arXiv:1101.4041v1 [math.PR] 20 Jan 2011

Randomly biased walks on subcritical trees

Gérard Ben Arous*, Alan Hammond †

January 24, 2011

Abstract

As a model of trapping by biased motion in random structure, we study the time taken for a biased random walk to return to the root of a subcritical Galton-Watson tree. We do so for trees in which these biases are randomly chosen, independently for distinct edges, according to a law that satisfies a logarithmic non-lattice condition. The mean return time of the walk is in essence given by the total conductance of the tree. We determine the asymptotic decay of this total conductance, finding it to have a pure power-law decay. In the case of the conductance associated to a single vertex at maximal depth in the tree, this asymptotic decay may be analysed by the classical defective renewal theorem, due to the non-lattice edge-bias assumption. However, the derivation of the decay for total conductance requires computing an additional constant multiple outside the power-law that allows for the contribution of all vertices close to the base of the tree. This computation entails a detailed study of a convenient decomposition of the tree, under conditioning on the tree having high total conductance. As such, our principal conclusion may be viewed as a development of renewal theory in the context of random environments.

For randomly biased random walk on a supercritical Galton-Watson tree with positive extinction probability, our main results may be regarded as a description of the slowdown mechanism caused by the presence of subcritical trees adjacent to the backbone that may act as traps that detain the walker. Indeed, this conclusion is exploited in [18] to obtain a stable limiting law for walker displacement in such a tree.

*Courant Institute of Mathematical Sciences, New York University. Research supported in part by the National Science Foundation under grants DMS-0806180 and OISE-0730136.

†Department of Statistics, Oxford University. Research supported in part by NSF grants DMS-0806180 and OISE-0730136 and by EPSRC grant EP/I004378/1. This research was initiated while the second author was at N.Y.U.

1 Introduction

The inquiry into drift and trapping for biased random walks on random structures has been pursued both by physicists and mathematicians. It was noted long ago in the physics literature [7] that, in disordered media, due to trapping in dead-end branches, the mean velocity would not be a monotone function of the bias. It was then pointed out that, in fact, strong biases could even produce sub-ballistic regimes, i.e., with zero velocity; (see [11], [12], or [19] for a physics survey of the general phenomena of anomalous diffusion). This hypothesis was confirmed for biased random walks on random trees by Lyons, Pemantle and Peres [25], and on supercritical percolation clusters, independently by Berger, Gantert and Peres in dimension two [6], and by Sznitman in all dimensions $d \geq 2$ [28]. For $d \geq 2$, the existence of a critical bias separating the sub-ballistic and ballistic regimes will be demonstrated in [17]. These studies naturally raise the question of understanding more deeply the means by which the walk is slowed down by trapping structures in this sub-ballistic regime. A. Sznitman mentions in [29] that this mechanism seems similar to that responsible for aging in Bouchaud's trap model ([4],[8]) whose main features are that the mean trapping times have a power-law distribution and that the tail of the distribution of the trapping times, conditionally on their mean, is exponential.

For random walks on supercritical Galton-Watson trees whose bias away from the root is constant, such a study of the trapping phenomenon was undertaken in [3]. In this context, the dead-ends responsible for the slowing down of the walk are simply subcritical Galton-Watson trees that hang off the backbone of the supercritical tree. Thus, the trapping times are related to return times to the root for biased random walks on subcritical trees. The paper [3] analyses a log-periodicity phenomenon which is very reminiscent of the classical lattice effect for random walks in random environments [21], [13], [14], [30]. This lattice effect, which has been discussed in the physics literature [27], causes holding times in traps to cluster around powers of the bias parameter. That is to say, although it may seem reasonable that such a walk has a scaling limit similar to the one obtained for Bouchaud trap models that is discussed in [4] and [5], the lattice effect is a discrete inhomogeniety in walker displacement that is persistent on all time scales and that prevents the existence of such a scaling limit.

In this paper, we study the question of how long it takes a biased random walk in a subcritical Galton-Watson tree to return to the root of the tree. In the context of biased walks on supercritical trees, our inquiry amounts to an investigation of the nature of the delaying mechanism of the subcritical trees that act as potential traps for the walk. We undertake the inquiry in

the case that the biases on edges that determine the distribution of the walk are random, rather than identically equal, imposing a non-lattice condition on this randomness that serves to eliminate the log-periodic effect from the return-time distribution.

