RIMS-1744

Biased random walk on critical Galton-Watson trees conditioned to survive

By

David A. CROYDON, Alexander FRIBERGH
and Takashi KUMAGAI

March 2012

京都大学　数理解析研究所

RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES
KYOTO UNIVERSITY, Kyoto, Japan

Biased random walk on critical Galton-Watson trees conditioned to survive

D. A. Croydon*, A. Fribergh† and T. Kumagai‡

March 17, 2012

Abstract

We consider the biased random walk on a critical Galton-Watson tree conditioned to survive, and confirm that this model with trapping belongs to the same universality class as certain one-dimensional trapping models with slowly-varying tails. Indeed, in each of these two settings, we establish closely-related functional limit theorems involving an extremal process and also demonstrate extremal aging occurs.

1 Introduction

Biased random walks in inhomogeneous environments are a natural setting to witness trapping phenomena. In the case of supercritical Galton-Watson trees with leaves (see [6], [8], [27]) or the supercritical percolation cluster on $\mathbb{Z}^d$ (see [17]), for example, it has been observed that dead-ends found in the environment can, for suitably strong biases, create a sub-ballistic regime that is characteristic of trapping. More specifically, for both of these models, the distribution of the time spent in individual traps has polynomial tail decay, and this places them in the same universality class as the one-dimensional heavy-tailed trapping models considered in [32]. Indeed, although the case of a deterministically biased random walk on a Galton-Watson tree with leaves is slightly complicated by a certain lattice effect, which means it can not be rescaled properly [6], in the case of randomly biased random walks on such structures, it was shown in [8] that precisely the same limiting behaviour as the one-dimensional models of [15] and [32] occurs. Moreover, there is evidence presented in [17] that suggests the biased random walk on a supercritical percolation cluster also has the same limiting behaviour. The universality class that connects these models was previously investigated in [3], [4] and [5], and is characterised by limiting stable subordinators and aging properties.

The aim of this paper is to investigate biased random walks on critical structures. To this end, we choose to study the biased random walk on a critical Galton-Watson tree conditioned to survive. With the underlying environment having radically different properties from its supercritical counterpart, we would expect different limiting behaviour, with more extreme trapping phenomena, to arise. It is further natural to believe that some of the properties of the biased random walk on the incipient infinite cluster for critical percolation on $\mathbb{Z}^d$, at least in high dimensions, would be similar to the ones proved in our context, as is observed to be the case for the unbiased random walk (compare, for instance, the results of [1] and [24]). Nevertheless,

*Dept of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom; d.a.croydon@warwick.ac.uk.
†CIMS, 251 Mercer Street, New York University, New York 10012-1185, U.S.A.; fribergh@cims.nyu.edu.
‡RIMS, Kyoto University, Kyoto 606-8502, Japan; kumagai@kurims.kyoto-u.ac.jp.

our current understanding of the geometry of this object is not sufficient to extend our results easily, and so we do not pursue this inquiry here. In particular, we anticipate that, as indicated by physicists in [2], for percolation close to criticality there is likely to be an additional trapping mechanism that occurs due to spatial considerations, which means that, even without taking the effect of dead-ends into account, it is more likely for the biased random walk to be found in certain regions of individual paths than others (see [9] for a preliminary study in this direction).

Our main model – the biased random walk on critical Galton-Watson trees conditioned to survive – is presented in the next section, along with a summary of the results we are able to prove for it. This is followed in Section 1.2 with an introduction to a one-dimensional trapping model in which the trapping time distributions have slowly-varying tails. This latter model, which is of interest in its own right, is of particular relevance for us, as it allows us to comprehensively characterise the universality class into which the Galton-Watson trees we consider fall. Furthermore, the arguments we apply for the one-dimensional model provide a useful template for the more complicated tree framework.

1.1 Biased random walk on critical Galton-Watson trees

Before presenting the Galton-Watson tree framework, we recall some classical results for sums of random variables whose distribution has a slowly-varying tail. Let $(X_i)_{i=1}^{\infty}$ be independent random variables, with distributional tail $\bar{F}(u) = 1 - F(u) = \mathbf{P}(X_i > u)$ satisfying: $\bar{F}(0) = 1$, $\bar{F}(u) > 0$ for all $u > 0$,

$$\lim_{u\to\infty} \frac{\bar{F}(uv)}{\bar{F}(u)} = 1, \tag{1.1}$$

for any $v > 0$, and $\bar{F}(u) \to 0$ as $u \to \infty$. A typical example is when the distribution in question decays logarithmically slowly, such as

$$\bar{F}(u) \sim \frac{1}{(\ln u)^{\gamma}}, \tag{1.2}$$

for some $\gamma > 0$, where throughout the article $f \sim g$ will mean $f(x)/g(x) \to 1$ as $x \to \infty$. A first scaling result for sums of the form $\sum_{i=1}^{n} X_i$ was obtained in [11], and this was subsequently extended by [22] to a functional result. In particular, in [22] it was established that if $L(x) := 1/\bar{F}(x)$, then

$$\left(\frac{1}{n} L\left(\sum_{i=1}^{nt} X_i\right)\right)_{t\geq 0} \to (m(t))_{t\geq 0} \tag{1.3}$$

in distribution with respect to the Skorohod J_1 topology (as an aid to the reader, we provide in the appendix a definition of the Skorohod J_1 and M_1 topologies, the latter of which is applied in several subsequent results), where $m = (m(t))_{t\geq 0}$ is an extremal process. To define m more precisely, suppose that $(\xi(t))_{t\geq 0}$ is the symmetric Cauchy process, i.e., the Lévy process with Lévy measure given by $\mu((x,\infty)) = x^{-1/2}$ for $x > 0$, and then set

$$m(t) = \max_{0<s\leq t} \Delta\xi(s),$$

where $\Delta\xi(s) = \xi(s) - \xi(s^-)$. (Observe that $(m(t))_{t\geq 0}$ is thus the maximum process of the Poisson point process with intensity measure $x^{-2}dxdt$.) We will prove that, in addition to appearing in the limit at (1.3), this extremal process arises in the scaling limits of a biased random walk on a critical Galton-Watson tree and, as is described in the next section, a one-dimensional directed trap model whose holding times have a slowly-varying mean.

We continue by introducing some relevant branching process and random walk notation, following the presentation of [10]. Let Z be a critical ($\mathbf{E}Z = 1$) offspring distribution in the domain of attraction of a stable law with index $\alpha \in (1, 2]$, by which we mean that there exists a sequence $a_n \uparrow \infty$ such that

$$\frac{Z[n] - n}{a_n} \xrightarrow{d} X, \tag{1.4}$$

where $Z[n]$ is the sum of n i.i.d. copies of Z and $\mathbf{E}(e^{-\lambda X}) = e^{-\lambda^\alpha}$ for $\lambda \geq 0$. Note that, by results of [16, Chapters XIII and XVII], this is equivalent to the probability generating function of Z satisfying

$$f(s) := \mathbf{E}(s^Z) = \sum_{k=0}^{\infty} p_k s^k = s + (1-s)^\alpha L(1-s), \qquad \forall s \in (0, 1), \tag{1.5}$$

where $L(x)$ is slowly varying as $x \to 0^+$, and the non-triviality condition $\mathbf{P}(Z = 1) \neq 1$ holding. We point out that the condition $\mathbf{E}(Z^2) < \infty$ is sufficient for the previous statements to hold with $\alpha = 2$.

Denote by $(Z_n)_{n\geq 0}$ the corresponding Galton-Watson process, started from $Z_0 = 1$. It has been established in [29, Lemma 2] that if $q_n := \mathbf{P}(Z_n > 0)$, then

$$q_n^{\alpha-1} L(q_n) \sim \frac{1}{(\alpha - 1)n}, \tag{1.6}$$

as $n \to \infty$, where L is the function appearing in (1.5). It is also well known that the branching process $(Z_n)_{n\geq 0}$ can be obtained as the generation size process of a Galton-Watson tree, $\mathcal{T}$ say, with offspring distribution Z. In particular, to construct the random rooted graph tree $\mathcal{T}$, start with a single ancestor (or root), and then suppose that individuals in a given generation have offspring independently of the past and each other according to the distribution of Z, see [25, Section 3] for details. The vertex set of $\mathcal{T}$ is the entire collection of individuals, edges are the parent-offspring bonds, and Z_n is the number of individuals in the nth generation of $\mathcal{T}$. From (1.6), it is clear that $\mathcal{T}$ will be a finite graph $\mathbf{P}$-a.s. However, in [23], Kesten showed that it is possible to make sense of conditioning $\mathcal{T}$ to survive or 'grow to infinity'. More specifically, there exists a unique (in law) random infinite rooted locally-finite graph tree $\mathcal{T}^*$ that satisfies, for any $n \in \mathbb{Z}_+$,

$$\mathbf{E}\left(\phi(\mathcal{T}^*|_n)\right) = \lim_{m\to\infty} \mathbf{E}\left(\phi(\mathcal{T}|_n) | Z_{m+n} > 0\right),$$

where ϕ is a bounded function on finite rooted graph trees of n generations, and $\mathcal{T}|_n$, $\mathcal{T}^*|_n$ are the first n generations of $\mathcal{T}$, $\mathcal{T}^*$ respectively. We will write $d_{\mathcal{T}^*}$ to represent the shortest path graph distance on $\mathcal{T}^*$.

Given a particular realisation of $\mathcal{T}^*$, we will denote by $X = ((X_n)_{n\geq 0}, P_x^{\mathcal{T}^*}, x \in \mathcal{T}^*)$ the discrete-time biased random walk on $\mathcal{T}^*$, and define this as follows. First, fix a bias parameter $\beta > 1$, and assign to each edge connecting a vertex x in generation k to a vertex y in generation $k + 1$ a conductance $c(x, y) := \beta^k =: c(y, x)$. The transition probabilities of X are then determined by

$$P^{\mathcal{T}^*}(x, y) := \frac{c(x, y)}{\sum_{y'\sim x} c(x, y')}, \qquad \forall x \sim y,$$

where the notation $x \sim y$ means that x and y are connected by an edge in $\mathcal{T}^*$. Thus, when at a vertex x that is not equal to the root of $\mathcal{T}^*$, the probability of jumping to a neighbouring vertex further away from the root than x is β times more likely than jumping towards the root. Using the usual terminology for random walks in random environments, we will say that $P_x^{\mathcal{T}^*}$

is the quenched law of the biased random walk on $\mathcal{T}^*$ started from x. Moreover, we introduce the annealed law for the process started from ρ, the root of the tree $\mathcal{T}^*$, by setting

$$\mathbb{P}_\rho(\cdot) := \int P_\rho^{\mathcal{T}^*}(\cdot)\mathrm{d}\mathbf{P}. \tag{1.7}$$

It will be this law under which we investigate the rate at which the process X, which we call the biased random walk on a critical Galton-Watson tree conditioned to survive, escapes from the root.

The main result we prove for the process X concerns the time it takes to progress along the backbone. To be more specific, as is described in more detail in Section 3.1, $\mathbf{P}$-a.s. the tree $\mathcal{T}^*$ admits a unique backbone, that is, a semi-infinite path starting from the root, $\{\rho = \rho_0, \rho_1, \rho_2, \dots\}$ say. We define $(\Delta_n)_{n\geq 0}$ by setting

$$\Delta_n := \inf\{m \geq 0 :\ X_m = \rho_n\} \tag{1.8}$$

to be the first time the process X reaches level n along this path. For this process, we are able to prove the following functional limit theorem.

Theorem 1.1. *Let $\alpha \in (1, 2]$. As $n \to \infty$, the laws of the processes*

$$\left(\frac{(\alpha-1)\ln_+ \Delta_{nt}}{n \ln \beta}\right)_{t\geq 0}$$

under $\mathbb{P}_\rho$ converge weakly with respect to the Skorohod J_1 topology on $D([0,\infty),\mathbb{R})$ to the law of $(m(t))_{t\geq 0}$.

It is interesting to observe that this result is extremely explicit compared to its supercritical counterparts. Indeed, notwithstanding the fact the lattice-effect that was the source of somewhat complicated behaviour in [6] does not occur in the critical setting, the above scaling limit clearly describes the β-dependence of the relevant slowdown effect. Note that, unlike in the supercritical case where there is a ballistic phase, this slowdown effect occurs for any non-trivial bias parameter, i.e. for any $\beta > 1$. Furthermore, we remark that the dependence on α is natural: as α decreases and the leaves get thicker (in the sense that tree's Hausdorff dimension of $\alpha/(\alpha-1)$ increases, see [12, 21]), the biased random walk moves more slowly away from its start point.

As suggested by comparing Theorem 1.1 with (1.3), the critical Galton-Watson tree case is closely linked with a sum of independent and identically-distributed random variables where $\bar{F}$ is asymptotically equivalent to $\ln\beta/(\alpha-1)\ln x$. Although the logarithmic rate of decay is relatively easy to guess, finding the correct constant is slightly subtle, particularly for $\alpha \neq 2$. This is because, unlike in the supercritical case and the critical case with $\alpha = 2$, when $\alpha \neq 2$ it can happen that there are multiple deep traps emanating from a single backbone vertex. As a result, we have to take special care which of these have actually been visited when determining the time spent there, meaning that the random variable which actually has the $\ln\beta/(\alpha-1)\ln x$ tail behaviour is not environment measurable (see Lemma 3.11). To highlight the importance of this consideration, which is also relevant albeit in a simpler way for $\alpha = 2$, in Theorem 3.14 we show that the constant that appears differs by a factor α when Δ_n is replaced by its quenched mean $E_\rho^{\mathcal{T}^*}\Delta_n$.

Theorem 1.1 readily implies the following corollary for the projection, $(\pi(X_m))_{m\geq 0}$, of the process $(X_m)_{m\geq 0}$ onto the backbone (roughly, $\pi(X_m)$ is the vertex on the backbone from which the trap X_m is located in emanates, see Section 3.2 for a precise definition). To state this, we define the right-continuous inverse $(m^{-1}(t))_{t\geq 0}$ of $(m(t))_{t\geq 0}$ by setting

$$m^{-1}(t) := \inf\{s \geq 0 :\ m(s) > t\}. \tag{1.9}$$

Corollary 1.2. *Let $\alpha \in (1,2]$. As $n \to \infty$, the laws of the processes*

$$\left(\frac{d_{\mathcal{T}^*}(\rho, \pi(X_{e^{nt}})) \ln \beta}{(\alpha - 1)n} \right)_{t \geq 0}$$

under $\mathbb{P}_\rho$ converge weakly with respect to the Skorohod M_1 topology on $D([0,\infty), \mathbb{R})$ to the law of $(m^{-1}(t))_{t \geq 0}$.

Remark 1.3. *Since the height of the leaves in which the random walk can be found at time e^n (see the localisation result of Lemma 4.5) will typically be of order n, some further argument will be necessary to deduce a limit result for the graph distance $d_{\mathcal{T}^*}(\rho, X_n)$ itself.*

Another characteristic property that we are able to show is that the random walk also exhibits extremal aging.

Theorem 1.4. *Let $\alpha \in (1,2]$. For any $0 < a < b$, we have*

$$\lim_{n \to \infty} \mathbb{P}_\rho \left(\pi(X_{e^{an}}) = \pi(X_{e^{bn}}) \right) = \frac{a}{b}.$$

Although regular aging has previously been observed for random walks in random environments in the sub-ballistic regime on $\mathbb{Z}$ (see [14]), as far as we know, this is the first example of a random walk in random environment where extremal aging has been proved. As already hinted at, this kind of behaviour, as well as that demonstrated in Theorem 1.1 and Corollary 1.2, places the biased random walk on a critical Galton-Watson tree conditioned to survive in a different universality class to that of the supercritical structures discussed previously. In the class of critical Galton-Watson trees we have instead the spin glass models considered in [7] and [20], and the trap models with slowly-varying tails we introduce in the next section.

1.2 One-dimensional directed trap model with slowly-varying tails

In this section, we describe the one-dimensional trap model with which we want to compare to our main model, and the results we are able to prove for it. To start with a formal definition, let $\tau = (\tau_x)_{x \in \mathbb{Z}}$ be a family of independent and identically-distributed strictly positive (and finite) random variables whose distribution has a slowly-varying tail, in the sense described by (1.1), built on a probability space with measure $\mathbf{P}$; the sequence $\tau = (\tau_x)_{x \in \mathbb{Z}}$ will represent the trap environment. For a fixed bias parameter $\beta > 1$, the directed trap model is then the continuous-time Markov process $X = (X_t)_{t \geq 0}$ with state space $\mathbb{Z}$, given by $X_0 = 0$ and with jump rates

$$c(x,y) := \begin{cases} \left(\frac{\beta}{\beta+1} \right) \tau_x^{-1}, & \text{if } y = x+1, \\ \left(\frac{1}{\beta+1} \right) \tau_x^{-1}, & \text{if } y = x-1, \end{cases}$$

and $c(x,y) = 0$ otherwise. To be more explicit, for a particular realisation of τ we will write P_x^τ for the law of the Markov chain with the above transition rates, started from x; similarly to describing $P_x^{\mathcal{T}^*}$ in the previous section, we call this the quenched law for the directed trap model. The corresponding annealed law $\mathbb{P}_x$ is obtained by integrating out the environment similarly to (1.7), i.e.

$$\mathbb{P}_x(\cdot) := \int P_x^\tau(\cdot) \mathrm{d}\mathbf{P}.$$

In studying the rate of escape of the above directed trap model, it is our initial aim to determine the rate of growth of

$$\Delta_n := \inf\{t \geq 0 : X_t = n\},$$

that is, the hitting times of level n by the process X. The following theorem contains our main conclusion in this direction. As in the statement at (1.3), we define $L(x) = 1/\bar{F}(x)$.

Theorem 1.5. *As $n \to \infty$, the laws of the processes*

$$\left(\frac{1}{n} L\left(\Delta_{nt}\right)\right)_{t \geq 0}$$

under $\mathbb{P}_0$ converge weakly with respect to the Skorohod J_1 topology on $D([0,\infty), \mathbb{R})$ to the law of the extremal process $(m(t))_{t \geq 0}$.

Similarly to [22, Remark 2.4], we note that the proof of the above result may be significantly simplified in the case when $\bar{F}$ decays logarithmically. The reason for this is that, in the logarithmic case, the hitting time Δ_n is very well-approximated by the maximum holding time within the first n vertices, and so the functional scaling limit for $(\Delta_n)_{n \geq 0}$ can be readily obtained from a simple study of the maximum holding time process. For general slowly varying functions, the same approximation does not provide tight enough control on Δ_n to apply this argument, and so a more sophisticated approach is required.

As a simple corollary of Theorem 1.5, it is also possible to obtain a scaling result for the process X itself. The definition of m^{-1} should be recalled from (1.9). We similarly define the right-continuous inverse $\bar{F}^{-1}$ of $\bar{F}$, only with $>$ replaced by $<$.

Corollary 1.6. *As $n \to \infty$, the laws of the processes*

$$\left(\frac{1}{n} X_{\bar{F}^{-1}(1/nt)}\right)_{t \geq 0}$$

under $\mathbb{P}_0$ converge weakly with respect to the Skorohod M_1 topology on $D([0,\infty), \mathbb{R})$ to the law of $(m^{-1}(t))_{t \geq 0}$.

Remark 1.7. *(i) Although the preceding corollary does look somewhat awkward, it becomes much clearer for concrete choices of $\bar{F}$. For example, if $\bar{F}$ has the form described at (1.2), then the above result concerns the distributional limit of*

$$\left(\frac{1}{n} X_{e^{(nt)^{1/\gamma}}}\right)_{t \geq 0}.$$

Moreover, it can be deduced from the above result that, as $t \to \infty$, the random variable $\bar{F}(t) X_t$ converges in distribution under $\mathbb{P}_0$ to $m^{-1}(1)$, which is easily checked to have a mean one exponential distribution.
(ii) In a number of places in the proofs of Theorem 1.5 and Corollary 1.6, we are slightly cavalier about assuming that $\bar{F}(\bar{F}^{-1}(x)) = x$ for $x \in (0,1)$. This is, of course, only true in general when $\bar{F}$ is continuous. In the case when this condition is not satisfied, however, we can easily overcome the difficulties that arise by replacing $\bar{F}$ with any non-increasing continuous function $\bar{G}$ that satisfies $\bar{G}(0) = 1$ and $\bar{G}(u) \sim \bar{F}(u)$ as $u \to \infty$. For example, one could define such a $\bar{G}$ by setting $\bar{G}(u) := (\frac{1}{u} \int_0^u L(v) dv)^{-1}$.

The extremal aging result we are able to prove in this setting is as follows.

Theorem 1.8. *For any $0 < a < b$, we have*

$$\lim_{n \to \infty} \mathbb{P}_0 \left(X_{\bar{F}^{-1}(1/na)} = X_{\bar{F}^{-1}(1/nb)} \right) = \frac{a}{b}.$$

Remark 1.9. *Note that if $\bar{F}_n$ and $\bar{F}$ are not continuous and eventually strictly decreasing, a minor modification to the proof of the above result (cf. Remark 1.7(ii)) is needed.*

1.3 Article outline and notes

The remainder of the article is organised as follows. In Section 2, we study the one-dimensional trap model introduced in Section 1.2 above, proving Theorem 1.5 and Corollary 1.6. In Section 3, we then adapt the relevant techniques to derive Theorem 1.1 and Corollary 1.2 for the Galton-Watson tree model. The arguments of both these sections depend on the extension of the limit at (1.3) that is proved in Section 5. Before this, in Section 4, we derive the extremal aging results of Theorems 1.4 and 1.8. Finally, as noted earlier, the appendix recalls some basic facts concerning Skorohod space.

