adapting the renewal argument used in [12, Theorem 4.1]. This is then extended to the quenched result Theorem 1.1 in Section 3 by applying the argument used in [4] which largely involves showing that multiple copies of the walk see sufficiently different areas of the tree. This confirms [2, Conjecture 3.1]. We conclude, in Section 4, that the upper bound on the bias is sharp (except for a single technical case that we omit).

2. An annealed invariance principle

The aim of this section is to prove the annealed analogue of Theorem 1.1. We first show that the time spent in a single trap has finite variance. Following this, we extend the result to show that the times between regenerations of the random walk have finite variance. By exploiting the regeneration structure, we are then able to approximate B_t^n by an i.i.d. sum of terms with finite variance. The annealed functional CLT then follows from standard results concerning i.i.d. sums.

We now show that the time spent in a trap has finite variance. Let $\overline{\mathcal{T}}^h$ be the tree formed by attaching an additional vertex $\overline{\rho}$ (as the parent of the root ρ) to an h-GW tree $\mathcal{T}^h$. For a fixed tree $\mathfrak{T}$ and vertex $x \in \mathfrak{T}$, let $\tau_x^+ := \inf\{k > 0 \colon X_k = x\}$ denote the first return time to x. Let ξ^f, ξ^g, and 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 $\mathcal{T}$. Since the generation sizes of $\mathcal{T}^g$ are dominated by those of $\mathcal{T}$, it follows that ξ^g is stochastically dominated by ξ. Using Bayes' law, we have $\boldsymbol{P}(\xi = k) = p_k(1-q^k)(1-q)^{-1} \le cp_k$, therefore, both ξ and ξ^g inherit the moment bounds of ξ^f. Furthermore, $\boldsymbol{P}(\xi^h = k) = p_k q^k$; therefore, ξ^h automatically has exponential moments.

Lemma 2.1. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta \in (\mu^{-1}, f'(q)^{-1/2})$. Then we have*

$$\boldsymbol{E}[E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[(\tau_{\overline{\rho}}^+)^2]] < \infty.$$

Proof. We can write

$$\tau_{\overline{\rho}}^+ = \sum_{x \in \overline{\mathcal{T}}^h} v_x,$$

where $v_x = \sum_{k=1}^{\tau_{\overline{\rho}}^+} \mathbf{1}_{\{X_k = x\}}$ is the number of visits to x before returning to $\overline{\rho}$ with $\mathbf{1}_A$ the indicator function of the event A. Recall that $c(x)$ denotes the set of children of x. It then follows that

$$\begin{aligned}
\boldsymbol{E}[E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[(\tau_{\overline{\rho}}^+)^2]] &= \boldsymbol{E}\Bigg[\sum_{x,y \in \overline{\mathcal{T}}^h} E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[v_x v_y]\Bigg] \\
&\le C_\beta \boldsymbol{E}\Bigg[\sum_{x,y \in \overline{\mathcal{T}}^h} (|c(x)|\beta + 1)(|c(y)|\beta + 1)\beta^{|x|+|y|}\Bigg] \\
&= C_\beta \boldsymbol{E}\Bigg[\sum_{x \in \overline{\mathcal{T}}^h} (|c(x)|\beta + 1)\beta^{|x|} \sum_{y \in \overline{\mathcal{T}}^h} (|c(y)|\beta + 1)\beta^{|y|}\Bigg], \qquad (2.1)
\end{aligned}$$

where the inequality follows from [5, Lemma 4.4]. Letting Z_k^h denote the size of the kth generation of $\overline{\mathcal{T}}^h$ and collecting terms in each generation, we have

$$\sum_{x \in \overline{\mathcal{T}}^h} (|c(x)|\beta + 1)\beta^{|x|} = 1 + \sum_{k \ge 1} Z_k^h(\beta^k + \beta^{k-1}) \le (1 + \beta^{-1}) \sum_{k \ge 0} Z_k^h \beta^k.$$

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

By [5, Lemma 4.1], since ξ^h has exponential moments, we have $\boldsymbol{E}[Z_k^h Z_j^h] \le Cf'(q)^j$ whenever $j \ge k$. Substituting this and the above inequality into (2.1), we have

$$\begin{aligned}\boldsymbol{E}[E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[(\tau_{\overline{\rho}}^+)^2]] &\le C_\beta \sum_{k\ge 0} \beta^k \sum_{j\ge k} \boldsymbol{E}[Z_k^h Z_j^h]\beta^j \\ &\le C_\beta \sum_{k\ge 0} \beta^k \sum_{j\ge k} (f'(q)\beta)^j \\ &\le C_{\beta, f'(q)} \sum_{k\ge 0} (f'(q)\beta^2)^k,\end{aligned}$$

which is finite by the assumption that $\beta < f'(q)^{-1/2}$. □

Let $r(0) := 0$, $r(n) := \inf\{k > r(n-1)\colon X_k, X_{k-1} \in \mathcal{T}^g\}$ for $n \ge 1$, and $Y_n := X_{r(n)}$. Then Y_n is a β-biased random walk on $\mathcal{T}^g$ coupled to X_n. Write $\zeta_0^Y := 0$ and, for $m = 1, 2, \ldots$, let

$$\zeta_m^Y := \inf\{k > \zeta_{m-1}^Y \colon |Y_j| < |Y_k| \le |Y_l| \text{ for all } j < k \le l\}$$

be regeneration times for the walk Y. We can then define $\zeta_k^X := \inf\{m \ge 0\colon X_m = Y_{\zeta_k^Y}\}$ to be the corresponding regeneration times for X and $\varrho_k := X_{\zeta_k^X}$ the regeneration points. From [10, Proposition 3.4], it follows that there exists, $\boldsymbol{P}$-a.s., an infinite sequence of regeneration times $\{\zeta_k^X\}_{k\ge 1}$ and the sequence

$$\{(\zeta_{k+1}^X - \zeta_k^X), (|\varrho_{k+1}| - |\varrho_k|)\}_{k\ge 1}$$

is i.i.d. (as is the corresponding sequence for Y). Furthermore, letting

$$\varphi_t := \sup\{j \ge 0\colon \zeta_j^X \le t\} \tag{2.2}$$

be the number of regenerations by time t, we see that φ_t is nondecreasing and diverges $\mathbb{P}$-a.s.

By [10, Theorems 3.1 and 4.1], whenever $\mu > 1$ and $\mu^{-1} < \beta < f'(q)^{-1}$, we see that the time and distance between regenerations of X both have finite means with respect to $\mathbb{P}$ and there exists $\nu \in (0, 1)$ such that $|X_n|n^{-1}$ converges $\mathbb{P}$-a.s. to ν. Let

$$\chi_j := |X_{\zeta_j^X}| - |X_{\zeta_{j-1}^X}| - \nu(\zeta_j^X - \zeta_{j-1}^X) = |\varrho_j| - |\varrho_{j-1}| - \nu(\zeta_j^X - \zeta_{j-1}^X).$$

From the previous remark, it follows that χ_j are i.i.d. with respect to $\mathbb{P}$. By the strong law of large numbers and the definition of ν, we see that χ_j are centred (see [10, Theorems 3.1 and 4.1]). We will show that the χ_j have a finite second moment and their sum

$$\Sigma_m := \sum_{j=2}^m \chi_j = (|X_{\zeta_m^X}| - \nu\zeta_m^X) - (|\varrho_1| - \nu\zeta_1^X)$$

can be used to approximate B_t^n.