This paper is the first of two. Indeed, its sequel [18] exploits the understanding of the trapping mechanism that we obtain in the present article to prove a stable scaling limit for the randomly biased random walk on supercritical Galton-Watson trees with leaves; in this way, the two papers form a counterpart to [3], in that they show how edge-bias randomization dissipates the persistent discrete inhomogeneity that obtains in the case of constant bias.

Randomly biased random walks on infinite trees have been studied by [24], who provide a criterion for recurrence, or transience, valid for a very broad class of trees. In the case of supercritical Galton-Watson trees, randomly biased walk is studied in the recurrent regime by [15], who investigate the high-n asymptotic of the maximal displacement of the walk during $[0, n]$. In [1], the same model is analysed in the case of zero extinction probability and in the transient regime, the author presenting criteria for zero speed and for positive speed. The tree being leafless in [1], this work is not concerned with trapping, but rather is a contribution to understanding how increasing bias causes more rapid progress for the walk on the backbone of a supercritical Galton-Watson tree. Regarding this last question, the monotonicity of speed as a function of bias is not known even when the bias is a constant: see Question 2.1 of [26].

We now define the general class of biased random walks on finite rooted trees in which we are interested. Let T be a finite rooted tree, with vertex set $V(T)$, edge set $E(T)$, and root ϕ. Such a tree will be called weighted if, to each unoriented edge $e \in E(T)$ is assigned a number $\beta_e \in [q, Q]$, that we will call the bias of the edge e. Here, $Q \geq q > 1$ are fixed constants. In this paper, in referring to a tree, we will generally mean a weighted tree. We define the law $\mathbb{P}_{T,\beta}$ of the β-biased random walk $\{X_i : i \in \mathbb{N}\}$ on the set of vertices of the tree T as a Markov chain on $V(T)$ with the following transition rules.

Definition 1 *If a vertex $v \in V(T)$, $v \neq \phi$, has offspring $v_1, \ldots, v_k$, then, for each $n \in \mathbb{N}$,*

$$\begin{aligned}
\mathbb{P}_{T,\beta}\Big(X_{n+1} = \overleftarrow{v} \Big| X_n = v\Big) &= \frac{1}{1 + \sum_{i=1}^k \beta_{v,v_i}}, \\
\mathbb{P}_{T,\beta}\Big(X_{n+1} = v_j \Big| X_n = v\Big) &= \frac{\beta_{v,v_j}}{1 + \sum_{i=1}^k \beta_{v,v_i}}, \qquad \textit{for } 1 \leq j \leq k,
\end{aligned}$$

where $\overleftarrow{v}$ denotes the parent of v, i.e., the neighbor of v which is the closest to the root. The jump-law from the root is given by

$$\mathbb{P}_{T,\beta}\Big(X_{n+1}=\phi_j\Big|X_n=\phi\Big)=\frac{\beta_{\phi,\phi_j}}{\sum_{i=1}^{k}\beta_{\phi,\phi_i}},\qquad \text{for } 1\le j\le k(\phi).$$

For $v\in V(T)$, we write $\mathbb{P}^v_{T,\beta}$ for the law of $\mathbb{P}_{T,\beta}$ given that $X(0)=v$. We write $\mathbb{E}^v_{T,\beta}$ for the expectation value under $\mathbb{P}^v_{T,\beta}$.

It is easy to see these biased random walks are reversible and to compute their invariant measures. Indeed, they fall in the general class of "random walks on weighted graphs"(see [22], and [2] for recent expositions).

Definition 2 *For every vertex $v\in V(T)$, define $P_{\phi,v}$ to be the unique simple path from the root to v, and $\omega(v)$ to be the product of the biases β_e of the edges e along the path $P_{\phi,v}$. If $e=(v,w)$ is an edge of T, with $v=\overleftarrow{w}$, we define the weight (or conductance) of e by $\mu(e)=\mu(v,w)=\omega(w)$. Finally, we define the unnormalized measure μ on $V(T)$ by setting $\mu(v)$ equal to the sum of the conductances of the edges adjacent to the vertex $v\in V(T)$.*

It is clear that the measure μ is reversible and thus invariant for the Markov chain X_n. Moreover, by the general random walks on weighted graphs, the transition rules between the adjacent vertices v and w can simply be rewritten in terms of the conductance of the edge between v and w as well as the measure $\mu(v)$:

$$\mathbb{P}_{T,\beta}\Big(X_{n+1}=w\Big|X_n=v\Big)=\frac{\mu(v,w)}{\mu(v)}. \tag{1}$$

We will be interested in the first return time to the root of the walk that starts at the root, where the return time is naturally defined as $\inf\{n>1, X_n=\phi\}$ under the law $\mathbb{P}^{\phi}_{T,\beta}$.