We finish the introduction with some notes about the conventions used in this article. Firstly, there are two widely used versions of the geometric distribution with a given parameter, one with support $0, 1, 2, \ldots$ and one with support $1, 2, 3, \ldots$. In the course of this work, we will use both, and hope that, even without explanation, it is clear from the context which version applies when. Secondly, there are many instances when for brevity we use a continuous variable where a discrete argument is required, in such places x, say, should be read as $\lfloor x \rfloor$. Finally, we recall that $f \sim g$ will mean $f(x)/g(x) \to 1$ as $x \to \infty$.

2 Directed trap model with slowly-varying tails

This section is devoted to the proof of Theorem 1.5 and Corollary 1.6. To this end, we start by deriving some slight adaptations of results from [32] regarding the trap environment. First, define a level n critical depth for traps of the environment by setting

$$g(n) := \bar{F}^{-1}(n^{-1} \ln n). \tag{2.1}$$

We will say that there are deep traps at the sites $\mathcal{D} := \{x \in \mathbb{Z} : \tau_x > g(n)\}$, and consider the following events: for $n \in \mathbb{N}$, $T \in (0, \infty)$,

$$\mathcal{E}_1(n, T) := \left\{ \min_{\substack{x_1, x_2 \in \mathcal{D} \cap [1, nT]: \\ x_1 \neq x_2}} |x_1 - x_2| > n^{\kappa} \right\},$$

$$\mathcal{E}_2(n) := \left\{ \mathcal{D} \cap [-(\ln n)^{1+\gamma}, 0] = \emptyset \right\},$$

where $\kappa, \gamma \in (0, 1)$ are fixed. The event $\mathcal{E}_1(n, T)$ requires that the distance between any two deep traps in the interval $[1, nT]$ is large, and the event $\mathcal{E}_2(n)$ will help to ensure that the time the process X spends outside of the strictly positive integers is negligible.

Lemma 2.1. *Fix $T \in (0, \infty)$. As $n \to \infty$, the $\mathbf{P}$-probability of the events $\mathcal{E}_1(n, T)$ and $\mathcal{E}_2(n)$ converge to one.*

Proof. To check the result for $\mathcal{E}_1(n, T)$, we simply observe that

$$\mathbf{P}\left(\mathcal{E}_1(n, T)^c\right) \leq \sum_{\substack{\{x_1, x_2\} \subseteq [1, nT]: \\ 0 < |x_1 - x_2| \leq n^{\kappa}}} \mathbf{P}\left(\tau_{x_1}, \tau_{x_2} > g(n)\right) \leq T n^{1+\kappa} \bar{F}\left(g(n)\right)^2 \leq \frac{T (\ln n)^2}{n^{1-\kappa}} \to 0.$$

Similarly, we have that $\mathbf{P}(\mathcal{E}_2(n)^c) \leq n^{-1}(1 + (\ln n)^{1+\gamma}) \ln n$, which also converges to 0. □

We continue by introducing the embedded discrete-time random walk associated with X and some of its properties, which will be useful throughout the remainder of the section. In particular, first let $S(0) = 0$ and $S(n)$ be the time of the nth jump of X; this is the clock process

corresponding to X. The embedded discrete-time random walk is then the process $Y = (Y_n)_{n\geq 0}$ defined by setting $Y_n := X_{S(n)}$. Clearly Y is a biased random walk on $\mathbb{Z}$ under P_0^τ for $\mathbf{P}$-a.e. realisation of τ, and thus satisfies, P_0^τ-a.s.,

$$\frac{Y_n}{n} \to \frac{\beta - 1}{\beta + 1} > 0.$$

Whilst this result already tells us that the embedded random walk Y drifts off to $+\infty$ and that the time it takes to hit level n, that is,

$$\Delta_n^Y := \inf\{k \geq 0 : Y_k = n\},$$

is finite for each n, P_0^τ-a.s., we further require that it does not backtrack too much, in the sense that, for each $T \in (0, \infty)$,

$$\mathcal{E}_3(n, T) := \left\{ \min_{0\leq i<j\leq \Delta_{nT}^Y} (Y_j - Y_i) > -(\ln n)^{1+\gamma} \right\}$$

occurs with high probability. This is the content of the following lemma, which is essentially contained in [32, Lemma 3].

Lemma 2.2. *Fix $T \in (0, \infty)$. As $n \to \infty$, the $\mathbb{P}_0$-probability of the event $\mathcal{E}_3(n, T)$ converges to one.*

Let us now introduce the total time the biased random walk X spends at a site $x \in \mathbb{Z}$,

$$T_x := \int_0^\infty \mathbf{1}_{\{X_t = x\}} dt.$$

To study this, first observe that the clock process $S = S(n)_{n\geq 0}$ can be written

$$S(n) = \sum_{i=0}^{n-1} \tau_{Y_i} \mathbf{e}_i,$$

where $(\mathbf{e}_i)_{i\geq 0}$ is an independent sequence of mean one exponential random variables under P_0^τ, independent of Y. Moreover, for $x \in \mathbb{Z}$, let $G(x) = \#\{n \geq 0 : Y_n = x\}$ be the total number of visits of the embedded random walk Y to x. By applying the fact that Y is a random walk with a strictly positive bias, we have that if $x \geq 0$, then $G(x)$ has the geometric distribution with parameter $p = (\beta - 1)/(\beta + 1)$ (again for $\mathbf{P}$-a.e. realisation of τ). It follows that T_x is equal in distribution under $\mathbb{P}_0$ to the random variable

$$\tau_x \sum_{i=1}^{G(x)} \mathbf{e}_i, \tag{2.2}$$

which is almost-surely finite. We will use this characterisation of the distribution of T_x to check that the time spent by X in traps that are not deep is asymptotically negligible, in the sense described by the following event: for $n \in \mathbb{N}$, $T \in (0, \infty)$,

$$\mathcal{E}_4(n, T) := \left\{ \sum_{i=0}^{\Delta_{nT}^Y - 1} \tau_{Y_i} \mathbf{e}_i \mathbf{1}_{\{\tau_{Y_i} \leq g(n)\}} < \bar{F}^{-1}(n^{-1}(\ln n)^{1/2}) \right\},$$

In particular, by similar arguments to [32, Lemma 4], we deduce the following.

Lemma 2.3. *Fix $T \in (0,\infty)$. As $n \to \infty$, the $\mathbb{P}_0$-probability of the event $\mathcal{E}_4(n,T)$ converges to one.*

Proof. We start by checking that

$$\mathbf{E}(\tau_0 \mathbf{1}_{\{\tau_0 \le g(n)\}}) = o(n^{-1}\bar{F}^{-1}(n^{-1}(\ln n)^{1/2})). \tag{2.3}$$

To this end, let $\rho, \varepsilon \in (0,1)$, and observe that

$$\begin{aligned}
\mathbf{E}(\tau_0 \mathbf{1}_{\{\tau_0 \le g(n)\}}) &\le g(n) \sum_{j=0}^{\infty} \rho^j \mathbf{P}(\tau_0 > \rho^{j+1} g(n)) \\
&\le c_1 g(n) \sum_{j=0}^{\infty} \rho^{j-\varepsilon(j+1)} \bar{F}(g(n)) \\
&\le \frac{c_2 g(n) \ln n}{n},
\end{aligned}$$

where the second inequality is an application of the representation theorem for slowly varying functions ([28, Theorem 1.2], for example), which implies that, for any $\varepsilon > 0$, there exists a constant $c_3 \in (0,\infty)$ such that

$$\frac{\bar{F}(v)}{\bar{F}(u)} \le c_3 \left(\frac{u}{v}\right)^{\varepsilon}, \tag{2.4}$$

for all $0 < v \le u$. Again applying (2.4), we have that $g(n) \le c_4 \bar{F}^{-1}(n^{-1}(\ln n)^{1/2})(\ln n)^{-1/2\varepsilon}$. Hence, if ε is chosen small enough, then (2.3) holds as desired.

To proceed, note that, on $\mathcal{E}_3(n,T)$, we have that

$$\sum_{i=0}^{\Delta^Y_{nT}-1} \tau_{Y_i} \mathbf{e}_i \mathbf{1}_{\{\tau_{Y_i} \le g(n)\}} \le \sum_{x=-(\ln n)^{1+\gamma}}^{nT-1} T_x \mathbf{1}_{\{\tau_x \le g(n)\}}.$$

Consequently, because $E^\tau_0 T_x = \tau_x E^\tau_0 G(x) \le \frac{\beta+1}{\beta-1}\tau_x$, it follows that

$$E^\tau_0 \left(\sum_{i=0}^{\Delta^Y_{nT}-1} \tau_{Y_i} \mathbf{e}_i \mathbf{1}_{\{\tau_{Y_i} \le g(n)\}} \mathbf{1}_{\mathcal{E}_3(n,T)} \right) \le \frac{\beta+1}{\beta-1} \sum_{x=-(\ln n)^{1+\gamma}}^{nT-1} \tau_x \mathbf{1}_{\{\tau_x \le g(n)\}}.$$

Combining this bound with (2.3) and using Markov's inequality yields

$$\mathbb{P}_0\left(\mathcal{E}_3(n,T) \cap \mathcal{E}_4(n,T)^c\right) \le \frac{\beta+1}{(\beta-1)\bar{F}^{-1}(n^{-1}(\ln n)^{1/2})} \mathbf{E}\left(\sum_{x=-(\ln n)^{1+\gamma}}^{nT-1} \tau_i \mathbf{1}_{\{\tau_x \le g(n)\}} \right) = o(1).$$

On recalling the conclusion of Lemma 2.2, this completes the proof. □

As a consequence of the previous result, to deduce a scaling limit for the sequence $(\Delta_n)_{n\ge 0}$, it will suffice to study sums of the form $\sum_{x=1}^n T_x \mathbf{1}_{\{\tau_x > g(n)\}}$. In fact, the backtracking result of Lemma 2.2 will further allow us to replace T_x in this expression by

$$\tilde{T}_x := \int_{\Delta_x}^{\Delta_{x,(\ln x)^{1+\gamma}}} \mathbf{1}_{\{X_t = x\}} dt, \tag{2.5}$$

where $\Delta_{x,(\ln x)^{1+\gamma}}$ is the first time after Δ_x that X leaves the interval $[x-(\ln x)^{1+\gamma}, x+(\ln x)^{1+\gamma}]$. This is particularly useful because, by applying the fact that deep traps are separated by a

distance that is polynomial in n (see Lemma 2.1), it will be possible to decouple the random variables $(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}})_{x \geq 1}$ in such a way that enables us to deduce functional scaling results for their sums from those for independent sums proved in Section 5. Before commencing this program in Lemma 2.5, however, we derive a preliminary lemma that suitably describes the asymptotic behaviour of the distributional tail

$$\bar{F}_n(u) := \mathbb{P}_0 \left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}} > u \right).$$

(Clearly, the definition of $\bar{F}_n$ is independent of the particular $x \geq 1$ considered.)

Lemma 2.4. *For every $\varepsilon > 0$, there exists a constant c such that, for any $u \geq c(g(n) \vee 1)$,*

$$(1-\varepsilon)\bar{F}_n(u) \leq \bar{F}(u) \leq (1+\varepsilon)\bar{F}_n(u).$$

Proof. For $x \geq 1$, let $\tilde{G}(x)$ be the total number of visits of the embedded random walk Y to x up until the first time after Δ_x^Y that it leaves the interval $[x - (\ln n)^{1+\gamma}, x + (\ln n)^{1+\gamma}]$. Then, similarly to (2.2), we have that $\tilde{T}_x$ is distributed as $\tau_x \sum_{i=1}^{\tilde{G}(x)} \mathbf{e}_i$. Hence, setting $\Gamma := \sum_{i=1}^{\tilde{G}(x)} \mathbf{e}_i$, we can use the independence of Γ and τ_x under $\mathbb{P}_0$ to write

$$\begin{aligned} \bar{F}_n(u) &= \mathbb{P}_0 \left(\tau_x \Gamma > u,\, \tau_x > g(n) \right) \\ &= \int_0^{u/g(n)} \bar{F}\left(uv^{-1}\right) \mathbb{P}_0 \left(\Gamma \in dv \right) + \int_{u/g(n)}^{\infty} \bar{F}\left(g(n)\right) \mathbb{P}_0 \left(\Gamma \in dv \right). \end{aligned}$$

It follows that

$$\left| \frac{\bar{F}_n(u)}{\bar{F}(u)} - 1 \right| \leq \int_0^{\infty} \left| \frac{\bar{F}\left(uv^{-1}\right)}{\bar{F}(u)} - 1 \right| \mathbb{P}_0 \left(\Gamma \in dv \right) + \left(\frac{\bar{F}\left(g(n)\right)}{\bar{F}(u)} + 1 \right) \mathbb{P}_0 \left(\Gamma \geq u/g(n) \right). \tag{2.6}$$

The first term on the right-hand side of (2.6) is independent of n, and so it will be enough for our purposes to show that it converges to 0 as $u \to \infty$. To do this, first note that, by the monotonicity of $\bar{F}$, (1.1) holds uniformly for $v \in [v_0, v_1]$ for any $0 < v_0 \leq v_1 < \infty$. Hence, the lim sup as $u \to \infty$ of the term of interest is bounded above by

$$\mathbb{P}_0 \left(\Gamma \notin [v_0, v_1] \right) + \limsup_{u \to \infty} \int_0^{v_0} \frac{\bar{F}(uv^{-1})}{\bar{F}(u)} \mathbb{P}_0 \left(\Gamma \in dv \right) + \limsup_{u \to \infty} \int_{v_1}^{\infty} \frac{\bar{F}(uv^{-1})}{\bar{F}(u)} \mathbb{P}_0 \left(\Gamma \in dz \right),$$

for any $0 < v_0 \leq v_1 < \infty$. Now, if $v_0 < 1$, then $\bar{F}(uv^{-1}) \leq \bar{F}(u)$ for all $v \in [0, v_0]$, and so the first limsup is bounded above by $\mathbb{P}_0(\Gamma \leq v_0)$. Furthermore, if v_1 is chosen to be no less than 1, then we can apply the bound at (2.4) to estimate $\bar{F}(uv^{-1})/\bar{F}(u)$ by cv^{ε} for $v \geq v_1$. Thus

$$\limsup_{u \to \infty} \int_0^{\infty} \left| \frac{\bar{F}\left(uv^{-1}\right)}{\bar{F}(u)} - 1 \right| \mathbb{P}_0 \left(\Gamma \in dv \right) \leq 2\mathbb{P}_0 \left(\Gamma \notin [v_0, v_1] \right) + c \int_{v_1}^{\infty} v^{\varepsilon} \mathbb{P}_0 \left(\Gamma \in dv \right).$$

Since $\mathbb{E}_0(\Gamma^{\varepsilon}) \leq \mathbb{E}_0(1+\Gamma) = 1 + \mathbb{E}_0(\tilde{G}(x))\mathbb{E}_0(\mathbf{e}_1) < \infty$, by taking v_0 arbitrarily small and v_1 arbitrarily large, the upper bound here can be made arbitrarily small, meaning that

$$\lim_{u \to \infty} \int_0^{\infty} \left| \frac{\bar{F}\left(uv^{-1}\right)}{\bar{F}(u)} - 1 \right| \mathbb{P}_0 \left(\Gamma \in dv \right) = 0,$$

as desired.

For the second term on the right-hand side of (2.6), we apply (2.4) and Markov's inequality to deduce that, if $u \geq g(n)$, then

$$\left(\frac{\bar{F}\left(g(n)\right)}{\bar{F}(u)} + 1\right) \mathbb{P}_0\left(\Gamma \geq u/g(n)\right) \leq \left(c\left(\frac{u}{g(n)}\right)^{\varepsilon} + 1\right) \frac{g(n)\mathbb{E}_0\Gamma}{u},$$

where $\varepsilon \in (0,1)$ is fixed. Thus, since the above bound is small whenever $u/g(n)$ is large (it was already noted in the previous paragraph that Γ has a finite first moment), the proof is complete. □

Lemma 2.5. *As $n \to \infty$, the laws of the processes*

$$\left(\frac{1}{n} L\left(\sum_{x=1}^{nt} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right)_{t \geq 0}$$

under $\mathbb{P}_0$ converge weakly with respect to the Skorohod J_1 topology on $D([0,\infty),\mathbb{R})$ to the law of $(m(t))_{t\geq 0}$.

Proof. First, fix $T \in (0,\infty)$ and suppose $(f_x)_{x\geq 1}$ is a collection of bounded, continuous functions on $\mathbb{R}$. We then have that

$$\begin{aligned}
&\mathbb{E}_0\left(\mathbf{1}_{\mathcal{E}_1(n,T)} \prod_{x=1}^{nT} f_x\left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right) \\
&= \sum_B \mathbb{E}_0\left(\mathbf{1}_{\{\mathcal{D}\cap[1,nT]=B\}} \prod_{x=1}^{nT} f_x\left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right) \\
&= \sum_B \mathbb{E}_0\left(\prod_{x\in B} f_x\left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right) \mathbf{1}_{\{\tau_x > g(n)\}} \prod_{x\in[1,nT]\setminus B} f_x(0)\mathbf{1}_{\{\tau_x \leq g(n)\}}\right),
\end{aligned}$$

where the sums are over subsets $B \subseteq [1,nT]$ such that if $x_1, x_2 \in B$ and $x_1 \neq x_2$, then $|x_1 - x_2| > n^{\kappa}$. By applying the independence of traps at different sites and the disjointness of the intervals $([x-(\ln n)^{1+\gamma}, x+(\ln n)^{1+\gamma}])_{x\in B}$ for the relevant choices of B, the above sum can be rewritten as

$$\sum_B \prod_{x\in B} \mathbb{E}_0\left(f_x\left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}} \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right) \prod_{x\in[1,nT]\setminus B} \mathbb{E}_0\left(f_x(0)\mathbf{1}_{\{\tau_x \leq g(n)\}}\right).$$

In particular, it follows that

$$\mathbb{E}_0\left(\mathbf{1}_{\mathcal{E}_1(n,T)} \prod_{x=1}^{nT} f_x\left(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right) = \mathbb{E}_0\left(\mathbf{1}_{\mathcal{E}_1'(n,T)} \prod_{x=1}^{nT} f_x\left(\tilde{T}_x' \mathbf{1}_{\{\tau_x' > g(n)\}}\right)\right),$$

where we suppose that, under $\mathbb{P}_0$, the pairs of random variables $(\tilde{T}_x', \tau_x')$, $x \geq 1$, are independent and identically-distributed as $(\tilde{T}_1, \tau_1)$, and the event $\mathcal{E}_1'(n,T)$ is defined analogously to $\mathcal{E}_1(n,T)$ from these random variables. Consequently, under $\mathbb{P}_0$, the laws of $(\tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}})_{x=1}^{nT}$ conditional on $\mathcal{E}_1(n,T)$ and $(\tilde{T}_x' \mathbf{1}_{\{\tau_x' > g(n)\}})_{x=1}^{nT}$ conditional on $\mathcal{E}_1'(n,T)$ are identical.

By applying the conclusion of the previous paragraph, we obtain that, for any bounded function $H : D([0,T],\mathbb{R}) \to \mathbb{R}$ that is continuous with respect to the Skorohod J_1 topology,

$$\left|\mathbb{E}_0\left[H\left(\left(\frac{1}{n} L\left(\sum_{x=1}^{nt} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right)_{t\in[0,T]}\right)\right] - \mathbb{E}_0\left[H\left(\left(\frac{1}{n} L\left(\sum_{x=1}^{nt} \tilde{T}_x' \mathbf{1}_{\{\tau_x' > g(n)\}}\right)\right)_{t\in[0,T]}\right)\right]\right|$$

is bounded above by $2\|H\|_\infty \mathbf{P}\left(\mathcal{E}_1(n,T)^c\right)$. Since Lemma 2.1 tells us that this upper bound converges to 0 as $n \to \infty$, to complete the proof it will thus suffice to establish the result with $(\tilde{T}_x, \tau_x)_{x\geq 1}$ replaced by $(\tilde{T}'_x, \tau'_x)_{x\geq 1}$. However, because we are assuming that $(\tilde{T}'_x, \tau'_x)_{x\geq 1}$ are independent, the tail asymptotics proved in Lemma 2.4 allow us to derive the relevant scaling limit for the sums involving $(\tilde{T}'_x, \tau'_x)_{x\geq 1}$ by a simple application of Theorem 5.1 (with $h_1(n) = \ln n$ and $h_2(n) = 0$). □

We are now in a position to prove Theorem 1.5 by showing that the rescaled sums considered in the previous lemma suitably well approximate the sequence $(\Delta_n)_{n\geq 1}$.