By the remark preceding Lemma 2.1, the offspring distribution ξ^g of $\mathcal{T}^g$ has exponential moments. Since Y is a random walk on $\mathcal{T}^g$, by [11, Proposition 3], we have $\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^k] < \infty$ for all $k \in \mathbb{Z}$ whenever $\beta > \mu^{-1}$.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

Let $\eta_k := r(k+1) - r(k)$ denote the total time taken between X making the kth and $(k+1)$th transition along the backbone. This time consists of

$$N_k := \sum_{j=r(k)+1}^{r(k+1)} \mathbf{1}_{\{X_j = Y_k\}}$$

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

$$\eta_k := 1 + \sum_{j=1}^{N_k} \gamma_{k,j}, \tag{2.3}$$

where $\gamma_{k,j} := \tau_k^{(j)} - \tau_k^{(j-1)}$ is the duration of the jth such excursion.

Proposition 2.1. *Under the assumptions of Theorem 1.1, we have*

$$\mathbb{E}[(\zeta_2^X - \zeta_1^X)^2] < \infty.$$

Proof. The law of $\zeta_2^X - \zeta_1^X$ under $\mathbb{P}$ is identical to its law under $\mathbb{P}_\rho(\cdot \mid \zeta_1^Y = 1)$. That is, by the independence structure, we can condition the first regeneration vertex to be the first vertex reached by Y without changing the law of $\zeta_2^X - \zeta_1^X$. Therefore, we see that $\mathbb{E}[(\zeta_2^X - \zeta_1^X)^2]$ can be written as

$$\mathbb{E}[(\zeta_2^X - \zeta_1^X)^2 \mid \zeta_1^Y = 1] = \mathbb{E}\left[\left(\sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} \eta_k\right)^2 \,\middle|\, \zeta_1^Y = 1\right] \leq \mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} \eta_k^2 \,\middle|\, \zeta_1^Y = 1\right]$$

by convexity. Using convexity again with the decomposition (2.3), we can write this as

$$\begin{aligned}
&\mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} \left(1 + \sum_{j=1}^{N_k} \gamma_{k,j}\right)^2 \,\middle|\, \zeta_1^Y = 1\right] \\
&\qquad \leq \mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} (N_k + 1)\left(1 + \sum_{j=1}^{N_k} \gamma_{k,j}^2\right) \,\middle|\, \zeta_1^Y = 1\right].
\end{aligned}$$

The excursion times $\gamma_{k,j}$ are distributed as the first return time to $\overline{\rho}$ for a walk started from $\overline{\rho}$ on $\overline{\mathcal{T}}^h$. Moreover, under $\mathbb{P}$ 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

$$\boldsymbol{E}[E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[(\tau_{\overline{\rho}}^+)^2]]\mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} (N_k + 1)^2 \,\middle|\, \zeta_1^Y = 1\right],$$

where, by Lemma 2.1, we have $\boldsymbol{E}[E_{\overline{\rho}}^{\overline{\mathcal{T}}^h}[(\tau_{\overline{\rho}}^+)^2]] < \infty$.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

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

$$\mathcal{L}(z, j) \coloneqq \sum_{i=0}^{j} \mathbf{1}_{\{Y_j = z\}}, \qquad \mathcal{L}(z) \coloneqq \mathcal{L}(z, \infty)$$

the local times of the vertex z. Write

$$W_{z,l} \coloneqq \sum_{j=0}^{\infty} \mathbf{1}_{\{X_j = z,\, X_{j+1} \notin \mathcal{T}^g,\, \mathcal{L}(z,j) = l\}}$$

for the number of excursions from z (by X) on the lth visit to z (by Y) for $l = 1, \ldots, \mathcal{L}(z)$ and $M \coloneqq |\{Y_k\}_{k=1}^{\zeta_2^Y - 1}|$ the number of distinct vertices visited by Y between time 1 and time $\zeta_2^Y - 1$. Then

$$\begin{aligned}
&\mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{\zeta_2^Y - \zeta_1^Y} (N_k + 1)^2 \mid \zeta_1^Y = 1\right] \\
&\quad = \mathbb{E}\left[(\zeta_2^Y - \zeta_1^Y) \sum_{k=1}^{M} \sum_{l=1}^{\mathcal{L}(z_k)} (W_{z_k,l} + 1)^2 \mid \zeta_1^Y = 1\right] \\
&\quad = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mathbb{E}[(\zeta_2^Y - \zeta_1^Y)\, \mathbf{1}_{\{k \le M,\, l \le \mathcal{L}(z_k)\}} (W_{z_k,l} + 1)^2 \mid \zeta_1^Y = 1] \\
&\quad \le \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} (\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^2\, \mathbf{1}_{\{k \le M,\, l \le \mathcal{L}(z_k)\}} \mid \zeta_1^Y = 1] \mathbb{E}[(W_{z_k,l} + 1)^4 \mid \zeta_1^Y = 1])^{1/2} \quad (2.4)
\end{aligned}$$

from the Cauchy–Schwarz inequality. Conditional on $\zeta_1^Y = 1$, for all $1 \le k \le M$, we have $\mathcal{L}(z_k) \le \zeta_2^Y - \zeta_1^Y$; moreover, $M \le \zeta_2^Y - \zeta_1^Y$. Therefore,

$$\mathbf{1}_{\{k \le M,\, l \le \mathcal{L}(z_k)\}} \le \mathbf{1}_{\{k, l \le \zeta_2^Y - \zeta_1^Y\}}.$$

Since the root does not have a parent, without any further information concerning the number of children from a given vertex, we see that the walk is more likely to take an excursion into one of the neighbouring traps when at the root than from this vertex. We can, therefore, stochastically dominate the number of excursions from a vertex by the number of excursions from the root to see that $\mathbb{E}[(W_{z_k,l} + 1)^4] \le \mathbb{E}[(W_{z_0,1} + 1)^4]$. Using this, the Cauchy–Schwarz inequality, and the fact that $\mathbb{P}(\zeta_1^Y = 1) > 0$, (2.4) can be bounded above by

$$\begin{aligned}
&\mathbb{P}(\zeta_1^Y = 1)^{-1} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} (\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^2\, \mathbf{1}_{\{k, l \le \zeta_2^Y - \zeta_1^Y\}}] \mathbb{E}[(W_{z_k,l} + 1)^4])^{1/2} \\
&\qquad \le C \mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^4]^{1/4} \mathbb{E}[(W_{z_0,1} + 1)^4]^{1/2} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \mathbb{P}(k, l \le \zeta_2^Y - \zeta_1^Y)^{1/4}.
\end{aligned}$$

Since the offspring distribution ξ^g has exponential moments, we see that the time between regenerations has finite fourth moments according to [11, Proposition 3]. That is, $\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^4] < \infty$.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

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

$$p_{\mathrm{ex}} := \frac{Z_1 - Z_1^g}{Z_1}.$$

Using properties of geometric random variables, we therefore have

$$\mathbb{E}[(W_{z_0,1} + 1)^4] \le C\mathbb{E}[(1 - p_{\mathrm{ex}})^{-4}] \le C\mathbb{E}[Z_1^4] < \infty$$

since $Z_1 \overset{\mathrm{D}}{=} \xi$, which has exponential moments (where '$\overset{\mathrm{D}}{=}$' denotes equality in distribution).