As we have mentioned, the random trees that we consider are subcritical Galton-Watson trees.

Definition 3 *Let $h=\{h_i : i\in\mathbb{N}\}$, $\sum_{i=0}^{\infty}h_i=1$, denote an offspring distribution. We write $\mathbb{P}_h$ for the Galton-Watson tree with offspring distribution h, that is, for the law on rooted trees in which each vertex has an independent and h-distributed number of offspring.*

We will make the following assumption on the offspring distribution.

Hypothesis 1 *Let $\{h_i : i\in\mathbb{N}\}$, $\sum_{k=1}^{\infty}kh_k<1$, be a subcritical offspring distribution for which there exists $c>0$ such that $\sum_{l\ge k}h_k\le\exp\{-ck\}$ for each $k\in\mathbb{N}$. We write $m_h=\sum_{k=1}^{\infty}kh_k$ for the mean number of offspring.*

We now describe the law of the random biases.

Definition 4 *Let ν be a probability distribution on $(1,\infty)$. For a tree T sampled from $\mathbb{P}_h$, let $(\beta_e)_{e\in E(T)}$ be independent and identically distributed random variables, with common distribution ν. We write $\mathbb{P}_{h,\nu}$ for the Galton-Watson tree with offspring distribution h equipped with this set of random biases.*

We make the following important non-lattice assumption on the distribution ν.

Hypothesis 2 *There exist $Q > q > 1$ such that the support of the measure ν is contained in $[q, Q]$. Moreover, the support of $\nu \circ \log^{-1}$ is non-lattice. That is, the $\mathbb{Z}$-linear span of $\log \mathrm{supp}(\nu)$ is dense in $\mathbb{R}$.*

The fact that the distribution of the bias is supported in $(1,\infty)$ means that the walk is biased away from the root. We now need to introduce the following important exponent.

Definition 5 *Let $\chi > 0$ be the unique value satisfying*

$$\int_1^\infty y^\chi \nu(dy) = \frac{1}{m_h}. \tag{2}$$

Our first result shows that, under the assumptions above, the mean return time to the root has a pure power-law tail. The regularity of the tail of the random variable constrasts with the case of constant bias studied in [3].

Definition 6 *Under the law $\mathbb{P}^\phi_{T,\beta}$, let $H_\phi = \inf\{n \geq 1, X_n = \phi\}$ denote the hitting time of the root.*

Here and throughout, by $A(x) \sim B(x)$ is meant $\frac{A(x)}{B(x)} \to 1$ as $x \to \infty$.

Theorem 1 *Assume Hypotheses 1 and 2. Then, here exists a constant $c_1 \in (0,\infty)$ such that*

$$\mathbb{P}_{h,\nu}\Big(\mathbb{E}^\phi_{T,\beta}(H_\phi) > x\Big) \sim c_1 x^{-\chi}.$$

where $c_1 \in (0,\infty)$ is a constant that will be specified after Theorem 2.

The electrical resistance theory of random walks, or, equivalently, the theory of random walk on weighted graphs (see [2] or [22]), may be used to compute mean return times in terms of the total conductance (or weight) of the tree. Indeed, defining $\omega(T)$ to be this total weight $\sum_{e\in E(T)} c(e)$, we see that

$$\omega(T) = \sum_{v\in V(T)} \omega(v). \tag{3}$$

The mean return time is then given by

$$\mathbb{E}^{\phi}_{T,\beta}(H_\phi) = \frac{2}{N}(\omega(T) - 1) \tag{4}$$

where N be the number of offspring of the root. This formula arises by averaging the return time over the offspring of the root visited at the first step of the walk, with the commute-time formula of electrical resistance theory being used to express the mean return time from offspring to root. See the upcoming (7) for the relevant expression for this mean return time.

Using (4), Theorem 1 is a simple consequence of the following pure power-law tail for the quantity $\omega(T)$.