Proof of Theorem 1.5. Fix $T \in (0,\infty)$ and observe that, on $\mathcal{E}_2(n) \cap \mathcal{E}_3(n,T) \cap \mathcal{E}_4(n,T)$, we have that

$$\sum_{x=1}^{nt-(\ln n)^{1+\gamma}} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}} \leq \Delta_{nt} \leq \sum_{x=1}^{nt} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}} + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2}), \qquad \forall t \in [0,T]. \quad (2.7)$$

By reparameterising the time-scales in the obvious way, it is clear that

$$d_{J_1}\left(\left(\frac{1}{n}L\left(\sum_{x=1}^{nt-(\ln n)^{1+\gamma}} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right)_{t\in[0,T]}, \left(\frac{1}{n}L\left(\sum_{x=1}^{nt} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right)\right)_{t\in[0,T]}\right), \quad (2.8)$$

where d_{J_1} is the Skorohod J_1 distance on $D([0,T],\mathbb{R})$ (as defined in the appendix at (A.1)), is bounded above by

$$\frac{(\ln n)^{1+\gamma}}{n} + \frac{1}{n}L\left(\sum_{x=1}^{nT} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right) - \frac{1}{n}L\left(\sum_{x=1}^{n(T-\varepsilon)} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}\right),$$

for large n. (Note that the first term above relates to the distortion of the time scale needed to compare the two processes.) By Lemma 2.5, this bound converges in distribution under $\mathbb{P}_0$ to $m(T) - m(T-\varepsilon)$. Now, in the limit as $\varepsilon \to 0$, $m(T) - m(T-\varepsilon)$ converges to 0 in probability. It readily follows that, as $n \to \infty$, so does the expression at (2.8). Hence, the theorem will follow from Lemmas 2.1, 2.2, 2.3 and 2.5, if we can show that

$$\sup_{x\in[0,\Xi]} \frac{1}{n}\left|L\left(x + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2})\right) - L(x)\right| \to 0,$$

in $\mathbb{P}_0$ probability, where $\Xi := \sum_{x=1}^{nT} \tilde{T}_x \mathbf{1}_{\{\tau_x > g(n)\}}$. To check this, we start by noting that Lemma 2.5 implies, for any $\lambda > 0$,

$$\mathbb{P}_0\left(\Xi \leq \bar{F}^{-1}(1/n\lambda)\right) = \mathbb{P}_0\left(n^{-1}L(\Xi) \leq \lambda\right) \to \mathbb{P}_0(m(T) \leq \lambda).$$

By choosing λ suitably large, the limiting probability can be made arbitrarily close to 1. Thus the problem reduces to showing that, for any $\lambda \in (0,\infty)$,

$$\sup_{x\in[0,\bar{F}^{-1}(1/n\lambda)]} \frac{1}{n}\left|L\left(x + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2})\right) - L(x)\right| \to 0.$$

Let $\varepsilon \in (0, \lambda)$, then, since $\bar{F}^{-1}(n^{-1}(\ln n)^{1/2}) \leq \bar{F}^{-1}(1/n\varepsilon)$ for large enough n, we have that

$$\begin{aligned}
&\sup_{x \in [\bar{F}^{-1}(1/n\varepsilon), \bar{F}^{-1}(1/n\lambda)]} \frac{1}{n} \left| L\left(x + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2})\right) - L(x) \right| \\
&\quad \leq \quad \frac{1}{n} L\left(\bar{F}^{-1}(1/n\lambda)\right) \sup_{x \geq \bar{F}^{-1}(1/n\varepsilon)} \left| \frac{L(x + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2}))}{L(x)} - 1 \right| \\
&\quad \leq \quad \lambda \sup_{x \geq \bar{F}^{-1}(1/n\varepsilon)} \left| \frac{L(2x)}{L(x)} - 1 \right|,
\end{aligned}$$

which converges to 0 as $n \to \infty$ by (1.1). Moreover, we also have that

$$\sup_{x \in [0, \bar{F}^{-1}(1/n\varepsilon)]} \frac{1}{n} \left| L\left(x + \bar{F}^{-1}(n^{-1}(\ln n)^{1/2})\right) - L(x) \right| \leq \frac{1}{n} L(2\bar{F}^{-1}(1/n\varepsilon)) \sim \frac{1}{n} L(\bar{F}^{-1}(1/n\varepsilon)),$$

where the asymptotic equivalence is an application of (1.1). In particular, since the right-hand side above is equal to ε, which can be chosen arbitrarily small, the result follows. □

From this, the proof of Corollary 1.6 is relatively straightforward.

Proof of Corollary 1.6. Define $X^* = (X_t)_{t \geq 0}$ to be the running supremum of X, i.e. $X_t^* := \max_{s \leq t} X_s$. Since $X_t^* \geq n$ if and only if $\Delta_n \leq t$, we obtain that $(X_t^* + 1)_{t \geq 0}$ is the inverse of $(\Delta_n)_{n \geq 0}$ in the sense described at (1.9). Thus, because the inverse map is continuous with respect to the Skorohod M_1 topology (at least on the subset of functions $f \in D([0, \infty), \mathbb{R})$ that satisfy $\limsup_{t \to \infty} f(t) = \infty$, see [30]), it is immediate from Theorem 1.5 that, as $n \to \infty$, the laws of the processes

$$\left(\frac{1}{n} X^*_{\bar{F}^{-1}(1/nt)} \right)_{t \geq 0}$$

under $\mathbb{P}_0$ converge weakly with respect to the Skorohod M_1 topology on $D([0, \infty), \mathbb{R})$ to the law of $(m^{-1}(t))_{t \geq 0}$. Thus, to complete the proof, it will suffice to demonstrate that, for any $T \in (0, \infty)$,

$$\sup_{t \in [0,T]} \frac{1}{n} \left| X^*_{\bar{F}^{-1}(1/nt)} - X_{\bar{F}^{-1}(1/nt)} \right| \to 0$$

in $\mathbb{P}_0$-probability as $n \to \infty$. To do this, we first fix $T \in (0, \infty)$ and set $N := nT \ln(nT)$. Theorem 1.5 then implies that $\mathbb{P}_0\left(\Delta_N \geq \bar{F}^{-1}(1/nT)\right) \to 1$ as $n \to \infty$. Moreover, on the set $\{\Delta_N \geq \bar{F}^{-1}(1/nT)\}$, it is the case that

$$\sup_{t \in [0,T]} \left| X^*_{\bar{F}^{-1}(1/nt)} - X_{\bar{F}^{-1}(1/nt)} \right| \leq \sup_{k \leq \Delta_N^Y} (Y_k^* - Y_k),$$

where Y^* is the running supremum of Y. Hence

$$\begin{aligned}
&\limsup_{n \to \infty} \mathbb{P}_0 \left(\sup_{t \in [0,T]} \frac{1}{n} \left| X^*_{\bar{F}^{-1}(1/nt)} - X_{\bar{F}^{-1}(1/nt)} \right| > \varepsilon \right) \\
&\quad \leq \quad \limsup_{n \to \infty} \mathbb{P}_0 \left(\frac{1}{n} \sup_{k \leq \Delta_N^Y} (Y_k^* - Y_k) > \varepsilon \right) \\
&\quad \leq \quad \limsup_{n \to \infty} \mathbb{P}_0 \left(n^{-1} (\ln N)^{1+\gamma} > \varepsilon, \mathcal{E}_3(N, 1) \right) \\
&\quad = \quad 0,
\end{aligned}$$

where we have applied the fact that $\mathbb{P}_0(\mathcal{E}_3(N, 1)^c) \to 0$, which is the conclusion of Lemma 2.2, and also that $n^{-1}(\ln N)^{1+\gamma} \to 0$, which is clear from the definition of N. □

3 Biased random walk on critical Galton-Watson trees

In this section, we explain how techniques similar to those of the previous section can be used to deduce the corresponding asymptotics for a biased random walk on a critical Galton-Watson tree conditioned to survive. Prior to proving our main results (Theorem 1.1 and Corollary 1.2), however, we proceed in the next two subsections to derive certain properties regarding the structure of the tree $\mathcal{T}^*$ and deduce some preliminary simple random walk estimates, respectively. These results establish information in the present setting that is broadly analogous to that contained in Lemmas 2.1-2.4 for the directed trap model.

3.1 Structure of the infinite tree

A key tool throughout this study is the spinal decomposition of $\mathcal{T}^*$ that appears as [23, Lemma 2.2], and which can be described as follows. First, $\mathbf{P}$-a.e. realisation of $\mathcal{T}^*$ admits a unique non-intersecting infinite path starting at the root. Conditional on this 'backbone', the number of children of vertices on the backbone are independent, each distributed as a size-biased random variable $\tilde{Z}$, which satisfies

$$\mathbf{P}\left(\tilde{Z}=k\right)=k\mathbf{P}(Z=k), \qquad k\geq 1. \tag{3.1}$$

Moreover, conditional on the backbone and the number of children of each backbone element, the trees descending from the children of backbone vertices that are not on the backbone are independent copies of the original critical branching process $\mathcal{T}$. To fix notation and terminology for this decomposition, we will henceforth suppose that $\mathcal{T}^*$ has been built by starting with a semi-infinite path, $\{\rho=\rho_0,\rho_1,\rho_2,\dots\}$ – this will form the backbone of $\mathcal{T}^*$. Then, after selecting $(\tilde{Z}_i)_{i\geq 0}$ independently with distribution equal to that of $\tilde{Z}$, to each backbone vertex ρ_i, we attach a collection of 'buds' ρ_{ij}, $j=1,\dots,\tilde{Z}_i-1$. Finally, we grow from each bud ρ_{ij} a 'leaf' $\mathcal{T}_{ij}$, that is, a Galton-Watson tree with initial ancestor ρ_{ij} and offspring distribution Z. See Figure 1 for a graphical representation of these definitions.

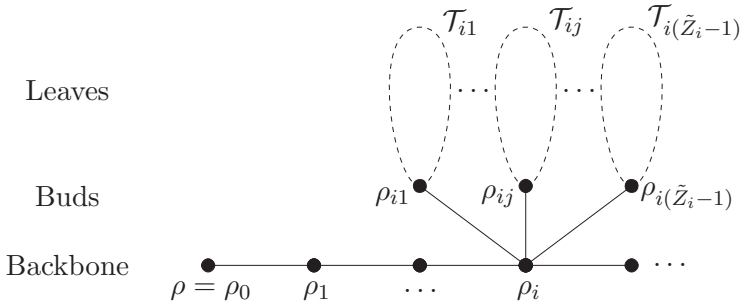

Figure 1: Decomposition of $\mathcal{T}^*$.

With this picture, it is clear how we can view $\mathcal{T}^*$ as an essentially one-dimensional trap model with the backbone playing the role of $\mathbb{Z}$ in the previous section. Rather than having an exponential holding time at each vertex ρ_i, however, we have a random variable representing the time it takes X to leave the tree $\mathcal{T}_i:=\{\rho_i\}\cup(\cup_{j=1,\dots,\tilde{Z}_i-1}\mathcal{T}_{ij})$ starting from ρ_i. As will be made precise later, key to determining whether this time is likely to be large or not are the heights of the leaves connected to ρ_i. For this reason, the rest of this section will be taken up with an investigation into the big, or perhaps more accurately tall, leaves of $\mathcal{T}^*$.

More concretely, we start by introducing a sequence of critical heights $(h_n)_{n\geq 1}$ by setting $h_n:=n(\ln n)^{-1}$ (roughly, β^{h_n} will play the role that the $g(n)$ introduced at (2.1) did in the

previous section), and define, for each $i \geq 0$,

$$N_n(i) := \# \left\{ 1 \leq j \leq \tilde{Z}_i - 1 : \, h(\mathcal{T}_{ij}) \geq h_n \right\},$$

where $h(\mathcal{T}_{ij})$ is the height of the tree $\mathcal{T}_{ij}$, so that $N_n(i)$ counts the number of big leaves emanating from the backbone vertex ρ_i. The random variables in the collection $(N_n(i))_{i \geq 0}$ are independent and identically-distributed. Moreover, it is possible to describe the asymptotic probability that one of these random variables is equal to zero, i.e. there is no big leaf at the relevant site.

Lemma 3.1. *Let $\alpha \in (1, 2]$. As $n \to \infty$, we have that*

$$\mathbf{P}\left(N_n(0) = 0\right) \sim 1 - \frac{\alpha}{(\alpha - 1)h_n}.$$

Proof. By conditioning on the number of buds attached to the root, we have

$$\mathbf{P}\left(N_n(0) = 0\right) = \mathbf{E}\left((1 - q_{h_n})^{\tilde{Z}-1} \right),$$

where, as introduced above (1.6), q_k is the probability that an unconditioned branching process with offspring distribution Z survives for at least k generations. By the size-biasing of (3.1), this can be rewritten as

$$\mathbf{P}\left(N_n(0) = 0\right) = \mathbf{E}\left(Z\,(1 - q_{h_n})^{Z-1} \right) = f'\left(1 - q_{h_n}\right),$$

where f' is the derivative of the generating function f, as defined at (1.5). Now, by [29, (2.1)], it holds that $f'(1 - x) \sim 1 - \alpha x^{\alpha - 1} L(x)$ as $x \to 0^+$, and so

$$\mathbf{P}\left(N_n(0) = 0\right) \sim 1 - \alpha q_{h_n}^{\alpha - 1} L(q_{h_n}).$$

From this, the proof is completed by recalling the tail decay at (1.6). □

It will be important for our future arguments that the sites from which big leaves emanate are not too close together, and that there are no big traps close to ρ. The final lemma of this section demonstrates that the sequence of critical heights we have chosen achieves this.

Lemma 3.2. *Let $\alpha \in (1, 2]$, $T \in (0, \infty)$ and $\varepsilon \in (0, 1)$. As $n \to \infty$,*

$$\mathbf{P}\left(\sum_{i=m}^{m+n^{\varepsilon}} \mathbf{1}_{\{N_n(i) \geq 1\}} \geq 2 \text{ for some } m \in \{0, 1, \ldots, Tn - n^{\varepsilon}\} \right) \to 0$$

and also

$$\mathbf{P}\left(\sum_{i=0}^{n^{\varepsilon}} \mathbf{1}_{\{N_n(i) \geq 1\}} \geq 1 \right) \to 0.$$

Proof. This is essentially the same as Lemma 2.1. □

3.2 Initial random walk estimates

This section collects together some preliminary results for the biased random walk $(X_m)_{m\geq 0}$ on $\mathcal{T}^*$, regarding in particular: the amount of backtracking performed by the embedded biased random walk on the backbone; the amount of time X spends in small leaves; the amount of time X spends close to the base of big leaves; and tail estimates for the amount of time X spends deep within big leaves.

To begin with, we introduce $Y = (Y_n)_{n\geq 1}$ to represent the jump process of $\pi(X)$, where $\pi : \mathcal{T}^* \to \{\rho_0, \rho_1, \dots\}$ is the projection onto the backbone, i.e. $\pi(x) = \rho_i$ for $x \in \mathcal{T}_i$. More precisely, set $S(0) = 0$,

$$S(n) = \inf\{m > S(n-1) : \pi(X_m) \neq \pi(X_{m-1})\}, \qquad \forall n \geq 1,$$

and then define $Y_n := X_{S(n)}$. From this construction, it is clear that, under either the quenched or annealed law, Y is simply a biased random walk on the semi-infinite line graph $\{\rho_0, \rho_1, \dots\}$, and so, as in the previous section, we can control the amount it backtracks. In particular, if we let

$$\Delta_n^Y := \inf\{m \geq 0 : \, Y_m = \rho_n\}$$

be the first time that the embedded random walk Y reaches level n along the backbone, then we have the following result, which is simply a restatement of Lemma 2.2. We recall that $d_{\mathcal{T}^*}$ is the shortest path graph distance on $\mathcal{T}^*$.

Lemma 3.3. *Let $\alpha \in (1, 2]$, $T \in (0, \infty)$ and $\gamma > 0$. As $n \to \infty$,*

$$\mathbb{P}_\rho\left(\min_{0\leq i<j\leq \Delta_{nT}^Y} (d_{\mathcal{T}^*}(\rho_0, Y_j) - d_{\mathcal{T}^*}(\rho_0, Y_i)) \leq -(\ln n)^{1+\gamma}\right) \to 0.$$

Our next goal is to show that the time the biased random walk X spends outside of the big leaves of $\mathcal{T}^*$ is unimportant, where we define the set of vertices in big leaves to be

$$\mathcal{B} := \left\{x \in \mathcal{T}_{ij} : \, i \geq 0, \, 1 \leq j \leq \tilde{Z}_i - 1, \, h(\mathcal{T}_{ij}) \geq h_n\right\}.$$

Key to doing this is the following equality, which is obtained by applying standard results for weighted random walks on graphs (cf. [6, Lemma 3.1]):

$$E_{\rho_{ij}}^{\mathcal{T}^*} \tau_{\rho_i} = 1 + \beta^{-i} \sum_{\substack{x,y\in\mathcal{T}_{ij}:\\ x\sim y}} c(x, y), \tag{3.2}$$

where for a vertex $x \in \mathcal{T}^*$, we define $\tau_x := \inf\{m \geq 0 : \, X_m = x\}$. For the statement of the next lemma, which is approximately analogous to Lemma 2.3, we recall the definition of $(\Delta_n)_{n\geq 0}$ from (1.8).

Lemma 3.4. *Let $\alpha \in (1, 2]$, $T \in (0, \infty)$ and $\varepsilon > 0$. As $n \to \infty$,*

$$\mathbb{P}_\rho\left(\sum_{m\leq \Delta_{nT}} \mathbf{1}_{\{X_m \notin \mathcal{B}\}} \geq \beta^{h_n(1+\varepsilon)}\right) \to 0.$$

Proof. We start by estimating the quenched expectation of the time X spends in a particular small leaf before reaching level nT along the backbone. Thus, suppose we have a leaf $\mathcal{T}_{ij}$ such that $i < nT$ and $h(\mathcal{T}_{ij}) < h_n$. Starting from the vertex ρ_i, the probability of hitting ρ_{ij} before ρ_{nT} can be computed exactly, by elementary means, as

$$P_{\rho_i}^{\mathcal{T}^*}\left(\tau_{\rho_{ij}} < \tau_{\rho_{nT}}\right) = \frac{1+\beta^{-1}+\cdots+\beta^{i+1-nT}}{1+(1+\beta^{-1}+\cdots+\beta^{i+1-nT})} \leq \frac{1}{2-\beta^{-1}}.$$

This means that the number of separate visits X makes to $\mathcal{T}_{ij}$ is stochastically dominated by a geometric random variable with parameter $1-(2-\beta^{-1})^{-1}$, and so its mean is bounded above by $\beta/(\beta-1)$. Moreover, the equality at (3.2) and our assumption on $h(\mathcal{T}_{ij})$ imply that, on each visit to $\mathcal{T}_{ij}$, the amount of time X spends there is bounded above by

$$E_{\rho_{ij}}^{\mathcal{T}^*}\tau_{\rho_i} \leq 1 + 2\beta^{h_n}\#\mathcal{T}_{ij},$$

where $\#\mathcal{T}_{ij}$ is the total number of vertices in $\mathcal{T}_{ij}$. Hence

$$E_{\rho}^{\mathcal{T}^*}\left(\sum_{m\leq\Delta_{nT}} \mathbf{1}_{\{X_m\in\mathcal{T}_{ij}\}}\right) \leq \frac{\beta}{\beta-1}\left(1+2\beta^{h_n}\#\mathcal{T}_{ij}\right). \tag{3.3}$$

As for the estimating time spent at a vertex ρ_i, where $0 < i < nT$, we start by noting that the total number of returns to ρ_i is a geometric random variable. Moreover, its parameter $P_{\rho_i}^{\mathcal{T}^*}(\tau_{\rho_i}^+ = \infty)$, where $\tau_{\rho_i}^+ := \inf\{m > 0 : X_m = \rho_i\}$ is the first return time to ρ_i, can easily be bounded below by the probability that X jumps from ρ_i to ρ_{i+1} on its first step times the probability that a biased random walk on $\mathbb{Z}$ never hits the vertex to the left of its starting point. Since the first of these quantities is given by $\beta/(\beta\tilde{Z}_i+1)$ and the second is equal to $1-\beta^{-1}$, it follows that

$$E_{\rho}^{\mathcal{T}^*}\left(\sum_{m\leq\Delta_{nT}} \mathbf{1}_{\{X_m=\rho_i\}}\right) \leq c_\beta\tilde{Z}_i. \tag{3.4}$$

A similar argument applies for $i = 0$.