It remains to show that

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\mathbb{P}(k, l \le \zeta_2^Y - \zeta_1^Y)^{1/4} \tag{2.5}$$

is finite. Note that $\mathbb{P}(k, l \le \zeta_2^Y - \zeta_1^Y) = \mathbb{P}(\zeta_2^Y - \zeta_1^Y \ge l)$ whenever $l \ge k$. Using Chebyshev's inequality, we can then bound (2.5) above by

$$2\sum_{k=1}^{\infty}\sum_{l=k}^{\infty}\mathbb{P}(\zeta_2^Y - \zeta_1^Y \ge l)^{1/4} \le 2\sum_{k=1}^{\infty}\sum_{l=k}^{\infty}\left(\frac{\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^j]}{l^j}\right)^{1/4}$$

for any integer j. In particular, from [11, Proposition 3], we see that $\mathbb{E}[(\zeta_2^Y - \zeta_1^Y)^j]$ is finite for any integer j. Choosing $j > 8$, we then find that this sum is finite, which completes the proof. □

Recall from (2.2) that φ_t is the number of regenerations by time t. We now show that $\Sigma_{\varphi_{nt}}$ approximates B_t^n. For $x \in \mathcal{T}$, let $\mathcal{T}_x$ denote the subtree consisting of all descendants of x in $\mathcal{T}$. Then, for $y \in \mathcal{T}^g$, let $\mathcal{T}_y^-$ be the branch at y; that is, the subtree rooted at y consisting only of y, the children of y not on $\mathcal{T}^g$, and their descendants; see Figure 2. The tree $\mathcal{T}_y^-$ then has the law of a tree rooted at y with some random number $\mathcal{M}_y$ of h-GW trees attached to y. Since $\mathcal{M}_y$ is dominated by ξ, from (1.1) we see that $\mathcal{M}_y$ has exponential moments. It therefore follows from [9, Theorem B] that there exists some constant C such that

$$\boldsymbol{P}(\mathcal{H}(\mathcal{T}_y^-) \ge n) \le Cf'(q)^n, \tag{2.6}$$

where, for a fixed rooted tree $\mathfrak{T}$, $\mathcal{H}(\mathfrak{T}) := \sup\{d(\rho, x)\colon x \in \mathfrak{T}\}$ is the height of $\mathfrak{T}$. Let $\mathcal{H}_n := \max\{\mathcal{H}(\mathcal{T}_y^-)\colon y \in \{Y_k\}_{k=0}^n\}$ denote the largest branch seen by Y by time n. It follows that

$$\sup_{t\in[0,T]}\left|B_t^n - \frac{\Sigma_{\varphi_{tn}}}{\sigma\sqrt{n}}\right| \le \frac{|\varrho_1| + \nu\zeta_1^X + \mathcal{H}_{nT}}{\sigma\sqrt{n}} + \sup_{j=1,\dots,\varphi_{nT}}\frac{|\varrho_{j+1}| - |\varrho_j| + \nu(\zeta_{j+1}^X - \zeta_j^X)}{\sigma\sqrt{n}}. \tag{2.7}$$

Up to time nT, the walk Y can only have visited at most nT vertices on $\mathcal{T}^g$, therefore, the probability that X has visited a branch of height at least $C\log(n)$ by time nT is at most $C_T n f'(q)^{C\log(n)}$. In particular, by the Borel–Cantelli lemma, choosing suitably large C, we see that there are almost surely only finitely many n such that Y has visited the root of a branch of height at least $C\log(n)$ by time nT. Since ϱ_1 and ζ_1^X do not depend on n and have finite mean, it follows that the first term in (2.7) converges $\mathbb{P}$-a.s. to 0.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

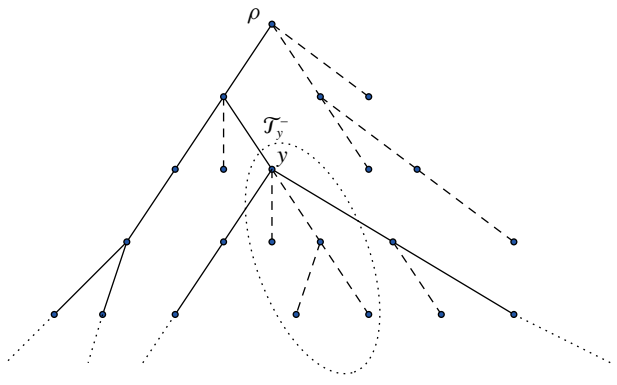


FIGURE 2: A section of a tree $\mathcal{T}$ rooted at ρ with the branch $\mathcal{T}_y^-$ rooted at y indicated by a dotted ellipse.

From [11, Proposition 3], for any $k \in \mathbb{Z}^+$, we have $\mathbb{E}[(|\varrho_2| - |\varrho_1|)^k] < \infty$; therefore, the distance between regeneration points is small. In particular, bounding φ_{nT} above by nT, using a union bound, and Markov's inequality, we have, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{j=1,\dots,\varphi_{nT}} \frac{|\varrho_{j+1}| - |\varrho_j|}{\sigma\sqrt{n}} > \varepsilon \right) \leq C_{T,\varepsilon} \mathbb{E}[(|\varrho_2| - |\varrho_1|)^2 \mathbf{1}_{\{|\varrho_2|-|\varrho_1|>\varepsilon\sqrt{n}\}}]$$

which converges to 0 as $n \to \infty$ by dominated convergence. Similarly, using Proposition 2.1, we see that the same holds for the supremum of $\zeta_{j+1}^X - \zeta_j^X$; therefore,

$$\mathbb{P}\left(\sup_{t\in[0,T]} \left| B_t^n - \frac{\Sigma_{\varphi_{tn}}}{\sigma\sqrt{n}} \right| > \varepsilon \right)$$

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

Lemma 2.2. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta \in (\mu^{-1}, f'(q)^{-1})$. Then ζ_n^X/n converges $\mathbb{P}$-a.s. to the deterministic process $t\mathbb{E}[\zeta_2 - \zeta_1]$ and φ_{nt}/n converges $\mathbb{P}$-a.s. to the deterministic linear process $t/\mathbb{E}[\zeta_2 - \zeta_1]$.*

Proof. Since the differences $\zeta_j^X - \zeta_{j-1}^X$ are i.i.d. with finite mean, from the law of large numbers and the fact that ζ_1^X/n converges $\mathbb{P}$-a.s. to 0, we have

$$\frac{\zeta_{nt}^X}{n} = \frac{\zeta_1^X}{n} + \sum_{k=2}^{nt} \frac{(\zeta_k^X - \zeta_{k-1}^X)}{n}$$

converges $\mathbb{P}$-a.s. to the desired limit. It therefore follows by continuity of the inverse at strictly increasing functions (see [13, Corollary 13.6.4]) that φ_{nt}/n converges $\mathbb{P}$-a.s. as stated. □

By Proposition 2.1 and the remark leading to it, we see that Σ_m is the sum of i.i.d. centred random variables with finite second moment. By Donsker's theorem, it therefore follows that $(\Sigma_{nt}/\sqrt{n})_{t\geq 0}$ converges to a scaled Brownian motion. By continuity of composition at continuous limits (see [13, Theorem 13.2.1]) and the previous remarks, we therefore have the following annealed CLT.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

Corollary 2.1. *Under the assumptions of Theorem 1.1, there exists a constant $\sigma^2 > 0$ such that the process*

$$B_t^n := \frac{|X_{\lfloor nt \rfloor}| - nt\nu}{\sigma\sqrt{n}}$$

converges in $\mathbb{P}$-distribution on $D([0,\infty),\mathbb{R})$ to a standard Brownian motion.