Theorem 2 *Assume Hypotheses 1 and 2. There exists a constant $d_1 \in (0, \infty)$ such that*

$$\mathbb{P}_{h,\nu}\Big(\omega(T) > u\Big) \sim d_1 u^{-\chi},$$

Remark. The two constants c_1 and d_1 in Theorems 1 and 2 are explicitly given by

$$c_1 = 2^{-\chi} d_1 \sum_{k=1}^{\infty} h_k k^{1-\chi} \tag{5}$$

and, denoting by $D(T) = \sup\{d(\phi, v) : v \in V(T)\}$ the depth of the tree T,

$$d_1 = \frac{1}{\chi m_h}\Big(\int_1^\infty y^\chi \log(y)\nu(dy)\Big)^{-1} \lim_{k\to\infty} \mathbb{E}_{h,\nu}\Big(\omega(T)^\chi \mathbb{1}_{D(T)=k}\Big). \tag{6}$$

It is natural to ask about not only the mean of the return time but also about how it is distributed. We will present here an asymptotically accurate description of this distribution in the case that the tree is large (in a natural sense). This description plays a central role in the analysis of randomly biased walk on a supercritical Galton-Watson tree undertaken in [18]. We will see that the behaviour of the return time distribution has two entirely distinct regimes, according to whether or not the walker has the opportunity to visit the base of the tree (and so explore most of the tree) before its return to the root. We will condition on the events that the tree is large and that the walker makes such a deep visit, and conclude that the tail of the return time is then exponential, completing the analogy with the Bouchaud trap model that we have mentioned.

Firstly, we introduce notation for conditioning that a tree be "large" enough. A tree T will be regarded as large if $\omega(T)$ is large.

Definition 7 *For any $u > 0$ we set $\mathbb{P}_{h,\nu,u} = \mathbb{P}_{h,\nu}(\,\cdot\,|\omega(T) > u)$.*

We then introduce our notation for the distribution of the random walks. We sample a tree T equipped with its biases (β_e)'s , choose a vertex $v \in V(T)$, and run the Markov chain on $V(T)$, started at v, that is specified in Definition 1. We denote by $\mathbb{P}_{h,\nu} \times \mathbb{P}^v_{T,\beta}$ the joint distribution of the tree, the biases and the random walk started at v. We will also need to sample the weighted tree from the conditioned measure $\mathbb{P}_{h,\nu,u}$. Naturally, we will then denote the joint distribution of the tree, the biases and the walk by $\mathbb{P}_{h,\nu,u} \times \mathbb{P}^v_{T,\beta}$.

A technical point is that, for a random tree T, we need to specify a selection rule for the starting point $v \in V(T)$ of the walk. A selection rule can be easily defined using the classical formal coding of Galton-Watson trees (see [23] Section 1.1). In Appendix A, we review this coding and the definition of a selection rule.

We also need to introduce notation for conditioning that the random walk explore enough of the tree. In the next definition, we again use the coding of trees in Appendix A to specify a lexicographical ordering on the set of vertices of any Galton-Watson tree.

Definition 8 *We call v_{base} the lexicographically minimal vertex of maximal depth in $V(T)$.*

Definition 9 *Under the law $\mathbb{P}^\phi_{T,\beta}$, the walk X is said to make a deep excursion into T if it hits v_{base} before returning to the root. We write $\mathcal{DE} = \{H_{v_{\text{base}}} < H_\phi\}$ for the event that X makes a deep excursion into T. We will write $\mathbb{P}^\phi_{T,\beta,\mathcal{DE}}$ the distribution of the random walk conditioned by the event that it makes a deep excursion.*

We remark that, for a high choice of u, it is reasonable to suppose that, under the law $\mathbb{P}_{h,\nu,u}$, the tree is typically long and thin, consisting of a long path connecting the root to the base vertex v_{base}, with only small subtrees hanging off this path. Indeed, in Theorem 6, we will present a result to this effect. For a typical large tree, then, a walk from ϕ that makes a deep excursion into T will be forced after reaching v_{base} to make a geometrically distributed number of highly ineffectual attempts to reach the root from v_{base}, so that H_ϕ will be approximately exponentially distributed in this case. This is the content of our next result.

Theorem 3 *Assume Hypothesis 1, and that there exist $Q > q > 1$ such that the support of the measure ν is contained in $[q, Q]$. For all $t > 0$,*

$$\left(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^\phi_{T,\beta,\mathcal{DE}}\right)\left(\frac{H_\phi}{\mathbb{E}^\phi_{T,\beta,\mathcal{DE}}(H_\phi)} > t\right) \to \exp\{-t\}$$

as $u \to \infty$.

We also present an asymptotically accurate expression for the mean return time $\mathbb{E}^{\phi}_{T,\beta,\mathcal{DE}}(H_\phi)$. To do so, we introduce some notation.