Piecing together the estimates at (3.3) and (3.4), we thus obtain

$$E_{\rho}^{\mathcal{T}^*}\left(\sum_{m\leq\Delta_{nT}} \mathbf{1}_{\{X_m\notin\mathcal{B}\}}\right) \leq c_\beta\beta^{h_n}\sum_{i=0}^{nT-1}\tilde{Z}_i\left[1+\max_{j=1,\ldots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right], \tag{3.5}$$

where c_β is a constant depending only on β. Now, to bound the summands, we consider the following probabilistic upper bound

$$\mathbf{P}\left(\tilde{Z}_i\left[1+\max_{j=1,\ldots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]\geq k\right) \leq \mathbf{P}\left(\tilde{Z}_i\geq k^{1/2}\right)+\mathbf{P}\left(\max_{j=1,\ldots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\geq k^{1/2}-1\right). \tag{3.6}$$

For the first of these terms, we apply the size-biasing of (3.1) and Markov's inequality to deduce

$$\mathbf{P}\left(\tilde{Z}_i\geq k^{1/2}\right) \leq \frac{\mathbf{E}(Z^{1+\alpha'})}{k^{\alpha'/2}}. \tag{3.7}$$

Since the expectation in (3.7) is finite for any $\alpha' \in (0, \alpha-1)$ (see [19, Section 35], for example), we fix an α' in this range to obtain a polynomial bound for the relevant probability. For the

second term of (3.6), we first condition on $\tilde{Z}_i$ to obtain

$$\begin{aligned}\mathbf{P}\left(\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\geq k^{1/2}-1\right) &= 1-\mathbf{E}\left(\left(1-\mathbf{P}\left(\#\mathcal{T}\geq k^{1/2}-1\right)\right)^{\tilde{Z}-1}\right)\\ &= 1-\mathbf{E}\left(Z\left(1-\mathbf{P}\left(\#\mathcal{T}\geq k^{1/2}-1\right)\right)^{Z-1}\right)\\ &= 1-f'\left(1-\mathbf{P}\left(\#\mathcal{T}\geq k^{1/2}-1\right)\right).\end{aligned}$$

From the proof of Lemma 3.1, we know that $f'(1-x)\sim 1-\alpha x^{\alpha-1}L(x)$ as $x\to 0^+$, and so

$$\mathbf{P}\left(\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\geq k^{1/2}-1\right)\sim\alpha\mathbf{P}\left(\#\mathcal{T}\geq k^{1/2}-1\right)^{\alpha-1}L\left(\mathbf{P}\left(\#\mathcal{T}\geq k^{1/2}-1\right)\right),\quad(3.8)$$

as $k\to\infty$. To establish a bound for $\mathbf{P}(\#\mathcal{T}\geq k)$ that decays polynomially quickly, first note that $\mathbf{P}(\#\mathcal{T}=k)=k^{-1}\mathbf{P}(S_k=-1)$, where $(S_k)_{k\geq 0}$ is a random walk on $\mathbb{Z}$ with step distribution $Z-1$ (see [13]). Moreover, by the local limit theorem of [19, Section 50], it is the case that $\mathbf{P}(S_k=-1)\sim ca_k^{-1}$, where a_k are the constants appearing in (1.4). Since $a_k\sim k^{1/\alpha}\ell(k)$ for some slowly varying function ℓ (see [19, Section 35], for example), it follows that if $\alpha'\in(0,1/\alpha)$, then there exists a constant c such that $\mathbf{P}(|\mathcal{T}|\geq k)\leq ck^{-\alpha'}$. Combining this estimate with (3.6), (3.7) and (3.8), we obtain that there exist constants c and $\delta>0$ such that

$$\mathbf{P}\left(\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]\geq k\right)\leq ck^{-\delta}.\quad(3.9)$$

Consequently, recalling (3.5),

$$\begin{aligned}&\mathbf{P}\left(\sum_{m\leq\Delta_{nT}}\mathbf{1}_{\{X_m\notin\mathcal{B}\}}\geq\beta^{h_n(1+\varepsilon)}\right)\\ &\leq \mathbf{P}\left(\sum_{m\leq\Delta_{nT}}\mathbf{1}_{\{X_m\notin\mathcal{B}\}}\geq nE_\rho^{\mathcal{T}^*}\left(\sum_{m\leq\Delta_{nT}}\mathbf{1}_{\{X_m\notin\mathcal{B}\}}\right)\right)\\ &\qquad+\mathbf{P}\left(E_\rho^{\mathcal{T}^*}\left(\sum_{m\leq\Delta_{nT}}\mathbf{1}_{\{X_m\notin\mathcal{B}\}}\right)\geq n^{-1}\beta^{h_n(1+\varepsilon)}\right)\\ &\leq n^{-1}+\mathbf{P}\left(\max_{i=0,\dots,nT-1}\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]\geq\frac{1}{c_\beta n^2T}\beta^{\varepsilon h_n}\right)\\ &\leq n^{-1}+nT\mathbf{P}\left(\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]\geq\frac{1}{c_\beta n^2T}\beta^{\varepsilon h_n}\right)\\ &\leq n^{-1}+cn^{1+2\delta}\beta^{-\varepsilon\delta h_n},\end{aligned}$$

and this converges to 0 as $n\to\infty$. □

The result above means that, in establishing the distributional convergence of Δ_n, we only have to consider the time the random walk X spends in big leaves. In fact, as we will now show, the time spent close to the backbone in big leaves is also negligible. To this end, let us start by introducing some notation and formalising some terminology. First, we will write y_{ij} for the

deepest vertex in $\mathcal{T}_{ij}$; that is, the vertex that maximises the distance from the root ρ_{ij}. So that this notion is well-defined, if there is more than one vertex at the deepest level of $\mathcal{T}_{ij}$, we choose y_{ij} to be the first in the usual lexicographical ordering of $\mathcal{T}_{ij}$, assuming that the offspring of each vertex have been labelled according to birth order. If the tree $\mathcal{T}_{ij}$ has height greater than or equal to h_n, then for a fixed $\delta \in (0,1)$ it is possible to define a unique vertex on the path from ρ_{ij} to y_{ij} at level h_n^δ in $\mathcal{T}_{ij}$. We shall denote this vertex x_{ij} and call it the 'entrance' to the leaf $\mathcal{T}_{ij}$. When we say that the leaf $\mathcal{T}_{ij}$ has been visited deeply, we will mean that X has hit x_{ij}. Moreover, by the 'time spent in the lower part of a big leaves emanating from ρ_i', we will mean

$$t_i' = \sum_{j=1}^{\tilde{Z}_i - 1} \mathbf{1}_{\{h(\mathcal{T}_{ij}) \geq h_n\}} \sum_{m=0}^{\infty} \mathbf{1}_{\{X_m \in \mathcal{T}_{ij} \setminus \mathcal{T}_{ij}(x_{ij})\}}, \tag{3.10}$$

where $\mathcal{T}_{ij}(x_{ij})$ is the part of the tree $\mathcal{T}_{ij}$ descending from the entrance x_{ij}.

To control the random variables $(t_i')_{i \geq 0}$ (which are identically-distributed apart from $i = 0$), we need to consider the structure of the trees $\mathcal{T}_{ij}' := \mathcal{T}_{ij} \setminus \mathcal{T}_{ij}(x_{ij})$, and for this, the construction of a Galton-Watson tree conditioned on its height given in [18] is helpful. In particular, in Section 2 of that article, the following algorithm is described. First, let (ξ_n, ζ_n), $n \geq 0$, be a sequence of independent pairs of random variables, with distribution given by

$$\mathbf{P}\left(\xi_{n+1} = j, \zeta_{n+1} = k\right) = c_n p_k \left(1 - q_n\right)^{j-1} \left(1 - q_{n+1}\right)^{k-j},$$

(recall that $q_n = \mathbf{P}(Z_n > 0)$ is the probability that the unconditioned branching process survives for at least n generations) for $1 \leq j \leq k$, where

$$c_n := \frac{\mathbf{P}\left(h(\mathcal{T}) = n\right)}{\mathbf{P}\left(h(\mathcal{T}) = n + 1\right)}. \tag{3.11}$$

Then, let $\tilde{\mathcal{T}}_0$ be a Galton-Watson tree of height 0, i.e. consisting solely of a root vertex, and, to construct $\tilde{\mathcal{T}}_{n+1}$, $n \geq 0$:

- let the first generation size of $\tilde{\mathcal{T}}_{n+1}$ be ζ_{n+1},
- let $\tilde{\mathcal{T}}_n$ be the subtree founded by the ξ_{n+1}th first generation particle of $\tilde{\mathcal{T}}_{n+1}$,
- attach independent Galton-Watson trees conditioned on having height strictly less than n to the $\xi_{n+1} - 1$ siblings to the left of the distinguished first generation particle,
- attach independent Galton-Watson trees conditioned on height strictly less that $n + 1$ to the $\zeta_{n+1} - \xi_{n+1}$ siblings to the right of the distinguished first generation particle.

It is shown in [18] that the tree $\tilde{\mathcal{T}}_n$ that results from this procedure has the same probabilistic structure as $\mathcal{T}$ conditioned to have height exactly equal to n. Before considering the implications of this result for the times $(t_i')_{i \geq 0}$, we derive the asymptotics of the constants $(c_n)_{n \geq 1}$ in our setting.

Lemma 3.5. *Let $\alpha \in (1, 2]$. The constants $(c_n)_{n \geq 1}$, as defined at (3.11), satisfy*

$$c_n \sim 1 + \frac{\alpha}{(\alpha - 1)n},$$

as $n \to \infty$.

Proof. First note that $\mathbf{P}(h(\mathcal{T}) = n) = q_n - q_{n+1}$. Moreover, if $f^{(n)}$ is the n-fold iteration of the generating function f, then we can write $q_n = 1 - f^{(n)}(0)$. It follows that

$$c_n = \frac{f^{(n+1)}(0) - f^{(n)}(0)}{f^{(n+2)}(0) - f^{(n+1)}(0)} = \frac{f(1-q_n) - 1 + q_n}{f(1-q_{n+1}) - 1 + q_{n+1}} = \frac{q_n^\alpha L(q_n)}{q_{n+1}^\alpha L(q_{n+1})} = \frac{q_n}{q_{n+1}} \times \frac{q_n^{\alpha-1} L(q_n)}{q_{n+1}^{\alpha-1} L(q_{n+1})},$$

where we have applied (1.5) to deduce the third equality. Now, by (1.6), the second term on the right-hand side satisfies

$$\frac{q_n^{\alpha-1} L(q_n)}{q_{n+1}^{\alpha-1} L(q_{n+1})} \sim \frac{n+1}{n} = 1 + \frac{1}{n}. \tag{3.12}$$

For the first term, again applying (1.5) and (1.6), it is the case that

$$\frac{q_n}{q_{n+1}} = \frac{q_n}{1 - f(1-q_n)} = \frac{1}{1 - q_n^{\alpha-1} L(q_n)} \sim 1 + \frac{1}{(\alpha-1)n}. \tag{3.13}$$

Multiplying the right-hand sides of (3.12) and (3.13) yields the result. □

Lemma 3.6. *Let $\alpha \in (1,2]$ and $T \in (0,\infty)$. As $n \to \infty$,*

$$\mathbb{P}_\rho \left(\sum_{i=0}^{nT-1} t'_i \geq \beta^{h_n} \right) \to 0.$$

Proof. Our first aim will be to show that

$$\mathbf{P}\left(h(\mathcal{T}'_{ij}) > x \middle| h(\mathcal{T}_{ij}) \geq h_n \right) \leq h_n^\delta \left(1 - \left(\inf_{h \geq h_n - h_n^\delta} c_h \right) f'\left(1 - q_{x - h_n^\delta} \right) \right), \tag{3.14}$$

for $h_n^\delta < x < h_n$. Fix an x in this range, and suppose for the moment that $h(\mathcal{T}_{ij}) = h \geq h_n$, so that x_{ij} is defined. Denote the path from ρ_{ij} to x_{ij} by $\rho_{ij} = w_0, w_1, \ldots, w_{h_n^\delta} = x_{ij}$. Now, remove the edges $\{w_{l-1}, w_l\}$, $l = 1, \ldots, h_n^\delta$ from $\mathcal{T}'_{ij}$, and denote by $\mathcal{T}_{ijl}$ the connected component containing w_l, so that $\mathcal{T}'_{ij}$ (minus the relevant edges) is the disjoint union of $\mathcal{T}_{ijl}$ over $l = 0, \ldots, h_n^\delta - 1$. From the procedure for constructing a Galton-Watson tree conditioned on its height described before Lemma 3.5, we deduce

$$\begin{aligned}
\mathbf{P}\left(h(\mathcal{T}'_{ij}) > x \middle| h(\mathcal{T}_{ij}) = h \right) &= \mathbf{P}\left(\max_{l=0,\ldots,h_n^\delta - 1} (h(\mathcal{T}_{ijl}) + l) > x \middle| h(\mathcal{T}_{ij}) = h \right) \\
&\leq \mathbf{P}\left(\max_{l=0,\ldots,h_n^\delta - 1} h(\mathcal{T}_{ijl}) > x + 1 - h_n^\delta \middle| h(\mathcal{T}_{ij}) = h \right) \\
&= 1 - \prod_{l=0}^{h_n^\delta - 1} \left[1 - \mathbf{P}\left(h(\mathcal{T}_{ijl}) > x + 1 - h_n^\delta \middle| h(\mathcal{T}_{ij}) = h \right) \right].
\end{aligned}$$

Moreover, if we suppose that $\mathcal{T}_{ij}$ conditioned on its height being equal to h has been built from

the random variables (ξ_n, ζ_n), $n \geq 0$, then we can write

$$\begin{aligned}
&\mathbf{P}\left(h(\mathcal{T}_{ijl}) < x+1-h_n^\delta \middle| h(\mathcal{T}_{ij}) = h\right) \\
&= \mathbf{E}\left(\mathbf{P}\left(h(\mathcal{T}_{ijl}) < x+1-h_n^\delta \middle| h(\mathcal{T}_{ij}) = h, \xi_{h-l}, \zeta_{h-l}\right)\right) \\
&= \mathbf{E}\left(\mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta \middle| h(\mathcal{T}) < h-l-1\right)^{\xi_{h-l}-1}\right. \\
&\qquad\qquad \left.\times\, \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta \middle| h(\mathcal{T}) < h-l\right)^{\zeta_{h-l}-\xi_{h-l}}\right) \\
&= \sum_{1\leq j\leq k} c_{h-l-1}p_k \mathbf{P}\left(h(\mathcal{T}) < h-l-1\right)^{j-1} \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta \middle| h(\mathcal{T}) < h-l-1\right)^{j-1} \\
&\qquad\qquad \times \mathbf{P}\left(h(\mathcal{T}) < h-l\right)^{k-j} \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta \middle| h(\mathcal{T}) < h-l\right)^{k-j} \\
&= \sum_{1\leq j\leq k} c_{h-l-1}p_k \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta\right)^{j-1} \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta\right)^{k-j} \\
&= \sum_{k\geq 1} c_{h-l-1} k p_k \mathbf{P}\left(h(\mathcal{T}) < x-h_n^\delta\right)^{k-1} \\
&= c_{h-l-1} f'\left(1-q_{x-h_n^\delta}\right).
\end{aligned}$$

Thus, combining these deductions, we obtain

$$\begin{aligned}
\mathbf{P}\left(h(\mathcal{T}'_{ij}) > x \middle| h(\mathcal{T}_{ij}) = h\right) &\leq 1 - \left(c_{h-l-1} f'\left(1-q_{x-h_n^\delta}\right)\right)^{h_n^\delta} \\
&\leq h_n^\delta \left(1 - \inf_{h' \geq h_n - h_n^\delta} c_{h'} f'\left(1-q_{x-h_n^\delta}\right)\right),
\end{aligned}$$

and, since this bound is independent of $h \geq h_n$, the bound at (3.14) follows.

Now, by arguing similarly to (3.5), it is possible to check that

$$E_\rho^{\mathcal{T}^*}\left(\sum_{i=0}^{nT-1} t'_i\right) \leq c_\beta \sum_{i=0}^{nT-1} \sum_{j=1}^{\tilde{Z}_i-1} \mathbf{1}_{\{h(\mathcal{T}_{ij})\geq h_n\}} \beta^{h(\mathcal{T}'_{ij})} \#\mathcal{T}_{ij},$$

where c_β is a constant depending only upon β. Thus, following the end of the proof of Lemma 3.4,

$$\begin{aligned}
&\mathbb{P}_\rho\left(\sum_{i=0}^{nT-1} t'_i \geq \beta^{h_n}\right) \\
&\leq n^{-1} + nT\mathbf{P}\left(Tc_\beta \sum_{j=1}^{\tilde{Z}_i-1} \mathbf{1}_{\{h(\mathcal{T}_{ij})\geq h_n\}} \beta^{h(\mathcal{T}'_{ij})} \#\mathcal{T}_{ij} \geq n^{-2}\beta^{h_n}\right) \\
&\leq n^{-1} + nT\mathbf{P}\left(Tc_\beta \tilde{Z}_i \max_{j=1,\ldots,\tilde{Z}_i-1} \#\mathcal{T}_{ij} \geq n^{-2}\beta^{h_n/2}\right) + nT\mathbf{P}\left(\max_{\substack{j=1,\ldots,\tilde{Z}_i-1: \\ h(\mathcal{T}_{ij})\geq h_n}} h(\mathcal{T}'_{ij}) \geq h_n/2\right).
\end{aligned}$$

Clearly the first term decays to zero, and, by applying (3.9), so does the second term. To deal with the third term, observe that, under the convention that $h(\mathcal{T}'_{ij}) = 0$ for j such that

$h(\mathcal{T}_{ij}) < h_n$,

$$\begin{aligned}
\mathbf{P}\left(\max_{\substack{j=1,\ldots,\tilde{Z}_i-1:\\ h(\mathcal{T}_{ij})\geq h_n}} h(\mathcal{T}'_{ij}) \geq h_n/2\right) &= 1 - \mathbf{E}\left(\left(1-\mathbf{P}\left(h(\mathcal{T}'_{ij})\geq h_n/2\right)\right)^{\tilde{Z}-1}\right)\\
&= 1 - f'\left(1-\mathbf{P}\left(h(\mathcal{T}'_{ij})\geq h_n/2\right)\right)\\
&\sim \alpha\mathbf{P}\left(h(\mathcal{T}'_{ij})\geq h_n/2\right)^{\alpha-1} L\left(\mathbf{P}\left(h(\mathcal{T}'_{ij})\geq h_n/2\right)\right)\\
&\sim \alpha q_{h_n}^{\alpha-1}\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)^{\alpha-1}\\
&\qquad L\left(q_{h_n}\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)\right),
\end{aligned}$$

where we have used that $f'(1-x) \sim 1 - \alpha x^{\alpha-1}L(x)$ as $x \to 0^+$, which we first recalled in the proof of Lemma 3.1, and (1.6) again. Since the representation theorem for slowly varying functions ([28, Theorem 1.2], for example) implies that, for any $\varepsilon > 0$,

$$L\left(q_{h_n}\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)\right) \leq \mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)^{-\varepsilon} L(q_{h_n}),$$

for large n, it follows that $\mathbf{P}(\max_{j=1,\ldots,\tilde{Z}_i-1:\, h(\mathcal{T}_{ij})\geq h_n} \max_{j\in B_i} h(\mathcal{T}'_{ij}) \geq h_n/2)$ is asymptotically less than

$$\alpha\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)^{\alpha-1-\varepsilon} q_{h_n}^{\alpha-1}L(q_{h_n}) \sim \frac{\alpha\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right)^{\alpha-1-\varepsilon}}{(\alpha-1)h_n}.$$

Finally, setting $x = h_n/2$ in (3.14) and applying Lemma 3.5 yields

$$\begin{aligned}
\mathbf{P}\left(h(\mathcal{T}'_{ij}) > h_n/2 \middle| h(\mathcal{T}_{ij})\geq h_n\right) &\leq h_n^{\delta}\left(1 - \inf_{h\geq h_n-h_n^{\delta}} c_h f'\left(1-q_{2^{-1}h_n-h_n^{\delta}}\right)\right)\\
&\sim \alpha h_n^{\delta} q_{2^{-1}h_n-h_n^{\delta}}^{\alpha-1} L\left(q_{2^{-1}h_n-h_n^{\delta}}\right)\\
&\sim c h_n^{\delta-1},
\end{aligned}$$

for a suitable choice of constant c, and so, by adjusting c as necessary, we obtain that, for large n,

$$nT\mathbf{P}\left(\max_{\substack{j=1,\ldots,\tilde{Z}_i-1:\\ h(\mathcal{T}_{ij})\geq h_n}} h(\mathcal{T}'_{ij}) \geq h_n/2\right) \leq cnh_n^{-1}h_n^{(\delta-1)(\alpha-1-\varepsilon)}.$$

Since this upper bound converges to 0 for any $\varepsilon < \alpha - 1$, this completes the proof. □

In deriving tail asymptotics for the time X spends in the big leaves emanating from a particular backbone vertex, it will be useful to have information about the set of big leaves that the biased random walk visits deeply before it escapes along the backbone, and the next two lemmas provide this. For their statement, we define the index set of big leaves emanating from ρ_i by

$$B_i := \left\{j = 1,\ldots,\tilde{Z}_i - 1 :\ h(\mathcal{T}_{ij}) \geq h_n\right\}$$

and the subset of those that are visited deeply by X before it escapes a certain distance along the backbone by

$$V_i := \left\{j \in B_i :\ \tau_{x_{ij}} < \tau_{z_i}\right\},$$

where $z_i := \rho_{i+1+h_n^{\delta}}$.

Lemma 3.7. *Let $\alpha \in (1,2]$ and $i \geq 0$. For any $A \subseteq B_i$, we have*

$$P_{\rho_i}^{\mathcal{T}^*}(V_i = A) = \frac{1}{1+\#B_i}\binom{\#B_i}{\#A}^{-1},$$

where $\#B_i$, $\#A$ represents the cardinality of B_i, A, respectively.

Proof. The lemma readily follows from the symmetry of the situation, which implies that, starting from ρ_i, the biased random walk X is equally likely to visit any one of x_{ij}, $j \in B_i$ and z_i first. □

Although the above lemma might seem simple, it allows us to deduce the distributional tail behaviour of the greatest height of a big leaf at a particular backbone vertex visited by the biased random walk X. Note that we continue to use the notation $q_n = \mathbf{P}(Z_n > 0)$.