3. A quenched invariance principle

We now extend Corollary 2.1 to a quenched functional CLT. We do this by adapting the technique used in [4] and [11], which exploited the fact that independent copies of the walk on the same tree visit disjoint regions of the tree at large times. For each $n \in \mathbb{N}$, write $\mathbb{B}_t^n(X)$ to be the linear interpolation satisfying

$$\mathbb{B}_{k/n}^n(X) = \frac{|X_k| - k\nu}{\sigma\sqrt{n}} \quad \text{for } k \in \mathbb{N}.$$

We then have $B_t^n = \mathbb{B}_t^n$ for $t > 0$ such that $nt \in \mathbb{N}$ and $|B_t^n - \mathbb{B}_t^n| \le n^{-1/2}(\nu+1)/\sigma$; therefore, it suffices to consider the interpolation. Lemma 3.1 yields a sufficient condition for proving Theorem 1.1 and follows from [4, Lemma 4.1]; therefore, we omit the proof.

Lemma 3.1. *Suppose that the assumptions of Theorem 1.1 hold and that, for any bounded Lipschitz function $F\colon C([0,T],\mathbb{R}) \to \mathbb{R}$ and $b \in (1,2)$, we have*

$$\sum_{k\ge 1} \operatorname{var}_{\boldsymbol{P}}(E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k \rfloor})]) < \infty. \tag{3.1}$$

Then, for $\boldsymbol{P}$-a.e. $\mathcal{T}$, the process $(B_t^n)_{t\ge 0}$ converges in $P^{\mathcal{T}}$-distribution on $D([0,\infty),\mathbb{R})$ to a standard Brownian motion.

We now complete the proof of the quenched functional CLT by following the method used in [11] to show that condition (3.1) holds for any bounded Lipschitz function $F\colon C([0,T],\mathbb{R}) \to \mathbb{R}$ and $b \in (1,2)$ under the assumptions of the theorem.

Proof of Theorem 1.1. For a fixed tree $\mathcal{T}$, let X^1 and X^2 be independent β-biased walks on $\mathcal{T}$, and Y^1 and Y^2 the corresponding backbone walks. For $i = 1, 2$, $k \in \mathbb{N}$, and $t, s \ge 0$, let

$$\mathbb{B}_{t,s}^{k,i} = \mathbb{B}_t^{\lfloor b^k \rfloor}(X_{\cdot+s}^i) - \mathbb{B}_s^{\lfloor b^k \rfloor}(X_\cdot^i)$$

be a random variable with law of the interpolation $\mathbb{B}^{\lfloor b^k \rfloor}$ started from the vertex X_s^i. Define

$$\vartheta_k^{Y^i} := \min\{m > \lfloor b^{k/4} \rfloor : m \in \{\zeta_j^{Y^i}\}_{j\ge 1}\} \quad \text{and} \quad \vartheta_k^{X^i} = \min\left\{m \ge 0 \colon X_m^i = Y^i_{\vartheta_k^{Y^i}}\right\}$$

to be the first regeneration time of Y^i after time $\lfloor b^{k/4} \rfloor$ and the corresponding time for X^i.

Let

$$\mathcal{A}_k^1 := \left\{Y^2_{\vartheta_k^{Y^2}} \notin \{Y_s^1 : s \le \vartheta_k^{Y^1}\}\right\} = \left\{X^2_{\vartheta_k^{X^2}} \notin \{X_s^1 : s \le \vartheta_k^{X^1}\}\right\},$$
$$\mathcal{A}_k^2 := \left\{Y^1_{\vartheta_k^{Y^1}} \notin \{Y_s^2 : s \le \vartheta_k^{Y^2}\}\right\} = \left\{X^1_{\vartheta_k^{X^1}} \notin \{X_s^2 : s \le \vartheta_k^{X^2}\}\right\},$$

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

and $\mathcal{A}_k := \mathcal{A}_k^1 \cap \mathcal{A}_k^2$ be the event that, after the first regeneration times after time $\lfloor b^{k/4} \rfloor$, the paths of Y^1 and Y^2 do not intersect. Write $\mathcal{B}^{k,i} := \{\vartheta_k^{Y^i} \le b^{k/3}\}$ to be the event that the first regeneration after time $b^{k/4}$ happens before time $b^{k/3}$.

Recall that for $x \in \mathcal{T}^g$ we denote by $\mathcal{H}(\mathcal{T}_x^-)$ the height of the branch attached to the vertex x. Using Lipschitz properties of $\mathbb{B}^{k,i}$, we have

$$\begin{aligned}\sup_{t\le T}\left|\mathbb{B}^{k,i}_{t,0} - \mathbb{B}^{k,i}_{t,\vartheta_k^{X^i}}\right| &\le \sup_{m\le Tb^k} b^{-k/2}\left||X^i_m| - m\nu - \left|X^i_{m+\vartheta_k^{X^i}}\right| + (m+\vartheta_k^{X^i})\nu + \left|X^i_{\vartheta_k^{X^i}}\right| - \vartheta_k^{X^i}\nu\right| \\ &= \sup_{m\le Tb^k} b^{-k/2}\left||X^i_m| - \left|X^i_{m+\vartheta_k^{X^i}}\right| + \left|X^i_{\vartheta_k^{X^i}}\right|\right| \\ &\le b^{-k/2}\max_{m\le Tb^k}\left|\left||Y^i_m| - \left|Y^i_{m+\vartheta_k^{Y^i}}\right|\right| + \left|Y^i_{\vartheta_k^{Y^i}}\right|\right| + b^{-k/2}\mathcal{H}^i_{Tb^k},\end{aligned}$$

where $\mathcal{H}^i_{Tb^k}$ is the height of the tallest branch seen up to time Tb^k by Y^i. Up to time Tb^k, the walk Y^i can visit at most $Tb^k + 1$ unique vertices. At the first hitting time of a vertex, the bud and backbone distribution from this vertex are independent of the past; therefore, by (2.6),

$$\begin{aligned}\mathbb{P}(\mathcal{H}^i_{Tb^k} \ge C\log(b^k)) &\le C_T b^k \boldsymbol{P}(\mathcal{H}(\mathcal{T}_\rho^-) \ge C\log(b^k)) \\ &\le C_T b^k f'(q)^{C\log(b^k)} \\ &\le C_T b^{-k} \quad \text{for sufficiently large } C. \end{aligned} \tag{3.2}$$