Definition 10 *Given a tree T and $v \in V(T)$, we define the descendent tree T_v of v to be the subgraph induced by the set of all descendents of v (among which, we include v). The root of T_v is taken to be v.*

Definition 11 *Let T be a weighted tree. Set v_{child} to be the neighbour of the root lying in the path $P_{\phi,v_{\text{base}}}$ from the root to v_{base}. The descendent tree $T_{v_{\text{child}}}$ is itself a weighted tree with root v_{child}. We set ω_* equal to the weight of this tree, as defined by the formula (3), where, to compute the summand on the right-hand-side, we note that, for each element $v \in V(T_{v_{\text{child}}})$, the weight of v is computed with respect to the root being v_{child}.*

We further write $p_{\text{de}} = \mathbb{P}^{v_{\text{child}}}_{T,\beta}(\mathcal{DE})$ for the probability that the walk starting at v_{child} makes a deep excursion into T.

The significance of ω_* is its appearance in the mean hitting time formula

$$E^{v_{\text{child}}}_{T,\beta}(H_\phi) = 2\omega_* - 1, \tag{7}$$

which follows from the commute-time formula in a reversible network, originally proved in [9], and presented as Theorem 3.3 in [2].

Theorem 4 *Assume the hypotheses of Theorem 3. For each $\epsilon > 0$,*

$$\mathbb{P}_{h,\nu,u}\left(\frac{\mathbb{E}^{\phi}_{T,\beta,\mathcal{DE}}(H_\phi)}{2\omega_* p_{\text{de}}^{-1}} \in \Big(1-\epsilon, 1+\epsilon\Big)\right) \to 1$$

as $u \to \infty$.

For all $t > 0$,

$$\Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{\phi}_{T,\beta,\mathcal{DE}}\Big)\left(\frac{H_\phi}{2\omega_* p_{\text{de}}^{-1}} > t\right) \to \exp\{-t\}$$

as $u \to \infty$.

The new information presented in Theorem 4 beyond that of Theorem 3 is that, for a typical large tree, under $\mathbb{P}^{\phi}_{T,\beta}$, if $\mathcal{DE}$ does not occur, then H_ϕ is negligible. Indeed, that H_ϕ is negligible in the event $\mathcal{DE}^c$ is equivalent to saying that most of the mean $\mathbb{E}^{\phi}_{T,\beta}(H_\phi)$ arises on the event $\mathcal{DE}$. It easy to see that this in turn is the same as saying that $\mathbb{E}^{\phi}_{T,\beta,\mathcal{DE}}(H_\phi)$ is well approximated by $2\omega_* p_{\text{de}}^{-1}$. (The key ingredients to see the latter equivalence are (7), and the fact that, under $\mathcal{DE}$, the walk from ϕ must immediately travel to v_{child}.)

1.1 The structure of the paper

We begin in Section 2 by introducing some necessary tools from classical defective renewal theory and the extension needed for the proof of Theorem 2, which forms the core of the argument for Theorem 1. We will illustrate how the defective renewal theory can be used straightforwardly to show that, for a vertex $v \in V(T)$ of maximal depth, $\omega(v)$ has a pure power-law tail. This power-law differs by a constant from the one given in Theorem 2 for $\omega(T)$, the reason being that several vertices near the base of the tree contribute to $\omega(T)$ on roughly equal terms. In order to better approximate $\omega(T)$ and thus prove Theorem 2, we introduce a convenient decomposition of the tree in Section 3. The main result, Theorem 6, about this decomposition states that the components are small conditionally on $\omega(T)$ being large. The proof of Theorem 6 is fairly involved and in places delicate, and it is deferred to Appendix B. In Section 4, we prove Theorem 2. In Section 5, we study the exponential tail of the return time, and prove Theorems 3 and 4.

In its exploration of time spent in traps, the paper [18] requires a different decomposition from the one introduced in Section 3. In Appendix C, we introduce this new decomposition of the tree, which we call the renewal decomposition. The renewal decomposition has components that enjoy more independence than the corresponding objects in the decompositions used in this paper. The analogue for the renewal decomposition of Theorem 6 is needed in [18]. Its proof also appears in Appendix C.

2 A first use of defective renewal theory

We recall and improve slightly on the classical non-lattice defective renewal theorem [16]. Suppose that a succession of light bulbs has been produced. Independently of the earlier ones, a given light bulb is *permanent* with probability $1-p$, in which case, it will shine eternally. A light bulb which is not permanent is called *temporary*; given that a bulb is temporary (and independently of the status of the preceding bulbs), it has a lifetime distributed according to the law μ. At time zero, the first bulb is installed and begins to shine. A bulb is replaced whenever it fails, until a permanent one is installed. The defective renewal theorem is concerned with the law of the time at which the final replacement is made (after which, a bulb will eternally shine). This is the time Z in the next lemma, whose first part is the classical defective renewal theorem.