Lemma 3.8. *Let $\alpha \in (1,2]$ and $i \geq 0$. For $x \geq h_n$,*

$$\mathbb{P}_\rho\left(\max_{j\in V_i} h(\mathcal{T}_{ij}) \geq x\right) = q_x^{\alpha-1} L(q_x).$$

Proof. Let $x \geq h_n$. By definition, we have that

$$\mathbb{P}_\rho\left(\max_{j\in V_i} h(\mathcal{T}_{ij}) < x\right) = \mathbf{E}\left(E_\rho^{\mathcal{T}^*}\left(\prod_{j\in V_i} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}\right)\right),$$

and decomposing the inner expectation over the possible values of V_i yields

$$E_\rho^{\mathcal{T}^*}\left(\prod_{j\in V_i} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}\right) = \sum_{A\subseteq B_i} E_\rho^{\mathcal{T}^*}\left(\mathbf{1}_{\{V_i=A\}}\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}\right).$$

Since $\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}$ is a measurable function of $\mathcal{T}^*$, this can be rewritten as

$$\begin{aligned} E_\rho^{\mathcal{T}^*}\left(\prod_{j\in V_i} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}\right) &= \sum_{A\subseteq B_i} P_\rho^{\mathcal{T}^*}(V_i = A)\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}} \\ &= \sum_{A\subseteq B_i} \frac{1}{\#B_i+1}\binom{\#B_i}{\#A}^{-1}\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}, \end{aligned}$$

where the second equality is an application of Lemma 3.7. Now, since

$$\mathbf{P}\left(\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}}\middle| B_i\right) = \mathbf{P}(h(\mathcal{T}) < x \,|\, h(\mathcal{T}) \geq h_n)^{\#A} = \left(1 - \frac{q_x}{q_{h_n}}\right)^{\#A},$$

for every $A \subseteq B_i$, it follows that

$$\begin{aligned}
\mathbb{P}_\rho\left(\max_{j\in V_i} h(\mathcal{T}_{ij}) < x\right) &= \mathbf{E}\left(\sum_{A\subseteq B_i} \frac{1}{\#B_i+1}\binom{\#B_i}{\#A}^{-1} \mathbf{P}\left(\prod_{j\in A} \mathbf{1}_{\{h(\mathcal{T}_{ij})<x\}} \,\middle|\, B_i\right)\right) \\
&= \mathbf{E}\left(\sum_{A\subseteq B_i} \frac{1}{\#B_i+1}\binom{\#B_i}{\#A}^{-1} \left(1-\frac{q_x}{q_{h_n}}\right)^{\#A}\right) \\
&= \mathbf{E}\left(\frac{1}{\#B_i+1}\sum_{l=0}^{\#B_i}\left(1-\frac{q_x}{q_{h_n}}\right)^{l}\right) \\
&= \mathbf{E}\left(\frac{q_{h_n}}{(\#B_i+1)q_x}\left(1-\left(1-\frac{q_x}{q_{h_n}}\right)^{\#B_i+1}\right)\right).
\end{aligned}$$

To continue, observe that, conditional on $\tilde{Z}_i$, $\#B_i$ is binomially distributed with parameters $\tilde{Z}_i - 1$ and q_{h_n}. Consequently, the probability we are trying to compute is equal to

$$\mathbf{E}\left(\sum_{l=0}^{\tilde{Z}-1}\binom{\tilde{Z}-1}{l} q_{h_n}^l \left(1-q_{h_n}\right)^{\tilde{Z}-1-l} \frac{q_{h_n}}{(l+1)q_x}\left(1-\left(1-\frac{q_x}{q_{h_n}}\right)^{l+1}\right)\right). \tag{3.15}$$

We break this into two terms. Firstly,

$$\begin{aligned}
&\mathbf{E}\left(\sum_{l=0}^{\tilde{Z}-1}\binom{\tilde{Z}-1}{l} q_{h_n}^l \left(1-q_{h_n}\right)^{\tilde{Z}-1-l} \frac{q_{h_n}}{(l+1)q_x}\right) \\
&= q_x^{-1}\mathbf{E}\left(\sum_{l=0}^{Z-1}\binom{Z}{l+1} q_{h_n}^{l+1} \left(1-q_{h_n}\right)^{Z-1-l} \mathbf{1}_{\{Z\geq 1\}}\right) \\
&= q_x^{-1}\mathbf{E}\left(1-\left(1-q_{h_n}\right)^{Z}\right) \\
&= q_x^{-1}\left(1-f\left(1-q_{h_n}\right)\right).
\end{aligned} \tag{3.16}$$

Secondly,

$$\begin{aligned}
&\mathbf{E}\left(\sum_{l=0}^{\tilde{Z}-1}\binom{\tilde{Z}-1}{l} q_{h_n}^l \left(1-q_{h_n}\right)^{\tilde{Z}-1-l} \frac{q_{h_n}}{(l+1)q_x}\left(1-\frac{q_x}{q_{h_n}}\right)^{l+1}\right) \\
&= q_x^{-1}\mathbf{E}\left(\sum_{l=0}^{Z-1}\binom{Z}{l+1} \left(q_{h_n}-q_x\right)^{l+1} \left(1-q_{h_n}\right)^{Z-1-l} \mathbf{1}_{\{Z\geq 1\}}\right) \\
&= q_x^{-1}\left(f\left(1-q_x\right)-f\left(1-q_{h_n}\right)\right).
\end{aligned} \tag{3.17}$$

Since taking the difference between (3.16) and (3.17) gives us (3.15), we have thus proved that

$$\begin{aligned}
\mathbb{P}_\rho\left(\max_{j\in V_i} h(\mathcal{T}_{ij}) < x\right) &= q_x^{-1}\left(1-f\left(1-q_x\right)\right) \\
&= q_x^{-1}\left(q_x - q_x^{\alpha} L\left(q_x\right)\right) \\
&= 1 - q_x^{\alpha-1} L\left(q_x\right),
\end{aligned}$$

where the second equality is a consequence of (1.5), and the lemma follows. □

With these preparations in place, we are now ready to study the asymptotic tail behaviour of

$$t_i := \sum_{j=1}^{\tilde{Z}_i - 1} \mathbf{1}_{\{h(\mathcal{T}_{ij}) \geq h_n\}} \sum_{m=0}^{\tau_{z_i}} \mathbf{1}_{\{X_m \in \mathcal{T}_{ij}(x_{ij})\}},$$

which can be interpreted as the length of time the X spends deep inside leaves emanating from ρ_i before escaping along the backbone. The next lemma gives an upper tail bound for this random variable.

Lemma 3.9. *Let $\alpha \in (1, 2]$ and $\varepsilon > 0$. There exists a constant $c_{\beta,\varepsilon}$ such that, for any $i \geq 0$ and x satisfying $\ln x \geq c_{\beta,\varepsilon} h_n$,*

$$\mathbb{P}_\rho \left(t_i \geq x \right) \leq \frac{(1+\varepsilon) \ln \beta}{(\alpha - 1) \ln x}.$$

Proof. First note that, by applying the commute time identity for random walks (for example, [26, Proposition 10.6]), we have that

$$E_{x_{ij}}^{\{\rho_i\} \cup \mathcal{T}_{ij}} \tau_{\rho_i} + E_{\rho_i}^{\{\rho_i\} \cup \mathcal{T}_{ij}} \tau_{x_{ij}} = \left(\beta^{-i} + \cdots + \beta^{-(i + h_n^\delta)} \right) \left(2\beta^i + \sum_{\substack{x,y \in \mathcal{T}_{ij}: \\ x \sim y}} c(x,y) \right),$$

where $E_\cdot^{\{\rho_i\} \cup \mathcal{T}_{ij}}$ refers to the random walk on the tree $\mathcal{T}_{ij}$ extended by adding the vertex ρ_i and the edge $\{\rho_i, \rho_{ij}\}$. Since $E_{x_{ij}}^{\{\rho_i\} \cup \mathcal{T}_{ij}} \tau_{\rho_i} = E_{x_{ij}}^{\mathcal{T}^*} \tau_{\rho_i}$, it follows that

$$E_{x_{ij}}^{\mathcal{T}^*} \tau_{\rho_i} \leq \frac{\beta}{\beta - 1} \left(2 + 2\beta^{h(\mathcal{T}_{ij})} \# \mathcal{T}_{ij} \right).$$

Thus, since the random walk X spends no time in $\mathcal{T}_{ij}(x_{ij})$ if $j \notin V_i$, we can bound the quenched expectation of t_i conditional on V_i as follows:

$$\begin{aligned} E_\rho^{\mathcal{T}^*} \left(t_i \,|\, V_i \right) &= \sum_{j=1}^{\tilde{Z}_i - 1} \sum_{m=0}^{\infty} P_\rho^{\mathcal{T}^*} \left(X_m \in \mathcal{T}_{ij}(x_{ij}),\, m \leq \tau_{z_i} \,|\, V_i \right) \\ &\leq \sum_{j \in V_i} E_{x_{ij}}^{\mathcal{T}^*} \left(\tau_{\rho_i} \right) E_\rho^{\mathcal{T}^*} \left(\upsilon_{ij} \mid V_i \right), \qquad (3.18) \\ &\leq c_\beta \sum_{j \in V_i} \left(1 + \# \mathcal{T}_{ij} \right) \beta^{h(\mathcal{T}_{ij})} E_\rho^{\mathcal{T}^*} \left(\upsilon_{ij} \mid V_i \right), \qquad (3.19) \end{aligned}$$

where υ_{ij} is the number of passages X makes from ρ_i to x_{ij} before it hits z_i, and the inequality at (3.18) is obtained by an application of the strong Markov property (that holds with respect to the unconditioned law). Now, υ_{ij} is clearly bounded above by the total number of visits to ρ_i, $N(\rho_i)$ say, and, by symmetry, this latter random variable satisfies $E_\rho^{\mathcal{T}^*}(N(\rho_i) \,|\, V_i) = E_\rho^{\mathcal{T}^*}(N(\rho_i) \,|\, \# V_i)$.

Consequently, we deduce that

$$\begin{aligned} E_\rho^{\mathcal{T}^*}(v_{ij} \,|\, V_i) &\leq \sum_{k=0}^{\#B_i} \mathbf{1}_{\{\#V_i=k\}} E_\rho^{\mathcal{T}^*}(N(\rho_i) \,|\, \#V_i = k) \\ &\leq \sum_{k=0}^{\#B_i} \mathbf{1}_{\{\#V_i=k\}} \frac{E_\rho^{\mathcal{T}^*}(N(\rho_i))}{P_\rho^{\mathcal{T}^*}(\#V_i = k)} \\ &\leq c_\beta \sum_{k=0}^{\#B_i} \mathbf{1}_{\{\#V_i=k\}} (\#B_i + 1)\tilde{Z}_i \\ &\leq c_\beta \tilde{Z}_i^2, \end{aligned}$$

where we have applied Lemma 3.7 and the argument at (3.4) to deduce $P_\rho^{\mathcal{T}^*}(\#V_i = k)^{-1} = \#B_i + 1$ and $E_\rho^{\mathcal{T}^*}(N(\rho_i)) \leq c_\beta \tilde{Z}_i$, respectively. Applying the above bound in combination with (3.19) yields

$$E_\rho^{\mathcal{T}^*}(t_i \,|\, V_i) \leq c_\beta \tilde{Z}_i^3 \left(1 + \max_{j=1,\ldots,\tilde{Z}_i - 1} \#\mathcal{T}_{ij}\right) \beta^{\max_{j \in V_i} h(\mathcal{T}_{ij})}.$$

Thus, for $\eta \in (0, \frac{1}{2})$, we can conclude

$$\begin{aligned} \mathbb{P}_\rho(t_i \geq x) &\leq \mathbb{P}_\rho\left(t_i \geq x^\eta E_\rho^{\mathcal{T}^*}(t_i \,|\, V_i)\right) + \mathbb{P}_\rho\left(E_\rho^{\mathcal{T}^*}(t_i \,|\, V_i) \geq x^{1-\eta}\right) \\ &\leq x^{-\eta} + \mathbb{P}_\rho\left(c_\beta \tilde{Z}_i^3 \left(1 + \max_{j=1,\ldots,\tilde{Z}_i - 1} \#\mathcal{T}_{ij}\right) \beta^{\max_{j \in V_i} h(\mathcal{T}_{ij})} \geq x^{1-\eta}\right) \\ &\leq x^{-\eta} + \mathbf{P}\left(c_\beta \tilde{Z}_i^3 \left(1 + \max_{j=1,\ldots,\tilde{Z}_i - 1} \#\mathcal{T}_{ij}\right)^2 \geq x^\eta\right) + \mathbb{P}_\rho\left(\beta^{\max_{j \in V_i} h(\mathcal{T}_{ij})} \geq x^{1-2\eta}\right) \\ &\leq x^{-\eta} + c_\beta x^{-\eta\delta/3} + q^{\alpha-1}_{(1-2\eta)\ln x/\ln\beta} L\left(q_{(1-2\eta)\ln x/\ln\beta}\right) \end{aligned}$$

for $(1-2\eta)\ln x/\ln\beta \geq h_n$, where the value of c_β has been updated from above and the constant δ is the one appearing in (3.9). We have also applied Lemma 3.8 in obtaining the final bound. Finally, (1.6) allows us to deduce from this that, as long as $(1-2\eta)\ln x/\ln\beta$ is sufficiently large, it holds that

$$\mathbb{P}_\rho(t_i \geq x) \leq \frac{(1+\eta)\ln\beta}{(1-2\eta)(\alpha-1)\ln x}.$$

The result follows. □

We can also prove a lower bound for the distributional tail of t_i that matches the upper bound proved above. Similarly to a proof strategy followed in [6], a key step in doing this is obtaining a concentration result to show that the time spent in a leaf visited deeply by the process X will be on the same scale as its expectation.

Lemma 3.10. *Let $\alpha \in (1, 2]$ and $\varepsilon > 0$. There exist constants n_0 and $c_{\beta,\varepsilon}$ such that, for any $i \geq 0$, $n \geq n_0$ and x satisfying $c_{\beta,\varepsilon} h_n \leq \ln x \leq n^2$,*

$$\mathbb{P}_\rho(t_i \geq x) \geq \frac{(1-\varepsilon)\ln\beta}{(\alpha-1)\ln x}.$$

Proof. Our first goal is to derive an estimate on the lower tail of the time that X spends in a big leaf $\mathcal{T}_{ij}$ before hitting ρ_i, given that it starts at the entrance vertex x_{ij}. To this end, we start by noting that under $P_{x_{ij}}^{\mathcal{T}^*}$ and conditional on the number of returns that the random walk X makes to $\mathcal{T}_{ij}(x_{ij})$ before hitting ρ_i, i.e.

$$\upsilon'_{ij} := \#\left\{m \leq \tau_{\rho_i} : \, X_{m-1} = x'_{ij}, \, X_m = x_{ij}\right\},$$

where x'_{ij} denotes the parent of x_{ij}, the random variable $\Sigma := \sum_{m=0}^{\tau_{\rho_i}} \mathbf{1}_{\{X_m \in \mathcal{T}_{ij}(x_{ij})\}}$ is distributed as $\upsilon'_{ij}+1$ independent copies of a random variable whose law is equal to that of $\tau_{x'_{ij}}$ under $P_{x_{ij}}^{\mathcal{T}^*}$. (This is a simple application of the strong Markov property.) In particular, we have that

$$E_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right) = \left(1+\upsilon'_{ij}\right) E_{x_{ij}}^{\mathcal{T}^*} \tau_{x'_{ij}},$$

and also

$$\mathrm{Var}_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right) = \left(1+\upsilon'_{ij}\right) \mathrm{Var}_{x_{ij}}^{\mathcal{T}^*}\left(\tau_{x'_{ij}}\right).$$

To control the right-hand sides of these quantities, we will apply the following moment bounds:

$$E_{x_{ij}}^{\mathcal{T}^*} \tau_{x'_{ij}} = 1 + \beta^{-(i+h_n^\delta)} \sum_{\substack{x,y \in \mathcal{T}_{ij}(x_{ij}): \\ x \sim y}} c(x,y) \geq \beta^{h(\mathcal{T}_{ij})-h_n^\delta}$$

and

$$E_{x_{ij}}^{\mathcal{T}^*}\left(\tau_{x'_{ij}}^2\right) \leq \frac{2\beta}{\beta-1} E_{x_{ij}}^{\mathcal{T}^*}\left(\tau_{x'_{ij}}\right)^2,$$

where the first moment lower bound is obtained by applying a formula similar to (3.2), and the second moment upper bound is an adaptation of a result derived in the proof of [6, Lemma 9.1]. As for the distribution of υ'_{ij} under $P_{x_{ij}}^{\mathcal{T}^*}$, it is clear this is geometric, with parameter given by

$$P_{x'_{ij}}^{\mathcal{T}^*}\left(\tau_{\rho_i} < \tau_{x_{ij}}\right) = \frac{\left(1+\beta^{-1}+\cdots+\beta^{-h_n^\delta+1}\right)^{-1}}{\beta^{h_n^\delta} + \left(1+\beta^{-1}+\cdots+\beta^{-h_n^\delta+1}\right)^{-1}} = \frac{\beta^{h_n^\delta} - \beta^{h_n^\delta-1}}{\beta^{2h_n^\delta} - \beta^{h_n^\delta-1}},$$

from which it follows that

$$E_{x'_{ij}}^{\mathcal{T}^*}\left(\upsilon'_{ij}+1\right) = \frac{\beta^{h_n^\delta}-\beta^{-1}}{1-\beta^{-1}} \geq \beta^{h_n^\delta} - \beta^{-1} \geq \frac{\beta^{h_n^\delta}}{2},$$

for $n \geq n_0$, where n_0 is a deterministic constant. Putting the above observations together, we deduce that, for $n \geq n_0$ and $\varepsilon > 0$,

$$\begin{aligned}
&P_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \leq \frac{\varepsilon}{4}\beta^{h(\mathcal{T}_{ij})}\right) \\
&\quad\leq P_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \leq \frac{\varepsilon}{4}\beta^{h(\mathcal{T}_{ij})}, \, \upsilon'_{ij}+1 \geq \varepsilon E_{x_{ij}}^{\mathcal{T}^*}(\upsilon'_{ij}+1)\right) + P_{x_{ij}}^{\mathcal{T}^*}\left(\upsilon'_{ij}+1 < \varepsilon E_{x_{ij}}^{\mathcal{T}^*}(\upsilon'_{ij}+1)\right) \\
&\quad\leq P_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \leq \frac{1}{2}\left(\upsilon'_{ij}+1\right) E_{x_{ij}}^{\mathcal{T}^*}\tau_{x'_{ij}}\right) + 1 - \left(1 - \frac{\beta^{h_n^\delta}-\beta^{h_n^\delta-1}}{\beta^{2h_n^\delta}-\beta^{h_n^\delta-1}}\right)^{\varepsilon E_{x_{ij}}^{\mathcal{T}^*}(\upsilon'_{ij}+1)} \\
&\quad\leq P_{x_{ij}}^{\mathcal{T}^*}\left(\left|\Sigma - E_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right)\right| \geq \frac{1}{2} E_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right)\right) + \varepsilon \\
&\quad\leq E_{x_{ij}}^{\mathcal{T}^*}\left(\frac{4\mathrm{Var}_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right)}{E_{x_{ij}}^{\mathcal{T}^*}\left(\Sigma \middle| \upsilon'_{ij}\right)^2}\right) + \varepsilon \\
&\quad\leq \frac{8\beta}{\beta-1} E_{x_{ij}}^{\mathcal{T}^*}\left(\frac{1}{\upsilon'_{ij}+1}\right) + \varepsilon \\
&\quad\leq c h_n^\delta \beta^{-h_n^\delta} + \varepsilon,
\end{aligned}$$

where c is a constant depending only on β and n_0 (and not ε).

Now, if we suppose $j_0 \in V_i$ is such that $h(\mathcal{T}_{ij_0}) = \max_{j\in V_i} h(\mathcal{T}_{ij})$, then

$$P_\rho^{\mathcal{T}^*}\left(t_i \le \frac{\varepsilon}{4}\beta^{\max_{j\in V_i} h(\mathcal{T}_{ij})}\,\middle|\, V_i\right) \le P_\rho^{\mathcal{T}^*}\left(\tau_{x_{ij_0}\to\rho_i} \le \frac{\varepsilon}{4}\beta^{h(\mathcal{T}_{ij_0})}\,\middle|\, V_i\right),$$

where $\tau_{x_{ij_0}\to\rho_i}$ is the amount of time X spends in $\mathcal{T}_{ij_0}(x_{ij_0})$ before $\inf\{m \ge \tau_{x_{ij_0}} : X_m = \rho_i\}$. By applying a strong Markov argument for the unconditioned law (cf. (3.18)), yields that the law of $\tau_{x_{ij_0}\to\rho_i}$ under $P_\rho^{\mathcal{T}^*}(\cdot\,|\,V_i)$ is the same as that of Σ (as defined above with $j = j_0$) under $P_{x_{ij_0}}^{\mathcal{T}^*}$, and thus the result of the previous paragraph implies that, for $n \ge n_0$ and $\varepsilon > 0$,

$$P_\rho^{\mathcal{T}^*}\left(t_i \le \frac{\varepsilon}{4}\beta^{\max_{j\in V_i} h(\mathcal{T}_{ij})}\,\middle|\, V_i\right) \le c h_n^\delta \beta^{-h_n^\delta} + \varepsilon.$$

Taking expectations with respect to $P_\rho^{\mathcal{T}^*}$ and $\mathbf{P}$ establishes that the same is true when $P_\rho^{\mathcal{T}^*}(\cdot|V_i)$ is replaced by the annealed law $\mathbb{P}_\rho$. Consequently, for any $n \ge n_0$, $\varepsilon > 0$ and $\ln x \ge h_n \ln\beta$,

$$\begin{aligned}\mathbb{P}_\rho\left(t_i \le x\right) &\le \mathbb{P}_\rho\left(t_i \le \frac{\varepsilon}{4}\beta^{\max_{j\in V_i} h(\mathcal{T}_{ij})}\right) + \mathbb{P}_\rho\left(\max_{j\in V_i} h(\mathcal{T}_{ij}) \le \frac{\ln(4x/\varepsilon)}{\ln\beta}\right)\\ &\le c h_n^\delta \beta^{-h_n^\delta} + \varepsilon + 1 - q_{\frac{\ln(4x/\varepsilon)}{\ln\beta}}^{\alpha-1} L\left(q_{\frac{\ln(4x/\varepsilon)}{\ln\beta}}\right),\end{aligned}$$

where we have applied Lemma 3.8 to deduce the second inequality. Finally, fix $\eta > 0$. If we set $\varepsilon = 1/(\ln x)^2$, then the second term is bounded above by $\eta/\ln x$ for any $\ln x \ge \eta^{-1}$. With this choice of ε, by (1.6), the fourth term is bounded above by $-(1-\eta)\ln(\beta)/(\alpha-1)\ln x$, uniformly over $\ln x \ge x_0$, for suitably large $x_0 = x_0(\eta)$. Moreover, it holds that, $c h_n^\delta \beta^{-h_n^\delta} = o(n^{-2}) = o(1/\ln x)$, uniformly over $\ln x \le n^2$, and this completes the proof. □

Finally for this section, we establish that the same distributional tail behaviour for the random variables

$$\tilde{t}_i := \sum_{j=1}^{\tilde{Z}_i - 1} \mathbf{1}_{\{h(\mathcal{T}_{ij}) \ge h_n\}} \sum_{m=\Delta_i}^{\Delta_{i,(\ln n)^{1+\gamma}}} \mathbf{1}_{\{X_m \in \mathcal{T}_{ij}(x_{ij})\}}, \tag{3.20}$$

where $\Delta_{i,(\ln n)^{1+\gamma}}$ is the first time after Δ_i that the process X hits a backbone vertex outside of the interval $\{\rho_{i-(\ln n)^{1+\gamma}}, \dots, \rho_{i+(\ln n)^{1+\gamma}}\}$. Given the backtracking result of Lemma 3.3, with high probability it is the case that $\tilde{t}_i$ will be identical to the t_i for all relevant indices i. However, the advantage of the sequence $(\tilde{t}_i)$ over (t_i) is that, similarly to the sequence of random variables $(\tilde{T}_x)$ introduced for the directed trap model at (2.5), at least when the traps are suitably well-spaced, it is possible to decouple the elements of $(\tilde{t}_i)$ in such a way as to be able to usefully compare them with an independent sequence.