Furthermore, by the Lipschitz property of Y^i, we have

$$b^{-k/2}\max_{m\le Tb^k}\left|\left||Y^i_m| - \left|Y^i_{m+\vartheta_k^{Y^i}}\right|\right| + \left|Y^i_{\vartheta_k^{Y^i}}\right|\right| \le 2\vartheta_k^{Y^i} b^{-k/2},$$

which is bounded above by $2b^{-k/6}$ on the event $\mathcal{B}^{k,i}$. Letting $\mathcal{C}^{k,i} := \{\mathcal{H}^i_{Tb^k} < C\log(b^k)\}$, we then have, on the event $\mathcal{B}^{k,i} \cap \mathcal{C}^{k,i}$,

$$\left|F(\mathbb{B}^{k,i}_{\cdot,0}) - F\left(\mathbb{B}^{k,i}_{\cdot,\vartheta_k^{X^i}}\right)\right| \le Cb^{-k/6}$$

for any bounded Lipschitz function $F\colon C([0,T],\mathbb{R}) \to \mathbb{R}$.

Using the Lipschitz and boundedness properties of F, we then have

$$\begin{aligned}\operatorname{var}_{\boldsymbol{P}}(E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k\rfloor})]) &= \boldsymbol{E}[E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k\rfloor})]^2] - \boldsymbol{E}[E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k\rfloor})]]^2 \\ &= \mathbb{E}[F(\mathbb{B}^{k,1})F(\mathbb{B}^{k,2})] - \mathbb{E}[F(\mathbb{B}^{k,1})]\mathbb{E}[F(\mathbb{B}^{k,2})] \\ &\le C(\mathbb{P}((\mathcal{B}^{k,1})^c) + \mathbb{P}((\mathcal{C}^{k,1})^c) + b^{-k/6}) \\ &\quad + \mathbb{E}\left[F\left(\mathbb{B}^{k,1}_{\cdot,\vartheta_k^{X^1}}\right)F\left(\mathbb{B}^{k,2}_{\cdot,\vartheta_k^{X^2}}\right)\right] - \mathbb{E}\left[F\left(\mathbb{B}^{k,1}_{\cdot,\vartheta_k^{X^1}}\right)\right]\mathbb{E}\left[F\left(\mathbb{B}^{k,2}_{\cdot,\vartheta_k^{X^2}}\right)\right].\end{aligned}$$

On the event $\mathcal{A}_k$, we see that $\mathbb{B}^{k,1}_{\cdot,\vartheta_k^{X^1}}$, $\mathbb{B}^{k,2}_{\cdot,\vartheta_k^{X^2}}$ are independent; therefore,

$$\mathbb{E}\left[F\left(\mathbb{B}^{k,1}_{\cdot,\vartheta_k^{X^1}}\right)F\left(\mathbb{B}^{k,2}_{\cdot,\vartheta_k^{X^2}}\right)\,\middle|\,\mathcal{A}_k\right] - \mathbb{E}\left[F\left(\mathbb{B}^{k,1}_{\cdot,\vartheta_k^{X^1}}\right)\,\middle|\,\mathcal{A}_k\right]\mathbb{E}\left[F\left(\mathbb{B}^{k,2}_{\cdot,\vartheta_k^{X^2}}\right)\,\middle|\,\mathcal{A}_k\right] = 0.$$

Using the Lipschitz property of F, we then have

$$\operatorname{var}_{\boldsymbol{P}}(E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k\rfloor})]) \le C(\mathbb{P}((\mathcal{A}^{k,1})^c) + \mathbb{P}((\mathcal{B}^{k,1})^c) + \mathbb{P}((\mathcal{C}^{k,1})^c) + b^{-k/6}).$$

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

For $i = 1, 2$, we see that Y^i are biased random walks on a supercritical GW tree without leaves $\mathcal{T}^g$, whose offspring law has exponential moments. It follows that the estimates $\mathbb{P}((\mathcal{A}^{k,1})^c)$, $\mathbb{P}((\mathcal{B}^{k,1})^c) \le b^{-\tilde{c}k}$ given in the proof of [11, Theorem 3] still hold. Combining these with (3.2), we see that there exists $c > 0$ such that, for sufficiently large k,

$$\operatorname{var}_{\boldsymbol{P}}(E^{\mathcal{T}}[F(\mathbb{B}^{\lfloor b^k \rfloor})]) \le Cb^{-ck},$$

which shows (3.1) and, therefore, the result follows from Lemma 3.1. □

4. Sharpness of the upper bound

In this section we show that when $\beta > f'(q)^{-1/2}$ the walk cannot be rescaled to a standard Gaussian, at least with deterministic centring. We do not study the boundary case $\beta = f'(q)^{-1/2}$ here; this remains an open problem. It is natural to expect that, in this case, $|X_n|$ belongs to the domain of partial attraction of a Gaussian, however, this requires further study. Similarly, for technical reasons, we exclude the $\beta = f'(q)^{-1}$ case. It is reasonable to expect that our proof can be extended to this case, however, more precise knowledge of the first-order behaviour is required for a full proof.

We say that a sequence of random variables $(W_n)_{n\ge 0}$ belongs to the domain of partial attraction of a Gaussian (denoted $W \in D_\partial(2)$) if there exist deterministic increasing divergent sequences n_k and a_k, and a deterministic sequence b_k such that $W_{n_k}/a_k - b_k$ converges in distribution to a standard Gaussian. Our main aim is to prove the following result.

Proposition 4.1. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta \in (f'(q)^{-1/2}, f'(q)^{-1})$. Then $|X_n| \notin D_\partial(2)$.*

We first note that if $\beta > f'(q)^{-1}$ then the result follows from [3] where non-Gaussian limits were established; we therefore restrict ourselves to the $\beta \in (f'(q)^{-1/2}, f'(q)^{-1})$ case.

Define $\gamma := \log(f'(q)^{-1})/\log(\beta)$, which, throughout this section, will be at most 2 by the assumption that $\beta > f'(q)^{-1/2}$. Let $\Delta_0 = 0$ and, for $k \ge 1$, denote by $\Delta_k := \inf\{m \ge 0 \colon X_m \in \mathcal{T}^g, |X_m| = k\}$ the first hitting time of the kth level of the backbone. First, we show that the time between regenerations of the walk has a heavy tail.

Lemma 4.1. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta > f'(q)^{-1/2}$. Then there exists a constant $c > 0$ such that $\mathbb{P}(\Delta_{|\varrho 2|} - \Delta_{|\varrho 1|} > t) \ge 1 \wedge ct^{-\gamma}$.*

Proof. With positive probability, the vertex ϱ_1 has at least one adjacent trap and the walk moves to it immediately. This trap is equal in distribution to $\mathcal{T}^h$, therefore, it suffices to show that $\boldsymbol{E}[P_\rho^{\overline{\mathcal{T}}^h}(\tau_{\overline{\rho}}^+ > t)] \ge 1 \wedge ct^{-\gamma}$ since the transition probabilities inside the traps are independent of the regenerations.