Lemma 3.11. *Let $\alpha \in (1,2]$ and $\varepsilon > 0$. There exist constants n_0 and $c_{\beta,\varepsilon}$ such that, for any $i \ge 0$, $n \ge n_0$ and x satisfying $c_{\beta,\varepsilon} h_n \le \ln x \le n^2$,*

$$\frac{(1-\varepsilon)\ln\beta}{(\alpha-1)\ln x} \le \mathbb{P}_\rho\left(\tilde{t}_i \ge x\right) \le \frac{(1+\varepsilon)\ln\beta}{(\alpha-1)\ln x}.$$

Proof. If the process X does not hit $\rho_{i-1-(\ln n)^{1+\gamma}}$ again after having hit ρ_i, and does not hit ρ_i again after having hit $\rho_{i+1+(\ln n)^{1+\gamma}}$, then $\tilde{t}_i$ is equal to t_i. Hence,

$$\mathbb{P}_\rho\left(t_i \ne \tilde{t}_i\right) \le 2\mathbf{E}\left(P_{\rho_{1+(\ln n)^{1+\gamma}}}^{\mathcal{T}^*}(\tau_\rho < \infty)\right).$$

An elementary calculation for the biased random walk on a line shows that the right-hand side here is equal to $\beta^{1-(\ln n)^{1+\gamma}} = o(n^{-2})$. Applying this fact, it is easy to deduce the result from Lemmas 3.9 and 3.10. □

3.3 Proof of main result for critical Galton-Watson trees

The purpose of this section is to complete the proof of our main results for biased random walks on critical Galton-Watson trees (Theorem 1.1 and Corollary 1.2).

Proof of Theorem 1.1. We start the proof by claiming that the conclusion of the lemma holds when the hitting time sequence $(\Delta_n)_{n\geq 0}$ is replaced by $(\sum_{i=0}^{n-1}\tilde{t}_i)_{n\geq 0}$. By imitating the proof of Lemma 2.5 with Lemma 3.2 in place of Lemma 2.1, to verify that this is indeed the case, it will be enough to prove the same result for $(\sum_{i=0}^{n-1}\tilde{t}'_i)_{n\geq 0}$, where $(\tilde{t}'_i)_{i\geq 0}$ in an independent sequence such that $\tilde{t}'_i \sim \tilde{t}_{1+(\ln n)^{1+\gamma}}$ for each i. (Note that, because the elements of the sequence $(\tilde{t}_i)_{i\geq 0}$ are only identically-distributed for $i \geq 1+(\ln n)^{1+\gamma}$, we do not take $\tilde{t}'_i \sim \tilde{t}_i$ for each i. By applying the second part of Lemma 3.2, which shows that with high probability there will be no big leaves in the interval close to ρ, it is easy to adapt the argument of Lemma 2.5 to overcome this issue.) Since the tail asymptotics of Lemma 3.11 mean that the relevant functional scaling limit for $(\sum_{i=0}^{n-1}\tilde{t}'_i)_{n\geq 0}$ is an immediate application of Theorem 5.1 (with $h_1(n) = \ln n$ and $h_2(n) = n^{-1}$), our claim holds as desired.

Now, fix $T \in (0,\infty)$. By Lemmas 3.3, 3.4 and 3.6, with probability converging to one we have that, for every $t \in [0,T]$,

$$\sum_{i=0}^{nt-1-(\ln n)^{1+\gamma}} \tilde{t}_i \leq \Delta_n \leq \sum_{i=0}^{nt-1} \tilde{t}_i + 2\beta^{2h_n}.$$

By repeating the proof of Theorem 1.5 exactly with the particular choice $L(x) := \log_+ x$, this, in conjunction with the conclusion of the previous paragraph, yields the result. □

Proof of Corollary 1.2. Since the proof is identical to that of Corollary 1.6, with $\bar{F}(x)$ being taken to be a distribution function that is asymptotically equivalent to $\ln\beta/(\alpha-1)\ln x$, we omit it. □

3.4 Growth rate of quenched mean hitting times

The purpose of this section is to compare the growth rate of $E_\rho^{\mathcal{T}^*}\Delta_n$, that is, the quenched expectations of the hitting times Δ_n, with the growth rate of Δ_n that was established in the previous section. Interestingly, in the result corresponding to Theorem 1.1 (see Theorem 3.14 below), an extra factor of α appears, meaning that the sequence of quenched expectations grows more quickly than the hitting times themselves. This is primarily due to the fact that the quenched expectation $E_\rho^{\mathcal{T}^*}\Delta_n$ feels all the big leaves at a particular backbone vertex, whereas the hitting time Δ_n only feels the big leaves that are deeply visited by X. Indeed, the extra α is most easily understood by comparing the following lemma, which describes the height of the biggest leaf at a particular backbone vertex, with Lemma 3.8, which concerns only deeply visited big leaves.

Lemma 3.12. *Let $\alpha \in (1,2]$. For any $i \geq 0$,*

$$\mathbf{P}\left(\max_{j=1,\ldots,\tilde{Z}_i-1} h(\mathcal{T}_{ij}) \geq x\right) \sim \alpha q_x^{\alpha-1} L\left(q_x\right),$$

as $x \to \infty$.

Proof. Conditioning on $\tilde{Z}_i$, we obtain

$$\begin{aligned}\mathbf{P}\left(\max_{j=1,\ldots,\tilde{Z}_i-1} h(\mathcal{T}_{ij}) < x\right) &= \mathbf{E}\left(\mathbf{P}\left(h(\mathcal{T}) < x\right)^{\tilde{Z}-1}\right)\\ &= \mathbf{E}\left(Z\left(1-q_x\right)^{Z-1}\right)\\ &= f'\left(1-q_x\right),\end{aligned}$$

where we have once again applied the size-biasing of (3.1) to obtain the second equality. Since we know from the proof of Lemma 3.1 that $f'(1-x) \sim 1 - \alpha x^{\alpha-1}L(x)$ as $x \to 0^+$, the proof is complete. □

In studying the quenched expectation of hitting times, we no longer need an argument that is so sophisticated as to consider the time spent in the individual leaves $\mathcal{T}_{ij}$ (which were defined after (3.1)). Instead, we will be concerned only with understanding the expected length of time the biased random walk X spends inside sets of the form $\mathcal{T}_i = \{\rho_i\} \cup (\cup_{j=1,\ldots,\tilde{Z}_i-1}\mathcal{T}_{ij})$. To this end, we introduce a stopping time

$$\sigma_i := \inf\{n \geq 0 : X_n \notin \mathcal{T}_i\}.$$

The expected time spent by X inside $\mathcal{T}_i$ on a single visit is thus given by $\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i$. Similarly to (3.2), we have that

$$\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i = \begin{cases} 1 + \frac{1}{\beta^{i-1}(1+\beta)}\sum_{x,y\in\mathcal{T}_i, x\sim y} c(x,y), & \text{if } i \geq 1,\\ 1 + \sum_{x,y\in\mathcal{T}_0, x\sim y} c(x,y), & \text{if } i = 0,\end{cases} \tag{3.21}$$

and this allows us to obtain the following distributional asymptotics.

Lemma 3.13. *Let $\alpha \in (1,2]$. For any $i \geq 0$,*

$$\mathbf{P}\left(\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i \geq x\right) \sim \frac{\alpha \ln \beta}{(\alpha-1)\ln x},$$

as $x \to \infty$.

Proof. If $i \geq 1$, then from (3.21) we are easily able to deduce that

$$\frac{\beta^{h(\mathcal{T}_i)}}{1+\beta} \leq \mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i \leq 1 + \frac{2\beta^{1+h(\mathcal{T}_i)}\#\mathcal{T}_i}{1+\beta}, \tag{3.22}$$

where $h(\mathcal{T}_i) = \mathbf{1}_{\{\tilde{Z}_i>1\}} + \max_{j=1,\ldots,\tilde{Z}_i-1} h(\mathcal{T}_{ij})$ is the height of $\mathcal{T}_i$. Hence, for any $\eta \in (0,1)$,

$$\begin{aligned}&\mathbf{P}\left(\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i \geq x\right)\\ &\leq \mathbf{P}\left(\beta^{\max_{j=1,\ldots,\tilde{Z}_i-1} h(\mathcal{T}_{ij})} \geq (x-1)^{1-\eta}\right) + \mathbf{P}\left(2\beta^2\#\mathcal{T}_i \geq (\beta+1)(x-1)^{\eta}\right)\\ &\leq \mathbf{P}\left(\max_{j=1,\ldots,\tilde{Z}_i-1} h(\mathcal{T}_{ij}) \geq \frac{(1-\eta)\ln(x-1)}{\ln\beta}\right) + cx^{-\delta\eta},\\ &\sim \frac{\alpha\ln\beta}{(\alpha-1)(1-\eta)\ln x},\end{aligned}$$

where we have applied (3.9) to deduce the second inequality for suitable constants c and $\delta > 0$, and Lemma 3.12 and (1.6) to obtain the asymptotic equivalence. Since (3.22) in conjunction with Lemma 3.12 and (1.6) also implies that

$$\mathbf{P}\left(\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i \geq x\right) \geq \mathbf{P}\left(\max_{j=1,\dots,\tilde{Z}_i-1} h(\mathcal{T}_{ij}) \geq \frac{\ln x + \ln(1+\beta)}{\ln \beta}\right) \sim \frac{\alpha \ln \beta}{(\alpha-1)\ln x},$$

the result follows in this case. The argument for $i = 0$ is similar. □

We are now ready to prove the main result of this section.

Theorem 3.14. *Let $\alpha \in (1, 2]$. As $n \to \infty$, the laws of the processes*

$$\left(\frac{(\alpha-1)\ln_+ E_\rho^{\mathcal{T}^*}\Delta_{nt}}{n\alpha \ln \beta}\right)_{t\geq 0}$$

under $\mathbf{P}$ converge weakly with respect to the Skorohod J_1 topology on $D([0,\infty),\mathbb{R})$ to the law of $(m(t))_{t\geq 0}$.

Proof. The embedded random walk on the backbone Y visits each site ρ_i, $i \geq 1$, a geometric parameter $(\beta-1)/(\beta+1)$ number of times in total and $\rho = \rho_0$ a geometric parameter $(\beta-1)/\beta$ number of times. Moreover, before visiting ρ_n, Y has to visit each element of $\{\rho_0, \dots, \rho_{n-1}\}$ at least once. This and the definition of Δ_n implies that

$$\sum_{i=0}^{n-1} \mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i \leq \mathbf{E}_\rho^{\mathcal{T}^*}\Delta_n \leq \frac{\beta+1}{\beta-1}\sum_{i=0}^{n-1}\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i. \tag{3.23}$$

Now, the random variables $\mathbf{E}_{\rho_i}^{\mathcal{T}^*}\sigma_i$ in these sums are independent and have slowly varying tails, as described by Lemma 3.13. Thus the result is a simple consequence of [22, Theorem 2.1] (or Theorem 5.1 below). □

Remark 3.15. *For comparison, recall the directed trap model of Section 2, but, so as to avoid having to consider the time that the biased random walk X spends at negative integers, replace $\mathbb{Z}$ by the half-line $\mathbb{Z}_+$. As in Theorem 1.5, we have that $(n^{-1}L(\Delta_{nt}))_{t\geq 0}$ converges in distribution under the annealed law $\mathbb{P}_0$ to $(m(t))_{t\geq 0}$. For the corresponding quenched expectation, similarly to (3.23), we have that $\sum_{i=0}^{n-1}\tau_i \leq E_0^\tau(\Delta_n) \leq \frac{\beta+1}{\beta-1}\sum_{i=0}^{n-1}\tau_i$. Thus, again applying [22, Theorem 2.1] (or Theorem 5.1 below), it is possible to check that $(n^{-1}L(E_0^\tau(\Delta_{nt})))_{t\geq 0}$ converges in distribution under $\mathbf{P}$ to $(m(t))_{t\geq 0}$. In particular, in contrast to the critical Galton-Watson tree case, the asymptotic behaviour of $E_0^\tau(\Delta_n)$ and Δ_n are identical. This is because, although certain big leaves will be avoided by certain realisations of the biased random walker in the tree setting, the geometry of the graph $\mathbb{Z}_+$ forces X, when travelling from 0 and n, to visit all the traps in between on every realisation.*

4 Extremal aging

In this section, we will prove Theorem 1.4 and Theorem 1.8, which state that the biased random walk on critical Galton-Watson tree conditioned to survive and the one-dimensional trap model, respectively, experience extremal aging. The phenomenon we describe for these models is similar to what happens in the trapping models considered by Onur Gun in his PhD thesis [20] and to results observed for spin glasses in [7].

4.1 Extremal aging for the one-dimensional trap model

We start by considering the one-dimensional trap model introduced in Section 1.2, with the goal of this section being to prove Theorem 1.8. The reason for proving this result before its counterpart for trees is that the simpler argument it requires will be instructive when it comes to tackling the more challenging tree case in the subsequent section.

Key to proving Theorem 1.8 is establishing that X localises at the closest trap to 0 of a sufficient depth. To describe this precisely, as we do in Lemma 4.2 below, we first introduce the notation

$$l(u) := \min\left\{x \geq 0 : \tau_x \geq \bar{F}^{-1}\left(\frac{1}{u}\right)\right\}.$$

From the independently and identically-distributed nature of the environment, we readily deduce the following preliminary lemma.

Lemma 4.1. *For any $0 < a < b$, we have*

$$\lim_{n\to\infty} \mathbf{P}\left(l(an) = l(bn)\right) = \frac{a}{b}.$$

We now establish the relevant localisation result for X.

Lemma 4.2. *For any $a > 0$, we have*

$$\lim_{n\to\infty} \mathbb{P}_0\left(X_{\bar{F}^{-1}(1/an)} = l(an)\right) = 1.$$

Proof. Our first aim is to show that X hits $l(an)$ before time $\bar{F}^{-1}(1/an)$ with high probability. Clearly, for any $T > 0$, we have that

$$\mathbf{P}\left(l(an) > nT\right) = \mathbf{P}\left(\tau_x < \bar{F}^{-1}\left(\frac{1}{an}\right) : x = 0, 1, \ldots, nT\right) = \left(1 - \frac{1}{an}\right)^{nT+1} \to e^{-T/a},$$

as $n \to \infty$. Moreover, by Lemma 4.1, for $\varepsilon \in (0,1)$ it holds that

$$\mathbf{P}\left(\tau_x \geq \bar{F}^{-1}\left(\frac{1}{an(1-\varepsilon)}\right) \text{ for some } x = 0, \ldots, l(an) - 1\right) = \mathbf{P}\left(l(an) \neq l(an(1-\varepsilon))\right) \to \varepsilon,$$

as $n \to \infty$. Recalling the notation $\tilde{T}_x$ introduced at (2.5) and applying these two results in conjunction with the bound at (2.7) yields

$$\limsup_{n\to\infty} \mathbb{P}_0\left(\Delta_{l(an)} > \sum_{x=1}^{nT} \tilde{T}_x \mathbf{1}_{\{\tau_x \leq \bar{F}^{-1}(1/an(1-\varepsilon))\}} + \bar{F}^{-1}\left(n^{-1}(\ln n)^{1/2}\right)\right) \leq \varepsilon + e^{-T/a}.$$

We know that $\bar{F}^{-1}(n^{-1}(\ln n)^{1/2}) \leq \bar{F}^{-1}(1/\varepsilon n)$ for large enough n, Markov's inequality thus implies

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{P}_0\left(\Delta_{l(an)} > \bar{F}^{-1}\left(\frac{1}{an}\right)\right) \\
&\quad\leq \limsup_{n\to\infty} \frac{1}{\bar{F}^{-1}(1/an)}\left[\mathbb{E}_0\left(\sum_{x=1}^{nT} \tilde{T}_x \mathbf{1}_{\{\tau_x \leq \bar{F}^{-1}(1/an(1-\varepsilon))\}}\right) + \bar{F}^{-1}\left(\frac{1}{\varepsilon n}\right)\right] + \varepsilon + e^{-T/a} \\
&\quad\leq \limsup_{n\to\infty} \frac{1}{\bar{F}^{-1}(1/an)}\left[c_\beta nT\mathbf{E}\left(\tau_0 \mathbf{1}_{\{\tau_0 \leq \bar{F}^{-1}(1/an(1-\varepsilon))\}}\right) + \bar{F}^{-1}\left(\frac{1}{\varepsilon n}\right)\right] + \varepsilon + e^{-T/a},
\end{aligned}$$

where c_β is a constant depending only on β. By proceeding as in the proof of Lemma 2.3 with $g(n)$ replaced by $\bar{F}^{-1}(1/an(1-\varepsilon))$, it is possible to check that

$$\mathbf{E}\left(\tau_0 \mathbf{1}_{\{\tau_0 \leq \bar{F}^{-1}(1/an(1-\varepsilon))\}}\right) \leq \frac{c_1 \bar{F}^{-1}(1/an(1-\varepsilon))}{an(1-\varepsilon)},$$

and so

$$\limsup_{n\to\infty} \mathbb{P}_0\left(\Delta_{l(an)} > \bar{F}^{-1}\left(\frac{1}{an}\right)\right) \leq \limsup_{n\to\infty} \frac{c_2 T \bar{F}^{-1}(1/an(1-\varepsilon)) + a\bar{F}^{-1}(1/\varepsilon n)}{a(1-\varepsilon)\bar{F}^{-1}(1/an)} + \varepsilon + e^{-T/a}.$$

Were $\limsup_{n\to\infty} \bar{F}^{-1}(1/an(1-\varepsilon))/\bar{F}^{-1}(1/an) > 0$, then there would exist a subsequence $(n_i)_{i\geq 0}$ and constant $c > 0$ such that $\bar{F}^{-1}(1/an_i(1-\varepsilon)) \geq c\bar{F}^{-1}(1/an_i)$. Applying the decreasing function $\bar{F}$ to both sides and then the slowly-varying property (1.1) yields that $a \leq a(1-\varepsilon)$, which is clearly a contradiction. Hence $\lim_{n\to\infty} \bar{F}^{-1}(1/an(1-\varepsilon))/\bar{F}^{-1}(1/an) = 0$. Similarly, one has that $\lim_{n\to\infty} \bar{F}^{-1}(1/\varepsilon n)/\bar{F}^{-1}(1/an) = 0$ for any $\varepsilon < a$. Thus, letting $T \to \infty$ and $\varepsilon \to 0$, the above estimate yields

$$\lim_{n\to\infty} \mathbb{P}_0\left(\Delta_{l(an)} > \bar{F}^{-1}\left(\frac{1}{an}\right)\right) = 0,$$

as desired.

Now, if $X_{\bar{F}^{-1}(1/an)} \neq l(an)$, then either X does not hit $l(an)$ before time $\bar{F}^{-1}(1/an)$, or it does hit $l(an)$ and spends less time than $\bar{F}^{-1}(1/an)$ there before moving to any other vertex. By the conclusion of the previous paragraph, the former event has probability 0 asymptotically, and so

$$\limsup_{n\to\infty} \mathbb{P}_0\left(X_{\bar{F}^{-1}(1/an)} \neq l(an)\right) \leq \limsup_{n\to\infty} \mathbf{E}\left(P^\tau_{l(an)}\left(\inf\{t : X_t \neq l(an)\} \leq \bar{F}^{-1}(1/an)\right)\right).$$

Since $\inf\{t : X_t \neq l(an)\}$ is exponential with mean $\tau_{l(an)}$ under $P^\tau_{l(an)}$, for any $\varepsilon > 0$ the right-hand side here is bounded above by

$$\limsup_{n\to\infty} \mathbf{E}\left(1 \wedge \frac{\bar{F}^{-1}(1/an)}{\tau_{l(an)}}\right) \leq \limsup_{n\to\infty} \left[\mathbf{P}\left(\tau_{l(an)} < \bar{F}^{-1}(1/an(1+\varepsilon))\right) + \frac{\bar{F}^{-1}(1/an)}{\bar{F}^{-1}(1/an(1+\varepsilon))}\right].$$

The probability in the previous expression is equal to $\mathbf{P}(l(an) \neq l(an(1+\varepsilon)))$, and, by Lemma 4.1, this is asymptotically bounded above by ε. Similarly to an observation made in the previous paragraph, we also have that $\lim_{n\to\infty} \bar{F}^{-1}(1/an)/\bar{F}^{-1}(1/an(1+\varepsilon)) = 0$, and thus we have established

$$\limsup_{n\to\infty} \mathbb{P}_0\left(X_{\bar{F}^{-1}(1/an)} \neq l(an)\right) \leq \varepsilon.$$

Since ε was arbitrary, this completes the proof. □

Combining Lemmas 4.1 and 4.2, we readily obtain Theorem 1.8.

4.2 Extremal aging for the critical Galton-Watson tree model

We now return to the setting of Section 1.1, so as to prove Theorem 1.4. Similarly to the strategy of the previous section, we will show that the biased random walk on a critical Galton-Watson tree localises in the first suitably big leaf it visits deeply. To describe this, we introduce the notation:

$$l(x) := \min\left\{i \geq 0 : \max_{j\in V_i} h(\mathcal{T}_{ij}) \geq x/\ln\beta\right\}.$$

Whilst the form of the following lemma is similar that of Lemma 4.1, we note that its proof is more involved. This is because, unlike the holding time means τ_x used to define l there, the random variables $\max_{j\in V_i} h(\mathcal{T}_{ij})$ are not environment measurable or independent.

Lemma 4.3. *Let $\alpha \in (1,2]$. For any $0 < a < b$, we have*

$$\lim_{n\to\infty} \mathbb{P}_\rho\left(l(an) = l(bn)\right) = \frac{a}{b}.$$

Proof. First, define

$$\tilde{V}_i := \left\{ j \in B_i :\ \tau_{x_{ij}} < \Delta_{i,(\ln n)^{(1+\gamma)}} \right\},$$

to be the set of big leaves visited by X before the stopping time $\Delta_{i,(\ln n)^{(1+\gamma)}}$ that was introduced at (3.20). Set $\tilde{H}_i := \max_{j\in\tilde{V}_i} h(\mathcal{T}_{ij})$ if $\tilde{V}_i \neq \emptyset$, and $\tilde{H}_i = 0$ otherwise; observe that if $\tilde{H}_i > 0$, then it is necessarily also the case that $\tilde{H}_i \geq h_n$. Moreover, for $T \in (0,\infty)$ and $\varepsilon \in (0,1)$, let

$$\mathcal{E}_1(n) := \left\{ \sum_{i=m}^{m+n^\varepsilon} \mathbf{1}_{\{N_n(i)\geq 1\}} \leq 1 :\ m = 0,1,\ldots,Tn - n^\varepsilon \right\} \cup \left\{ \sum_{i=0}^{n^\varepsilon} \mathbf{1}_{\{N_n(i)\geq 1\}} = 0 \right\}.$$

By proceeding as in the proof of Lemma 2.5, it is possible to show that, under $\mathbb{P}_\rho$, the random variables $(\tilde{H}_i, N_n(i))_{i=0}^{nT}$ conditional on $\mathcal{E}_1(n)$ have the same joint distribution as $(\tilde{H}'_i, N'_n(i))_{i=0}^{nT}$ conditional on $\mathcal{E}'_1(n)$, where $(\tilde{H}'_i, N'_n(i))_{i\geq 0}$ are independent copies of the pair of random variables $(\tilde{H}_{1+(\ln n)^{1+\gamma}}, N_n(1+(\ln n)^{1+\gamma}))$ and $\mathcal{E}'_1(n)$ is defined analogously to $\mathcal{E}_1(n)$ with the $N_n(i)$s replaced by $N'_n(i)$s. Consequently, if we set

$$\tilde{l}(x) := \min\left\{ i \geq 0 :\ \tilde{H}_i \geq x/\ln\beta \right\},$$

and define $\tilde{l}'(x)$ similarly from the random variables $\tilde{H}'_i$, then

$$\begin{aligned}
&\left|\mathbb{P}_\rho\left(\tilde{l}(an) = \tilde{l}(bn)\right) - \mathbb{P}_\rho\left(\tilde{l}'(an) = \tilde{l}'(bn)\right)\right| \\
&\quad\leq \left|\mathbb{P}_\rho\left(\tilde{l}(an) = \tilde{l}(bn),\ \tilde{l}(an) \leq nT,\ \mathcal{E}_1(n)\right) - \mathbb{P}_\rho\left(\tilde{l}'(an) = \tilde{l}'(bn),\ \tilde{l}'(an) \leq nT,\ \mathcal{E}'_1(n)\right)\right| \\
&\qquad + \left|\mathbb{P}_\rho\left(\tilde{l}(an) > nT,\ \mathcal{E}_1(n)\right) - \mathbb{P}_\rho\left(\tilde{l}'(an) > nT,\ \mathcal{E}'_1(n)\right)\right| \\
&\qquad + 2\mathbb{P}_\rho\left(\tilde{l}'(an) > nT,\ \mathcal{E}'_1(n)\right) + 2\mathbb{P}_\rho(\mathcal{E}_1(n)^c) \\
&\quad\leq 2\mathbb{P}_\rho\left(\tilde{l}'(an) > nT\right) + 2\mathbf{P}(\mathcal{E}_1(n)^c),
\end{aligned} \tag{4.1}$$

where we have applied the fact that $\{\tilde{l}(an) = \tilde{l}(bn),\ \tilde{l}(an) \leq nT\}$ and $\{\tilde{l}(an) > nT\}$ are both $(\tilde{H}_i)_{i=0}^{nT}$ measurable events. Now, similarly to the observation made in the proof of Lemma 3.11, if the process X does not hit $\rho_{i-1-(\ln n)^{1+\gamma}}$ again after having hit ρ_i, and does not hit ρ_i again after having hit $\rho_{i+1+(\ln n)^{1+\gamma}}$ – an event which has probability greater than $1-o(n^{-2})$ uniformly in i, then $\tilde{H}_i$ is equal to $\max_{j\in V_i} h(\mathcal{T}_{ij})$. Hence, applying (1.6) and Lemma 3.8, we obtain that, for any $x, \varepsilon > 0$,

$$\left|\mathbb{P}_\rho\left(\tilde{H}_i \geq xn\right) - \frac{1}{(\alpha-1)xn}\right| \leq \frac{\varepsilon}{n}$$

for large n (uniformly in i), and clearly the same bound holds when $\tilde{H}_i$ is replaced by $\tilde{H}'_i$. Applying the independence of the random variables $(\tilde{H}'_i)_{i\geq 0}$, it follows that

$$\lim_{n\to\infty} \mathbb{P}_\rho\left(\tilde{l}'(an) = \tilde{l}'(bn)\right) = \frac{a}{b},$$

and also

$$\begin{aligned}\mathbb{P}_\rho\left(\tilde{l}'(an) > nT\right) &= \mathbb{P}_\rho\left(\tilde{H}_{1+(\ln n)^{1+\gamma}} < \frac{an}{\ln\beta}\right)^{nT+1} \\ &\leq \left(1 - \frac{\ln\beta - \varepsilon a}{(\alpha-1)an}\right)^{nT+1} \\ &\sim e^{-T(\ln\beta-\varepsilon a)/(\alpha-1)a}.\end{aligned}$$

Combining these results with Lemma 3.2, which implies that $\mathbf{P}(\mathcal{E}_1(n)^c) \to 0$, and the estimate at (4.1), then letting $T \to \infty$, yields

$$\lim_{n\to\infty} \mathbb{P}_\rho\left(\tilde{l}(an) = \tilde{l}(bn)\right) = \frac{a}{b}. \tag{4.2}$$

Now, suppose that $\mathcal{E}_2(n)$ is the event that the embedded random walk on the backbone Y does not backtrack more that $(\ln n)^{1+\gamma}$ before hitting $\rho_{n(T+1)}$ – by Lemma 3.3, $\mathbb{P}_\rho(\mathcal{E}_2(n)) \to 1$. Moreover, on the event $\mathcal{E}_2(n)$, we have that $\tilde{H}_i = \max_{j\in V_i} h(\mathcal{T}_{ij})$ for $i \leq n(T+1) - 1 - h_n^\delta$. In particular, for large enough n, if $\mathcal{E}_2(n)$ holds and also $\tilde{l}(an) \leq nT$, then it must be the case that $l(an) = \tilde{l}(an)$. Hence, for large n,

$$\mathbb{P}_\rho\left(l(an) \neq \tilde{l}(an)\right) \leq \mathbb{P}_\rho\left(\tilde{l}(an) > nT\right) + \mathbb{P}_\rho\left(\mathcal{E}_2(n)^c\right).$$

Similarly to above, we have that the first term here can be bounded above by

$$\left|\mathbb{P}_\rho\left(\tilde{l}(an) > nT,\, \mathcal{E}_1(n)\right) - \mathbb{P}_\rho\left(\tilde{l}'(an) > nT,\, \mathcal{E}_1'(n)\right)\right| + \mathbb{P}_\rho\left(\tilde{l}'(an) > nT\right) + \mathbf{P}(\mathcal{E}_1(n)^c),$$

the limsup as $n \to \infty$ of which can be made arbitrarily small by choosing T suitably large. Hence

$$\lim_{n\to\infty} \mathbb{P}_\rho\left(l(an) \neq \tilde{l}(an)\right) = 0.$$

The lemma follows by applying this in conjunction with (4.2). □

Before proceeding to prove the analogue of Lemma 4.2 in the tree setting – see Lemma 4.5 below, we prove a preliminary estimate which rules out the possibility that any leaves have heights that are close to any particular level on the appropriate scale.

Lemma 4.4. *Let $\alpha \in (1,2]$. For any $a, T \in (0,\infty)$,*

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \mathbf{P}\left(\min_{i=0,1,\ldots,nT} \min_{j=1,\ldots,\tilde{Z}_i-1} |h(\mathcal{T}_{ij}) - an| \leq an\varepsilon\right) = 0.$$

Proof. First observe that

$$\begin{aligned}\mathbf{P}\left(\min_{j=1,\ldots,\tilde{Z}_i-1} |h(\mathcal{T}_{ij}) - an| \leq an\varepsilon\right) &= 1 - \mathbf{E}\left(\left(1 - q_{an(1-\varepsilon)} + q_{an(1+\varepsilon)}\right)^{\tilde{Z}-1}\right) \\ &= 1 - f'\left(1 - q_{an(1-\varepsilon)} + q_{an(1+\varepsilon)}\right) \\ &\sim \alpha\left(q_{an(1-\varepsilon)} - q_{an(1+\varepsilon)}\right)^{\alpha-1} L\left(q_{an(1-\varepsilon)} - q_{an(1+\varepsilon)}\right),\end{aligned}$$

where we again apply [29, (2.1)] to deduce the asymptotic equality. Now, from (1.6), one can check that

$$q_{an(1-\varepsilon)} - q_{an(1+\varepsilon)} \sim \frac{2\varepsilon q_{an}}{\alpha - 1},$$

which yields

$$\mathbf{P}\left(\min_{j=1,\dots,\tilde{Z}_i-1}|h(\mathcal{T}_{ij})-an|\le an\varepsilon\right)\sim\alpha\left(\frac{2\varepsilon q_{an}}{\alpha-1}\right)^{\alpha-1}L\left(\frac{2\varepsilon q_{an}}{\alpha-1}\right)\sim\frac{c_\alpha\varepsilon^{\alpha-1}}{an},$$

where c_α is a constant depending only on α. The lemma readily follows. □

Lemma 4.5. *Let $\alpha\in(1,2]$. For any $a>0$, we have*

$$\lim_{n\to\infty}\mathbb{P}_\rho\left(\pi(X_{e^{an}})=\rho_{l(an)}\right)=1.$$

Proof. Fix $\varepsilon>0$, and let i_0,j_0 be indices such that $x_{i_0j_0}$ is the first entrance to a big leaf with height greater than or equal to $an(1+\varepsilon)/\ln\beta$ visited by X (the relevant terminology was introduced just above (3.10)). Note that this implies $i_0=l(an(1+\varepsilon))$. Moreover, if $\{l(an(1+\varepsilon))\le nT\}$ holds and n is suitably large, then $\tau_{x_{i_0j_0}}\le\Delta_{n(T+1)}$. In particular, for large n,

$$\begin{aligned}\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}>\Delta_{n(T+1)}\right)&\le\mathbb{P}_\rho\left(l(an(1+\varepsilon))>nT\right)\\&\le\mathbb{P}_\rho\left(l(an(1+\varepsilon))\ne\tilde{l}(an(1+\varepsilon))\right)+\mathbb{P}_\rho\left(\tilde{l}(an(1+\varepsilon))>nT\right),\end{aligned}$$

where $\tilde{l}(an)$ was defined in the proof of Lemma 4.3, and the upper bound here converges to 0 as n and then T tend to infinity. Consequently, by applying Lemmas 3.3, 3.4, 3.6 similarly to the proof of Theorem 1.1, as well as Lemma 4.4, we obtain that if

$$\Theta:=\sum_{i=0}^{n(T+1)-1}\sum_{j=1}^{\tilde{Z}_i-1}\mathbf{1}_{\{h_n\le h(\mathcal{T}_{ij})\le an(1-\varepsilon)/\ln\beta\}}\sum_{m=\Delta_i}^{\Delta_{i,(\ln n)^{1+\gamma}}}\mathbf{1}_{\{X_m\in\mathcal{T}_{ij}(x_{ij})\}},$$

then

$$\lim_{\varepsilon\to0}\limsup_{T\to\infty}\limsup_{n\to\infty}\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}>\Theta+2\beta^{2h_n}\right)=0.$$

Now, by proceeding similarly to the proof of Lemma 3.4, we have that

$$\begin{aligned}E_\rho^{\mathcal{T}^*}\Theta&\le\frac{\beta}{\beta-1}\sum_{i=0}^{n(T+1)-1}\tilde{Z}_i\left(1+2e^{an(1-\varepsilon)}\#\mathcal{T}_{ij}\right)\\&\le\frac{2\beta n(T+1)e^{an(1-\varepsilon)}}{\beta-1}\max_{i=0,\dots,n(T+1)-1}\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right].\end{aligned}$$

Combining these observations yields

$$\begin{aligned}&\limsup_{\varepsilon\to0}\limsup_{n\to\infty}\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}>e^{an}\right)\\&\le\limsup_{\varepsilon\to0}\limsup_{T\to\infty}\limsup_{n\to\infty}\mathbb{P}_\rho\left(\Theta>2^{-1}e^{an}\right)\\&\le\limsup_{\varepsilon\to0}\limsup_{T\to\infty}\limsup_{n\to\infty}\left[\mathbb{P}_\rho\left(\Theta>nE_\rho^{\mathcal{T}^*}\Theta\right)+\mathbb{P}_\rho\left(E_\rho^{\mathcal{T}^*}\Theta>2^{-1}n^{-1}e^{an}\right)\right]\\&\le\limsup_{\varepsilon\to0}\limsup_{T\to\infty}\limsup_{n\to\infty}\mathbf{P}\left(\max_{i=0,\dots,n(T+1)-1}\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]>\frac{(\beta-1)e^{an\varepsilon}}{4\beta n^2(T+1)}\right)\\&\le\limsup_{\varepsilon\to0}\limsup_{T\to\infty}\limsup_{n\to\infty}n(T+1)\mathbf{P}\left(\tilde{Z}_i\left[1+\max_{j=1,\dots,\tilde{Z}_i-1}\#\mathcal{T}_{ij}\right]>\frac{(\beta-1)e^{an\varepsilon}}{2\beta n^2(T+1)}\right).\end{aligned}$$

Applying (3.9), we thus obtain

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}>e^{an}\right)=0.$$

The conclusion of the previous paragraph implies that

$$\begin{aligned}
&\limsup_{n\to\infty}\mathbb{P}_\rho\left(\pi(X_{e^{an}})\neq\rho_{l(an)}\right)\\
&\quad\le\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\left[\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}>e^{an}\right)+\mathbb{P}_\rho\left(\tau_{x_{i_0j_0}}\le e^{an},\ \pi(X_{e^{an}})\neq\rho_{l(an)}\right)\right]\\
&\quad\le\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbf{E}\left(P^{\mathcal{T}^*}_{x_{i_0j_0}}\left(\inf\{m:\ \pi(X_m)\neq\rho_{l(an)}\}\le e^{an}\right)\right)\\
&\quad\le\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbf{E}\left(P^{\mathcal{T}^*}_{x_{i_0j_0}}\left(\tau_{\rho_{l(an)}}\le e^{an}\right)\right).
\end{aligned}$$

It is plain to show that

$$P^{\mathcal{T}^*}_{x_{i_0j_0}}\left(\tau_{y_{i_0j_0}}>\tau_{\rho_{l(an)}}\right)\le c_1\beta^{-h^\delta_n},$$

for some constant c_1 depending only on β; indeed, this is nothing more than a computation for a biased random walk on $\mathbb{Z}$. Furthermore, another simple calculation for biased random walk on the line yields

$$P^{\mathcal{T}^*}_{y_{i_0j_0}}\left(\tau^+_{y_{i_0j_0}}>\tau_{\rho_{l(an)}}\right)\le c_2e^{-a(1+\varepsilon)n},$$

where $\tau^+_{y_{i_0j_0}}$ is the time of the first return to $y_{i_0j_0}$, so

$$P^{\mathcal{T}^*}_{y_{i_0j_0}}\left(\tau_{\rho_{l(an)}}\le e^{an}\right)\le\sum_{k=0}^{e^{an}/2}P^{\mathcal{T}^*}_{y_{i_0j_0}}\left(\tau^+_{y_{i_0j_0}}\le\tau_{\rho_{l(an)}}\right)^kP^{\mathcal{T}^*}_{y_{i_0j_0}}\left(\tau^+_{y_{i_0j_0}}>\tau_{\rho_{l(an)}}\right)\le c_3e^{-a\varepsilon n}.$$

Consequently,

$$\limsup_{n\to\infty}\mathbb{P}_\rho\left(\pi(X_{e^{an}})\neq\rho_{l(an)}\right)\le\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\left(c_1\beta^{-h^\delta_n}+c_3e^{-a\varepsilon n}\right)=0,$$

which completes the proof. □

Putting Lemma 4.3 and 4.5 together, we obtain Theorem 1.4.

5 A limit theorem for sums of independent random variables with slowly varying tail probability

In this section, we derive the limit theorem for sums of independent random variables with slowly varying tail probability that was applied in the proofs of Lemma 2.5 and Theorem 1.1. The result we prove here is a generalisation of [22, Theorem 2.1].

Let $(X_{i,j})_{i,j\in\mathbb{N}}$ be non-negative random variables such that for each $n\ge 1$, the elements of the collection $(X_{n,j})_{j\in\mathbb{N}}$ are independent and have common distribution function F_n. Moreover, suppose F is a distribution function such that $\bar{F}(x):=1-F(x)$ is slowly varying and $\bar{F}(x)>0$ for all $x>0$. Similarly writing $\bar{F}_n(x):=1-F_n(x)$, the main assumption of this section is that for each $\varepsilon>0$, there exist constants c_1,c_2 such that

$$(1-\varepsilon)\bar{F}_n(x)\le\bar{F}(x)\le(1+\varepsilon)\bar{F}_n(x),\qquad\forall x\in[c_1(g_1(n)\vee 1),c_2g_2(n)],\tag{5.1}$$

where $g_i(n) := \bar{F}^{-1}(n^{-1}h_i(n))$, $i = 1, 2$, with $h_1 : \mathbb{N} \to (0, \infty)$ a non-decreasing, divergent function satisfying $\lim_{n\to\infty} h_1(n)/n = 0$, and $h_2 : \mathbb{N} \to [0, \infty)$ a non-increasing function satisfying $\lim_{n\to\infty} h_2(n) = 0$. (Note that necessarily $\lim_{n\to\infty} g_i(n) = \infty$ for $i = 1, 2$.) Defining a function L by setting $L(x) := 1/\bar{F}(x)$, we then have the following scaling result for sums of the form

$$S_m^n := \sum_{j=1}^{m} X_{n,j}.$$

Theorem 5.1. *Assume that (5.1) holds. As $n \to \infty$,*

$$\left(\frac{1}{n} L\left(S_{nt}^n\right)\right)_{t\geq 0} \to (m(t))_{t\geq 0} \tag{5.2}$$

in distribution with respect to the Skorohod J_1 topology on $D([0, \infty), \mathbb{R})$.

Remark 5.2. *(i) Note that, similarly to Remark 1.9, if $\bar{F}_n$ and $\bar{F}$ are not continuous and eventually strictly decreasing, a minor modification to the proof of the above result (cf. Remark 1.7(ii)) is needed.*
(ii) The same conclusion holds if on the left-hand side of (5.2) we replace L by $L_n(x) = 1/\bar{F}_n(x)$.