Let $\mathfrak{H}$ denote the height of $\mathcal{T}^h$. Then, by [9, Theorem B], there exists a constant c such that $\boldsymbol{P}(\mathfrak{H} \ge m) \ge cf'(q)^m$. By comparison with a biased random walk on $\mathbb{Z}$, the walk reaches the deepest vertex δ in $\mathcal{T}^h$ with probability at least $p_\infty = 1 - \beta^{-1}$ independently of $\mathfrak{H}$. From δ, the walk has probability $\beta^{-\mathfrak{H}}$ of reaching $\overline{\rho}$ before returning to δ. Moreover, each excursion consists of at least two steps. Therefore, on the event that δ is reached on this first excursion into the trap, the time before reaching $\overline{\rho}$ is dominated by a geometric random variable with

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

termination probability $\beta^{-\mathfrak{H}}$. It follows that, for sufficiently large t,

$$\begin{aligned}\boldsymbol{E}[P_\rho^{\overline{\mathcal{T}}^h}(\tau_{\overline{\rho}}^+ > t)] &\geq p_\infty \boldsymbol{P}\left(\mathfrak{H} \geq \frac{\log(t)}{\log(\beta)}\right) P(\mathrm{geom}(\beta^{-\log(t)/\log(\beta)}) \geq t) \\ &\geq Cf'(q)^{\log(t)/\log(\beta)}(1-\beta^{-\log(t)/\log(\beta)})^t.\end{aligned}$$

This final expression is bounded below by $ct^{-\gamma}$, which completes the proof. □

We now prove a bound on the fluctuations of an i.i.d. sum of random variables with heavy tails.

Lemma 4.2. *Let $S_n \coloneqq \sum_{k=1}^n R_k$ denote the partial sum of a sequence of i.i.d. random variables satisfying $\mathbb{P}(|R_1| > t) \geq 1 \wedge ct^{-\gamma}$ for some positive constants c, γ. Then, for any $T > 0$, there exists $q_T > 0$ independent of n such that $\mathbb{P}(|S_n - c_n| > Tn^{1/\gamma}) \geq q_T$ for any deterministic sequence $(c_n)_{n\geq 1}$.*

Proof. Let $\tilde{S}_n = \sum_{k=1}^n \tilde{R}_k$ be an independent copy of S_n. Then

$$\begin{aligned}\mathbb{P}(|S_n - c_n| > Tn^{1/\gamma}) &= \tfrac{1}{2}(\mathbb{P}(|S_n - c_n| > Tn^{1/\gamma}) + \mathbb{P}(|\tilde{S}_n - c_n| > Tn^{1/\gamma})) \\ &\geq \tfrac{1}{2}\mathbb{P}(\{|S_n - c_n| > Tn^{1/\gamma}\} \cup \{|\tilde{S}_n - c_n| > Tn^{1/\gamma}\}) \\ &\geq \tfrac{1}{2}\mathbb{P}(|S_n - \tilde{S}_n| > 2Tn^{1/\gamma}).\end{aligned}$$

The sum $\hat{S}_n \coloneqq S_n - \tilde{S}_n$ is a sum of i.i.d. symmetric random variables $\hat{R}_k \coloneqq R_k - \tilde{R}_k$. Moreover, $\mathbb{P}(|\hat{R}_1| > t) \geq 1 \wedge Ct^{-\gamma}$ for some constant C.

Fix $s > 0$, for sufficiently large n, we have

$$\mathbb{P}\Big(\max_{k\leq n} |\hat{R}_k| > sn^{1/\gamma}\Big) = 1 - (1 - \mathbb{P}(|\hat{R}_1| > sn^{1/\gamma}))^n \geq 1 - \left(1 - \frac{Cs^{-\gamma}}{n}\right)^n, \tag{4.1}$$

which converges to $1 - \mathrm{e}^{-Cs^{-\gamma}}$ as $n \to \infty$ and is bounded away from 0.

Let $k_n^+ \coloneqq \inf\{k \geq 0 \colon \hat{S}_k > 2Tn^{1/\gamma}\}$, $k_n^- \coloneqq \inf\{k \geq 0 \colon \hat{S}_k < -2Tn^{1/\gamma}\}$, and $k_n^* \coloneqq \inf\{k \geq 0 \colon |\hat{R}_k| > 4Tn^{1/\gamma}\}$. Then note that $\{k_n^* \leq m\} \subseteq \{k_n^+ \leq m\} \cup \{k_n^- \leq m\}$ for any m. In particular, using (4.1) and the symmetry of $\hat{R}_k$ yields $\mathbb{P}(k_n^+ \leq m) = \mathbb{P}(k_n^- \leq m)$, and we see that $\mathbb{P}(k_n^+ \leq n)$ is bounded away from 0.

By the symmetry of $\hat{R}_k$, we have $\mathbb{P}(\hat{S}_m \geq 0) \geq \frac{1}{2}$ for any $m \in \mathbb{Z}^+$. It then follows that, by the independence of $\hat{R}_k$,

$$\mathbb{P}(|\hat{S}_n| > 2Tn^{1/\gamma}) \geq \mathbb{P}\left(\{k_n^+ \leq n\} \cap \left\{\sum_{j=k_n^+ + 1}^{n} \hat{R}_j \geq 0\right\}\right) \geq \mathbb{P}(k_n^+ \leq n) \min_{m\leq n} \mathbb{P}(\hat{S}_m \geq 0),$$

which is bounded away from 0 and completes the proof. □

For an increasing divergent sequence n_k, let $\psi_k \coloneqq \sup\{j \geq 0 \colon |\varrho_j| \leq n_k\}$ be the number of regenerations by level n_k and $\theta = \mathbb{E}[|\varrho_2| - |\varrho_1|]^{-1}$. We now show that ψ_k obeys a CLT.

Lemma 4.3. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta > 1$. Then $(\psi_k - \theta n_k)n_k^{-1/2}$ converges in distribution as $k \to \infty$, with respect to $\mathbb{P}$, to a scaled Gaussian.*