To the end of proving the above result, it is helpful to introduce $(\eta(t))_{t\geq 0}$ to represent a one-sided stable process with index 1/2, i.e., with Lévy measure given by $\mu((x, \infty)) = x^{-1/2}$ for $x > 0$. We will write $F_*(x) = P(\eta(1) \leq x)$ for the distribution function of $\eta(1)$. Briefly, the connection with $(m(t))_{t\geq 0}$ is that $m(t) = (\max_{0\leq s\leq t} \Delta\eta(s))^{1/2}$ (as processes) – here we recall $\Delta\eta(s) = \eta(s) - \eta(s^-)$. Moreover, if we set

$$\begin{aligned} \eta_{n,i} &= \eta(i/n) - \eta((i-1)/n), \qquad \forall i \geq 1, \\ m_n(t) &= \left(\max_{i\leq nt} \eta_{n,i}\right)^{1/2} \mathbf{1}_{[1/n,\infty)}(t), \end{aligned}$$

then $m_n \to m$ almost-surely in the Skorohod J_1 topology. Indeed, since $\eta_n(t) := \eta(\lfloor nt \rfloor/n) \to \eta(t)$ in the Skorohod J_1 topology, by the continuous mapping theorem

$$\max_{i\leq nt} \eta_{n,i} = \max_{0\leq s\leq t} \Delta\eta_n(s) \to \max_{0\leq s\leq t} \Delta\eta(s) \tag{5.3}$$

almost-surely as a process in the same topology.

We are now ready to present the key lemma needed to establish Theorem 5.1. (This corresponds to [22, Lemmas 2.2 and 2.3].) In its statement, we use the notation $\phi_n(x) := F_n^{-1}(F_*(x))$, and we also define $\phi(x) := F^{-1}(F_*(x))$ for its proof.

Lemma 5.3. *Under (5.1), we have the following.*
(i) For every $\lambda > 0$ and $T > 0$, as $n \to \infty$,

$$\sup_{0\leq x\leq T} \left|\frac{1}{n} L\left(\lambda\phi_n\left(n^2x^2\right)\right) - x\right| \to 0. \tag{5.4}$$

(ii) For each $\delta > 0$, $T > \delta$, there exist random constants $K_1, K_2 > 0$ and n_0 such that, for every $t \in [\delta, T]$ and $n \geq n_0$,

$$K_1\phi_n\left(n^2 m_n(t)^2\right) \leq \sum_{i\leq nt} \phi_n\left(n^2\eta_{n,i}\right) \leq K_2\phi_n\left(n^2 m_n(t)^2\right), \tag{5.5}$$

almost-surely.

Proof. We first give some preliminary computations. Rewriting (5.1), we have

$$\frac{F(x)-\varepsilon}{1-\varepsilon} \le F_n(x) \le \frac{F(x)+\varepsilon}{1+\varepsilon}, \qquad \forall x \in [c_1(g_1(n) \vee 1), c_2 g_2(n)].$$

Setting

$$z = F_*^{-1}(F(x)), \qquad \kappa_1(\varepsilon, z) = F_*^{-1}\left(\frac{F_*(z)-\varepsilon}{1-\varepsilon}\right), \qquad \kappa_2(\varepsilon, z) = F_*^{-1}\left(\frac{F_*(z)+\varepsilon}{1+\varepsilon}\right),$$

it follows that

$$\phi_n(\kappa_1(\varepsilon, z)) \le \phi(z) = x \le \phi_n(\kappa_2(\varepsilon, z)), \tag{5.6}$$

for $F(c_1 g_1(n)) < F_*(z) = F(x) < F(c_2 g_2(n))$. Since

$$z^{-1/2} \sim 1 - F_*(z) < 1 - F(c_1 g_1(n)) = \bar{F}(c_1 g_1(n)) \sim \bar{F}(g_1(n)) = n^{-1} h_1(n),$$

and similarly $z^{-1/2} \sim 1 - F_*(z) > n^{-1} h_2(n)$, for suitably large n, the inequality at (5.6) holds for all $c_1' n^2 / h_1(n)^2 < z < c_2' n^2 / h_2(n)^2$. Now, since $F_*^{-1}(x) \sim (1-x)^{-2}$ for $x \to 1^-$, we have

$$\begin{aligned}
\kappa_1(\varepsilon, n^2 x^2) &= F_*^{-1}\left(\frac{F_*(n^2x^2)-\varepsilon}{1-\varepsilon}\right) \\
&\sim \left(1 - \frac{F_*(n^2x^2)-\varepsilon}{1-\varepsilon}\right)^{-2} \\
&= \frac{(1-\varepsilon)^2}{(1-F_*(n^2x^2))^2} \\
&\sim (1-\varepsilon)^2 n^2 x^2,
\end{aligned}$$

so that $\kappa_1(\varepsilon, n^2x^2) \ge (1-\varepsilon)^3 n^2 x^2$ for large n. Similarly, $\kappa_2(\varepsilon, n^2x^2) \le (1+\varepsilon)^3 n^2 x^2$ for large n. Since ϕ_n, h_1 are non-decreasing and h_2 is non-increasing, by (5.6) we conclude

$$\phi_n\left((1-\varepsilon)^3 n^2 x^2\right) \le \phi\left(n^2 x^2\right) \le \phi_n\left((1+\varepsilon)^3 n^2 x^2\right), \qquad \forall n > h_1^{-1}(c_1''/x) \vee h_2^{-1}(c_2''/x). \tag{5.7}$$

Now let us prove (i). By the definition of ϕ_n and the fact that $1 - F_*(x) \sim x^{-1/2}$, we have, as in the proof of [22, Lemma 2.2], that $\lim_{x\to\infty} L(\lambda\phi(x^2))/x = 1$, from which it follows that $\lim_{n\to\infty} L(\lambda\phi(n^2x^2))/n = x$ for $\lambda > 0$. Noting that $L(\lambda\phi(n^2x^2))/n$ is monotone in x and the limiting function is continuous, this convergence is uniform in x on each finite interval. By (5.7),

$$n^{-1} L\left(\lambda\phi_n\left((1-\varepsilon)^3 n^2 x^2\right)\right) \le n^{-1} L\left(\lambda\phi(n^2x^2)\right) \le n^{-1} L\left(\lambda\phi_n\left((1+\varepsilon)^3 n^2 x^2\right)\right),$$

for $n > h_1^{-1}(c_1''/x) \vee h_2^{-1}(c_2''/x)$, and thus we obtain (5.4).

We next prove (ii). First, by taking $K_1 = 1$ the lower bound of (5.5) is clear. So we will prove the upper bound. As in the proof of [22, Lemma 2.3], noting that ϕ^{-1} is slowly varying and using the representation theorem of [28, Theorem 1.2], we have $\phi^{-1}(x) = c(x)\exp(\int_1^x \varepsilon(t)/t dt)$ where $c(x) \to c > 0$ and $\varepsilon(x) \to 0$ as $x \to \infty$. Thus, $\phi(x)$ may be expressed as

$$\phi(x) = \exp\left(\int_1^{q(x)x} \frac{\tilde{\varepsilon}(t)}{t} dt\right)$$

where $\tilde{\varepsilon}(x) \to \infty$ and $q(x) \to 1/c$ as $x \to \infty$. Using this and (5.7), we have, for all $a > 0$ and $M > 2$,

$$\begin{aligned} n^2 \frac{\phi_n(n^2 a)}{\phi_n(2n^2 a)} &\leq n^2 \frac{\phi\left(\frac{n^2 a}{(1-\varepsilon)^3}\right)}{\phi_n\left(\frac{2n^2 a}{(1+\varepsilon)^3}\right)} \\ &= n^2 \exp\left(- \int_{q\left(\frac{n^2 a}{(1-\varepsilon)^3}\right)\frac{n^2 a}{(1-\varepsilon)^3}}^{q\left(\frac{2n^2 a}{(1+\varepsilon)^3}\right)\frac{2n^2 a}{(1+\varepsilon)^3}} \frac{\tilde{\varepsilon}(t)}{t} dt \right) \\ &\leq n^2 \exp(-M \log(an^2)) \\ &= a^{-M} n^{-2(M-1)}, \end{aligned}$$

when $n > h_1^{-1}(c_1'''/\sqrt{a}) \vee h_2^{-1}(c_2'''/\sqrt{a}) \vee (c_M a^{-1/2})$ for some c_M depending on M. Thus

$$\lim_{n\to\infty} n^2 \frac{\phi_n(n^2 a)}{\phi_n(2n^2 a)} = 0, \qquad \forall a > 0. \tag{5.8}$$

Given this, the rest is a minor modification of the proof of [22, Lemma 2.3]. Let $a = m(\delta)^2/3 > 0$. Since $m_n \to m$ almost-surely (as discussed around (5.3)), there exists a random $n_1 \geq 1$ such that

$$m_n(t)^2 \geq m_n(\delta)^2 \geq 2a, \qquad \forall t \geq \delta,\ n \geq n_1. \tag{5.9}$$

Let $\mathcal{A}_{1,i} = \{\eta_{n,i} < n^{-2}\}$, $\mathcal{A}_{2,i} = \{n^{-2} \leq \eta_{n,i} \leq a\}$ and $\mathcal{A}_{3,i} = \{a < \eta_{n,i}\}$, and define

$$S_m^{n,k} = \sum_{i \leq m} \phi_n(n^2 \eta_{n,i}) \mathbf{1}_{\mathcal{A}_{k,i}}, \qquad k = 1, 2, 3.$$

Since ϕ_n is non-decreasing, $S_{nt}^{n,1} \leq nT\phi_n(1)$ for $t \leq T$. Further, for $n \geq a^{-1/2}$, we have that $n\phi_n(1)/\phi_n(2n^2 a) \leq n\phi_n(n^2 a)/\phi_n(2n^2 a)$, which goes to 0 as $n \to \infty$ by (5.8). Thus, there exists a random $n_2 \geq 1$ such that

$$S_{nt}^{n,1} \leq \phi_n(2n^2 a) \leq \phi_n(n^2 m_n(t)^2), \qquad \forall \delta \leq t \leq T,\ n \geq n_2, \tag{5.10}$$

where the last inequality is due to (5.9). Next, by (5.8), there exists a random $n_3 \geq 1$ such that

$$0 \leq \phi_n(n^2 x) \leq \phi_n(n^2 a) n^2 x \leq \phi_n(2n^2 a) x, \qquad \forall n^{-2} \leq x \leq a,\ n \geq n_3.$$

Thus, for $\delta \leq t \leq T$ and $n \geq n_3(\omega)$, we have

$$S_{nt}^{n,2} \leq \phi_n(2n^2 a) \sum_{i \leq nt} \eta_{n,i} = \phi_n(2n^2 a) \eta(\lfloor nt \rfloor / n) \leq \phi_n(n^2 m_n(t)^2) \eta(T), \tag{5.11}$$

where the last inequality is due to (5.9). Now, noting that there are only finitely many $t \in [0, T]$ such that $\Delta\eta(t) > a$, there exists a random $K_3 > 0$ such that $\sum_{i \leq nT} \mathbf{1}_{A_{3,i}} \leq K_3$ for large n, almost-surely. Using this and the definition of $m_n(t)$, there exists a random $n_4 \geq 1$ such that the following holds:

$$S_{nt}^{n,3} \leq K_3 \max_{i \leq nt} \phi_n(n^2 \eta_{n,i}) = K_3 \phi_n(n^2 m_n(t)^2), \qquad \forall \delta \leq t \leq T,\ n \geq n_4. \tag{5.12}$$

Combining (5.10), (5.11) and (5.12), we obtain

$$\sum_{i \leq nt} \phi_n(n^2 \eta_{n,i}) \leq K_2 \phi_n(n^2 m_n(t)^2), \qquad \forall \delta \leq t \leq T,\ n \geq n_2 \vee n_3 \vee n_4 =: n_0,$$

where $K_2 = 1 + \eta(T) + K_3$. Thus we have obtained the upper bound of (ii). □

Proof of Theorem 5.1. Given the above lemma, the proof of Theorem 5.1 is basically the same as that of [22, Theorem 2.1], and so we only sketch it briefly. Let

$$\zeta_n^{(n)}(t) = \frac{1}{n} L\left(\sum_{i \le nt} \phi_n\left(n^2 \eta_{n,i}\right)\right).$$

Then, by definition, $(\zeta_n^{(n)}(t))_{t\ge 0}$ is equal in law to $(\frac{1}{n}L(S^n_{nt}))_{t\ge 0}$. Further, as discussed around (5.3), $m_n \to m$ almost-surely with respect to the Skorohod J_1 topology. So, in order to complete the proof, it suffices to prove the following: for $T > 0$,

$$\sup_{0\le t\le T} \left|\zeta_n^{(n)}(t) - m_n(t)\right| \to 0, \tag{5.13}$$

almost-surely. Firstly,

$$\begin{aligned}\sup_{\delta\le t\le T} \left|\zeta_n^{(n)}(t) - m_n(t)\right| &\le \sup_{0\le t\le T} \max_{i=1,2} \left|n^{-1} L\left(K_i \phi_n(n^2 m_n(t)^2)\right) - m_n(t)\right| \\ &\le \sup_{0\le x\le m_n(T)} \max_{i=1,2} \left|n^{-1} L\left(K_i \phi_n(n^2 x^2)\right) - x\right|,\end{aligned}$$

where Lemma 5.3(ii) is used in the first inequality. This bound converges to 0 as $n \to \infty$ by Lemma 5.3(i). Secondly, using the monotonicity of $\zeta_n^{(n)}$ and m_n (and Lemma 5.3 again), we have

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{0\le t\le \delta} \left|\zeta_n^{(n)}(t) - m_n(t)\right| \le \lim_{\delta\to 0} \limsup_{n\to\infty} (\zeta_n^{(n)}(\delta) + m_n(\delta)) = \lim_{\delta\to 0} 2m(\delta) = 0,$$

almost-surely. We thus obtain (5.13). □

A Appendix: Skorohod topologies

For the convenience of the reader, we recall here, following [31], the definitions of the two topologies on the Skorohod space $D([0,\infty),\mathbb{R})$ that are applied in this article. We start by defining the Skorohod J_1 topology on $D([0,T],\mathbb{R})$, where $T \in (0,\infty)$, to be that induced by the metric

$$d_{J_1}(f,g) := \inf_{\lambda\in\Lambda} \left\{\|f\circ\lambda - g\| \vee \|\lambda - I\|\right\}, \tag{A.1}$$

where $\|\cdot\|$ is the uniform norm, I is the identity map on $[0,T]$ and Λ is the set of strictly increasing functions mapping $[0,T]$ onto itself. To define the M_1 topology on $D([0,T],\mathbb{R})$, we first introduce the notion of the completed graph of a function f in this space by defining

$$\Gamma_f := \left\{(t,x) \in [0,T]\times\mathbb{R} :\ x = \alpha f(t^-) + (1-\alpha) f(t) \text{ for some } \alpha \in [0,1]\right\}.$$

We then say that $u = (u_1(t), u_2(t))_{t\in[0,1]}$ is a parametric representation of Γ_f if it is a continuous bijection from $[0,1]$ to Γ_f whose first coordinate is non-decreasing, and define a metric on $D([0,T],\mathbb{R})$ by setting

$$d_{M_1}(f,g) := \inf_{u\in\Pi_f, v\in\Pi_g} \left\{\|u_1 - v_1\| \vee \|u_2 - v_2\|\right\},$$

where Π_f (resp. Π_g) is the set of parametric representations of f (resp. g). It is the topology that d_{M_1} induces that is the Skorohod M_1 topology on $D([0,T],\mathbb{R})$. Note that M_1 is a weaker

topology than J_1, in the sense that convergence in the latter implies convergence in former, but not vice versa.

To extend the above notions to $D([0,\infty),\mathbb{R})$, we characterise convergence in the Skorohod J_1 (or M_1) topology on this space by saying $f_n \to f$ if and only if $f_n \to f$ with respect to the Skorohod J_1 (or M_1) topology on $D([0,T],\mathbb{R})$ for every continuity point of f. (These topologies can also be described by metrics, see [31, Section 3], for example.) In particular, to establish weak convergence of a random sequence $(X^n)_{n\geq 1}$ to X with respect to the Skorohod J_1 (or M_1) topology on $D([0,\infty),\mathbb{R})$, we require that $(X^n)_{n\geq 1}$ converges weakly to X with respect to the Skorohod J_1 (or M_1) topology on $D([0,T],\mathbb{R})$ for every time T at which X is almost-surely continuous. Note that, since we only ever consider the limits $(m(t))_{t\geq 0}$ and $(m^{-1}(t))_{t\geq 0}$, which are both continuous at each fixed T with probability 1, in our setting we are always required to check that the relevant weak convergence of processes holds in $D([0,T],\mathbb{R})$ for every time T.

Acknowledgements

Part of this work was completed while D.C. was undertaking a three month JSPS Postdoctoral Fellowship at the Research Institute for Mathematical Sciences, Kyoto University, during which time he was most generously hosted by T.K. A.F. would like to thank Gérard Ben Arous for suggesting that these models would exhibit extremal aging.

References

[1] M. T. Barlow and T. Kumagai, *Random walk on the incipient infinite cluster on trees*, Illinois J. Math. **50** (2006), no. 1-4, 33–65 (electronic).

[2] M. Barma and D. Dhar, *Directed diffusion in a percolation network*, J. Phys. C: Solid State Phys. **16** (1983), 1451–1458.

[3] G. Ben Arous and J. Černý, *Dynamics of trap models*, Mathematical statistical physics, Elsevier B. V., Amsterdam, 2006, pp. 331–394.

[4] G. Ben Arous and J. Černý, *Scaling limit for trap models on $\mathbb{Z}^d$*, Ann. Probab. **35** (2007), no. 6, 2356–2384.

[5] G. Ben Arous and J. Černý, *The arcsine law as a universal aging scheme for trap models*, Comm. Pure Appl. Math. **61** (2008), no. 3, 289–329.

[6] G. Ben Arous, A. Fribergh, N. Gantert, and A. Hammond, *Biased random walks on Galton-Watson trees with leaves*, Ann. Probab. **40** (2012), no. 1, 280–338.

[7] G. Ben Arous and O. Gun, *Universality and extremal aging for dynamics of spin glasses on sub-exponential time scales*, Comm. Pure Appl. Math. **65** (2012), no. 1, 77–127.

[8] G. Ben Arous and A. Hammond, *Randomly biased walks on subcritical trees*, Comm. Pure Appl. Math., to appear.

[9] D. A. Croydon, *Slow movement of a random walk on the range of a random walk in the presence of an external field*, Preprint, arXiv:1203.0405.

[10] D. A. Croydon and T. Kumagai, *Random walks on Galton-Watson trees with infinite variance offspring distribution conditioned to survive*, Electron. J. Probab. **13** (2008), 1419–1441.

[11] D. A. Darling, *The influence of the maximum term in the addition of independent random variables*, Trans. Amer. Math. Soc. **73** (1952), 95–107.

[12] T. Duquesne and J.-F. Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), no. 4, 553–603.

[13] M. Dwass, *The total progeny in a branching process and a related random walk*, J. Appl. Probability **6** (1969), 682–686.

[14] N. Enriquez, C. Sabot, and O. Zindy, *Aging and quenched localization for one-dimensional random walks in random environment in the sub-ballistic regime*, Bull. Soc. Math. France **137** (2009), no. 3, 423–452.

[15] N. Enriquez, C. Sabot, and O. Zindy, *Limit laws for transient random walks in random environment on* $\mathbb{Z}$, Ann. Inst. Fourier (Grenoble) **59** (2009), no. 6, 2469–2508.

[16] W. Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971.

[17] A. Fribergh and A. Hammond, *Phase transition for the speed of the biased random walk on the supercritical percolation cluster*, Preprint, arXiv:1103.1371.

[18] J. Geiger and G. Kersting, *The Galton-Watson tree conditioned on its height*, Probability theory and mathematical statistics (Vilnius, 1998), TEV, Vilnius, 1999, pp. 277–286.

[19] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954, Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob.

[20] O. Gun, *Universality of transient dynamics and aging for spin glasses*, ProQuest LLC, Ann Arbor, MI, 2009, Thesis (Ph.D.)–New York University.

[21] B. Haas and G. Miermont, *The genealogy of self-similar fragmentations with negative index as a continuum random tree*, Electron. J. Probab. **9** (2004), no. 4, 57–97 (electronic).

[22] Y. Kasahara, *A limit theorem for sums of i.i.d. random variables with slowly varying tail probability*, J. Math. Kyoto Univ. **26** (1986), no. 3, 437–443.

[23] H. Kesten, *Sub-diffusive behavior of random walk on a random cluster*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 4, 425–487.

[24] G. Kozma and A. Nachmias, *The Alexander-Orbach conjecture holds in high dimensions*, Invent. Math. **178** (2009), no. 3, 635–654.

[25] J.-F. Le Gall, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), no. 1, 35–62.

[26] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson.

[27] R. Lyons, R. Pemantle, and Y. Peres, *Biased random walks on Galton-Watson trees*, Probab. Theory Related Fields **106** (1996), no. 2, 249–264.

[28] E. Seneta, *Regularly varying functions*, Springer-Verlag, Berlin, 1976, Lecture Notes in Mathematics, Vol. 508.

[29] R. S. Slack, *A branching process with mean one and possibly infinite variance*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **9** (1968), 139–145.

[30] W. Whitt, *Weak convergence of first passage time processes*, J. Appl. Probability **8** (1971), 417–422.

[31] W. Whitt, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002, An introduction to stochastic-process limits and their application to queues.

[32] O. Zindy, *Scaling limit and aging for directed trap models*, Markov Process. Related Fields **15** (2009), no. 1, 31–50.