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

Proof. Using the fact that $\{\psi_k \geq m\} = \{|\varrho_m| \leq n_k\}$, we have, for $\lambda \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\left(\frac{\psi_k - \theta n_k}{n_k^{1/2}} \geq \lambda\right) &= \mathbb{P}(\psi_k \geq \lambda n_k^{1/2} + \theta n_k) \\ &= \mathbb{P}\left(\sum_{j=1}^{\lambda n_k^{1/2}+\theta n_k} (|\varrho_j| - |\varrho_{j-1}|) \leq n_k\right) \\ &= \mathbb{P}\Bigg(\sum_{j=2}^{\lambda n_k^{1/2}+\theta n_k} (|\varrho_j| - |\varrho_{j-1}| - \mathbb{E}[|\varrho_j| - |\varrho_{j-1}|]) \\ &\qquad \leq -(\lambda n_k^{1/2} - 1)\mathbb{E}[|\varrho_2| - |\varrho_1|] - |\varrho_1|\Bigg) \\ &= \mathbb{P}\Bigg(\sum_{j=2}^{m_k} \frac{|\varrho_j| - |\varrho_{j-1}| - \mathbb{E}[|\varrho_j| - |\varrho_{j-1}|]}{m_k^{1/2}} \\ &\qquad \leq -\lambda\mathbb{E}[|\varrho_2| - |\varrho_1|]^{3/2} + \Theta_{k,\lambda}(|\varrho_1|)\Bigg), \end{aligned} \tag{4.2}$$

where $m_k = \lambda n_k^{1/2} + \theta n_k$ and $\Theta_{k,\lambda}(|\varrho_1|)$ converges in distribution to 0. The terms $|\varrho_j| - |\varrho_{j-1}| - \mathbb{E}[|\varrho_j| - |\varrho_{j-1}|]$ are i.i.d. and, by comparison with a biased random walk on $\mathbb{Z}$, have exponential moments. By the CLT we then see that (4.2) converges to $\Phi(c\lambda)$ for some constant c, where

$$\Phi(u) := \int_{-\infty}^{u} \frac{\mathrm{e}^{-v^2/2}}{\sqrt{2\pi}}\,\mathrm{d}v.$$

In particular, $(\psi_k - \theta n_k)n_k^{-1/2}$ converges in distribution to a scaled Gaussian. □

Next, we show that Δ_n can be approximated by an i.i.d. sum of heavy-tailed random variables and therefore cannot be rescaled to a standard Gaussian.

Lemma 4.4. *Suppose that $p_0 > 0$, $\mu > 1$, and $\beta \in (f'(q)^{-1/2}, f'(q)^{-1})$. Then $\Delta_n \notin D_\partial(2)$.*

Proof. By contradiction, suppose that $\Delta_{n_k}/a_k - b_k$ converges in distribution to a standard Gaussian for deterministic, increasing, divergent sequences n_k, a_k, and a deterministic sequence b_k. Write

$$\Xi_k := \sum_{j=2}^{\theta n_k} \frac{\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|}}{a_k} - b_k, \qquad \mathcal{E}_k := \frac{\Delta_{n_k} - \Delta_{|\varrho_{\psi_k}|} + \Delta_{|\rho_1|}}{a_k} + \sum_{j=\theta n_k+1}^{\psi_k} \frac{\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|}}{a_k}.$$

Then $\Xi_k + \mathcal{E}_k = \Delta_{n_k}/a_k - b_k$. We first prove that $a_k \geq c n_k^{1/\gamma}$ for some $c > 0$. Using this, we show that $\mathcal{E}_k$ is negligible. It then follows that Ξ_k cannot converge to a Gaussian since the terms $\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|}$ are i.i.d. and, by Lemma 4.1, heavy tailed.

Suppose that there exists an increasing sequence k_l such that $a_{k_l}/n_{k_l}^{1/\gamma} \to 0$. For $T > 0$, write $T_l := T n_{k_l}^{1/\gamma}/a_{k_l}$. Then, by assumption, we have

$$\mathbb{P}(|\Xi_{k_l}| \geq 2T_l) - \mathbb{P}\left(|\mathcal{E}_{k_l}| \geq \frac{T_l}{2}\right) \leq \mathbb{P}(|\Xi_{k_l} + \mathcal{E}_{k_l}| \geq T_l) = \mathbb{P}\left(\left|\frac{\Delta_{n_{k_l}}}{a_{k_l}} - b_{k_l}\right| \geq T_l\right) \to 0. \tag{4.3}$$

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

Since $\gamma > 2$, there exists an increasing sequence m_l satisfying $n_{k_l}^{1/2} \ll m_l \ll n_{k_l}^{1/\gamma}$. From Lemma 4.3 and the fact that $0 \le \Delta_{n_k} - \Delta_{|\varrho_{\psi_k}|} \le \Delta_{|\varrho_{\psi_k+1}|} - \Delta_{|\varrho_{\psi_k}|}$, we have

$$\mathbb{P}\left(|\mathcal{E}_{k_l}| \ge \frac{T_l}{2}\right) \le \mathbb{P}\left(\frac{\Delta_{|\varrho_1|}}{n_{k_l}^{1/\gamma}} + \sum_{j=1}^{m_l} \frac{\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|}}{n_{k_l}^{1/\gamma}} \ge \frac{T}{2}\right) + o(1), \tag{4.4}$$

which converges to 0 by Markov's inequality since $\mathbb{E}[\Delta_{|\varrho_2|} - \Delta_{|\varrho_1|}] < \infty$. By Lemmas 4.1 and 4.2, $\mathbb{P}(|\Xi_{k_l}| \ge 2T_l)$ is bounded below by a positive constant q_{2T}. This contradicts (4.3), therefore, it follows that $a_k \ge c n_k^{1/\gamma}$ for some $c > 0$.

Since $a_k \gg n_k^{1/2}$, similarly to (4.4), we see that $\mathcal{E}_k$ converges in distribution to 0. Finally, since $\{\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|}\}_{j=2}^{\infty}$ are i.i.d. variables with $\mathbb{P}(\Delta_{|\varrho_j|} - \Delta_{|\varrho_{j-1}|} > t) \ge 1 \wedge ct^{-\gamma}$, it follows that Ξ_k cannot converge in distribution to a standard Gaussian (see [8, Theorem IX.8.1]). We can, therefore, conclude that the assumed convergence fails. □

Write $\overline{X}_k := \sup\{|X_m| : m \le n_k, X_m \in \mathcal{T}^g\}$ for the furthest distance reached along the backbone. We conclude with the proof of Proposition 4.1 by showing that the walk never deviates too far from the furthest point reached along the backbone, using the fact that $\{\overline{X}_k \ge x\} = \{\Delta_x \le n_k\}$ for any $x \in \mathbb{Z}$.

Proof of Proposition 4.1. By contradiction, suppose that $X_{n_k}/a_k - b_k$ converges to a standard Gaussian for deterministic, increasing, divergent sequences n_k, a_k, and a deterministic sequence b_k.

The distance $|X_{n_k}| - \overline{X}_k$ is bounded above by the height of the largest trap seen up to time n_k. Up to time n_k, the walk can have visited at most n_k buds, each of which is an independent GW tree $\mathcal{T}^h$. By [9, Theorem B], we have

$$\boldsymbol{P}(\mathcal{H}(\mathcal{T}^h) \ge m) \sim c_h f'(q)^{-m} \quad \text{for some constant } c_h \text{ as } m \to \infty.$$

[-44pt] In particular, using a union bound, it follows that there exists $C > 0$ such that, with high probability, the walk does not visit the bud of a trap of height greater than $C \log(n_k)$ by time n_k.

The distance $\overline{X}_k - |X_{n_k}|$ is bounded above by the furthest distance moved back along the backbone. Since $\beta > 1$, by comparison with a biased random walk on $\mathbb{Z}$, it follows that, with high probability, up to time n_k the walk never backtracks further than $C \log(n_k)$ from the furthest point reached along the backbone. Combining with the above argument, with high probability $|\overline{X}_k - |X_{n_k}|| \le C \log(n_k)$.

To consider $\overline{X}_k$ instead of $|X_{n_k}|$, it now remains to show that $a_k \gg \log(n_k)$. We prove the stronger result that $a_k \ge c n_k^{1/\gamma}$. By contradiction, suppose that there exists an increasing subsequence k_l such that $a_{k_l}/n_{k_l}^{1/\gamma}$ converges to 0. Letting $x \in \mathbb{Z}$, we have

$$\begin{aligned}\mathbb{P}\left(\left|\frac{|X_{n_{k_l}}|}{a_{k_l}} - b_{k_l}\right| \ge x\right) \ge \mathbb{P}\left(\left|\frac{\overline{X}_{k_l}}{a_{k_l}} - b_{k_l}\right| \ge x + \frac{C\log(n_{k_l})}{a_{k_l}}\right) \\ - \mathbb{P}(|\overline{X}_{k_l} - |X_{n_{k_l}}|| \ge C\log(n_{k_l})),\end{aligned} \tag{4.5}$$

where $\mathbb{P}(|\overline{X}_{k_l} - X_{n_{k_l}}| \ge C\log(n_{k_l}))$ converges to 0 as $l \to \infty$. Using the fact that $\{\overline{X}_k \ge m\} = \{\Delta_m \le n_k\}$ and denoting $C_{x,l} := x a_{k_l} + C\log(n_{k_l})$, it follows that

$$\mathbb{P}\left(\left|\frac{\overline{X}_{k_l}}{a_{k_l}} - b_{k_l}\right| \ge x + \frac{C\log(n_{k_l})}{a_{k_l}}\right) \ge \mathbb{P}(\overline{X}_{k_l} \ge a_{k_l} b_{k_l} + C_{x,l}) + \mathbb{P}(\overline{X}_{k_l} \le a_{k_l} b_{k_l} - C_{x,l})$$

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

$$= \mathbb{P}(\Delta_{a_{k_l}b_{k_l}+C_{x,l}} - \Delta_{a_{k_l}b_{k_l}} \le n_{k_l} - \Delta_{a_{k_l}b_{k_l}}) + \mathbb{P}(\Delta_{a_{k_l}b_{k_l}} - \Delta_{a_{k_l}b_{k_l}-C_{x,l}} \le \Delta_{a_{k_l}b_{k_l}} - n_{k_l}).$$

From [10], it follows that $(\Delta_{a_{k_l}b_{k_l}+C_{x,l}} - \Delta_{a_{k_l}b_{k_l}})/C_{x,l}$ and $(\Delta_{a_{k_l}b_{k_l}} - \Delta_{a_{k_l}b_{k_l}-C_{x,l}})/C_{x,l}$ converge in distribution to ν^{-1}, therefore, the above expression can be bounded below by

$$(1 - o(1))\mathbb{P}(|\Delta_{a_{k_l}b_{k_l}} - n_{k_l}| \ge 2C_{x,l}\nu^{-1}),$$

which, by Lemmas 4.2 and 4.4, is bounded below by a constant $q > 0$ (independent of x) for sufficiently large l. In particular, this contradicts (4.5) since, as $l \to \infty$,

$$\mathbb{P}\left(\left|\frac{|X_{n_{k_l}}|}{a_{k_l}} - b_{k_l}\right| \ge x\right) \to 2\Phi(-x) = 2\int_x^\infty \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}\,\mathrm{d}x,$$

which can be made arbitrarily small by choosing x large. Therefore, we have $\liminf_k a_k/n_k^{1/\gamma} > 0$, which is sufficient to consider $\overline{X}_k$ instead of X_{n_k}.

We now prove that a CLT does not exist for $\overline{X}_k$. Using the fact that $\{\overline{X}_k \ge m\} = \{\Delta_m \le n_k\}$, we have

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\overline{X}_k}{a_k} - b_k\right| \ge x\right) &= \mathbb{P}(\Delta_{(b_k+x)a_k} \le n_k) + \mathbb{P}(\Delta_{(b_k-x)a_k} \ge n_k) \\ &= \mathbb{P}\left(\frac{\Delta_{(b_k+x)a_k} - \Delta_{b_k a_k}}{a_k} \le \frac{n_k - \Delta_{b_k a_k}}{a_k}\right) + \mathbb{P}\left(\frac{\Delta_{b_k a_k} - \Delta_{(b_k-x)a_k}}{a_k} \le \frac{\Delta_{b_k a_k} - n_k}{a_k}\right).\end{aligned}$$

From [10], we see that $(\Delta_{(b_k+x)a_k} - \Delta_{b_k a_k})/a_k$ and $(\Delta_{b_k a_k} - \Delta_{(b_k-x)a_k})/a_k$ converge in distribution to $\nu^{-1}x$, therefore, the result follows from Lemma 4.4 and the fact that the Gaussian distribution is continuous. □

Acknowledgements

I would like to thank my supervisor David Croydon for suggesting the problem, his support and many useful discussions. This work is supported by EPSRC as part of the MASDOC DTC at the University of Warwick (grant number EP/H023364/1).

References

[1] Athreya, K. B. and Ney, P. E. (2004). *Branching Processes*. Dover, Mineola, NY.

[2] Ben Arous, G. and Fribergh, A. (2016). Biased random walks on random graphs. In *Probability and Statistical Physics in St. Petersburg* (Proc. Sympos. Pure Math. **91**), American Mathematical Society, Providence, RI, pp. 99–153.

[3] Ben Arous, G., Fribergh, A., Gantert, N. and Hammond, A. (2012). Biased random walks on Galton–Watson trees with leaves. *Ann. Prob.* **40,** 280–338.

[4] Bolthausen, E. and Sznitman, A.-S. (2002). On the static and dynamic points of view for certain random walks in random environment. *Meth. Appl. Anal.* **9,** 345–375.

[5] Bowditch, A. (2018). Central limit theorems for biased randomly trapped walks on $\mathbb{Z}$. *Stoch. Process Appl.* Available at https://doi.org/10.1016/j.spa.2018.03.017.

[6] Bowditch, A. (2018). Escape regimes of biased random walks on Galton–Watson trees. *Prob. Theory Relat. Fields* **170,** 685–768.

[7] Dembo, A. and Sun, N. (2012). Central limit theorem for biased random walk on multi-type Galton-Watson trees. *Electron. J. Prob.* **17,** 75.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press

[8] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
[9] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Prob.* **23,** 1125–1138.
[10] Lyons, R., Pemantle, R. and Peres, Y. (1996). Biased random walks on Galton–Watson trees. *Prob. Theory Relat. Fields* **106,** 249–264.
[11] Peres, Y. and Zeitouni, O. (2008). A central limit theorem for biased random walks on Galton–Watson trees. *Prob. Theory Relat. Fields* **140,** 595–629.
[12] Sznitman, A.-S. (2000). Slowdown estimates and central limit theorem for random walks in random environment. *J. Europ. Math. Soc. (JEMS)* **2,** 93–143.
[13] Whitt, W. (2002). *Stochastic-Process Limits*. Springer, New York.

https://doi.org/10.1017/jpr.2018.38 Published online by Cambridge University Press