Lemma 1 *On a probability space* $(\Omega, \mathbb{P})$*, let* $\{X_i : i \in \mathbb{N}\}$ *be a sequence of independent and identically distributed random variables. The common*

distribution μ is assumed to have a non-negative, bounded and non-lattice support. Let Y denote an independent random variable taking values in the positive integers. Set $Z = \sum_{i=1}^{Y} X_i$.

1. *Suppose that, under $\mathbb{P}$, Y has the law of a geometric random variable of parameter $p \in (0,1)$: that is, for each $j \geq 0$, $\mathbb{P}(Y = j) = p^j(1-p)$. Define $\kappa > 0$ as the unique value such that $\mathbb{E}(e^{\kappa X}) = \frac{1}{p}$, and let $c_2 = \frac{1-p}{\kappa p \mathbb{E}(Xe^{\kappa X})}$. Then*
$$\mathbb{P}(Z \geq u) \sim c_2 \exp\{-\kappa u\},$$
where recall that $f(u) \sim g(u)$ denotes $\lim_{u\to\infty} \frac{f(u)}{g(u)} = 1$.

2. *Suppose that there exist $p \in (0,1)$, $c \in (0,\infty)$ and $k_0 \in \mathbb{N}$ such that, for $k \geq k_0$, $\mathbb{P}(Y = k) = cp^k$. Then*
$$\mathbb{P}(Z \geq u) \sim c_3 \exp\{-\kappa u\},$$
with $c_3 = \frac{cc_2}{1-p}$.

3. *Suppose that, for such p and c, we assume merely that $\mathbb{P}(Y = k) \sim cp^k$ as $k \to \infty$. In this case also,*
$$\mathbb{P}(Z \geq u) \sim c_3 \exp\{-\kappa u\}.$$

Proof. The first statement is (6.16) on page 377 of [16]. (Note that there is typographical error in (6.16) and that μ should be replaced by $\mu^\sharp$.)

In the second statement, Y assumes high values according to a law whose density is adjusted from that of a geometric law by multiplication by a factor of $\alpha = c/(1-p) \in (0,\infty)$. Suppose that $\alpha \in (0,1]$. Then we may equip the probability space $(\Omega, \mathbb{P})$ with a geometric random variable V and an independent event A satisfying $\mathbb{P}(A) = \alpha$ in such a way that $Y\mathbb{1}_{Y\geq k_0} = V\mathbb{1}_{V\geq k_0}\mathbb{1}_A$. Writing Q for the supremum of the support of μ, we have that, if $u \geq k_0 Q$, then $Z \geq u$ implies that $Y \geq k_0$. We find then that, for such u, $\mathbb{P}(Y \geq u) = \mathbb{P}(V \geq u)\mathbb{P}(A) = \alpha\mathbb{P}(V \geq u)$, so that the second statement of the lemma follows from the first one in this case.

In the case that $\alpha > 1$, we similarly equip $(\Omega, \mathbb{P})$ with a geometric random variable V and an independent event A for which $\mathbb{P}(A) = \alpha^{-1}$ such that $V\mathbb{1}_{V\geq k_0} = Y\mathbb{1}_{Y\geq k_0}\mathbb{1}_A$. We may then treat this case analogously to the preceding one.

The third statement may be reduced to the second one by a coupling argument. Let Y' denote a random variable satisfying $\mathbb{P}(Y' = k) = cp^k$ for

$k \geq k_0$, for some $k_0 \in \mathbb{N}$. We claim that a coupling $\mathbb{Q}$ of Y and Y' may be effected such that

$$\lim_{n\to\infty} \mathbb{Q}\Big(Y = Y' \Big| Y' = n\Big) = 1 \tag{8}$$

and

$$\lim_{n\to\infty} \mathbb{Q}\Big(Y = Y' \Big| Y = n\Big) = 1. \tag{9}$$

Indeed, we may construct $\mathbb{Q}$ to satisfy $\mathbb{Q}(Y = n, Y' = n) = p_n \wedge q_n$, with $p_n = \mathbb{P}(Y = n)$ and $q_n = \mathbb{P}(Y' = n)$. We then have that $\mathbb{Q}(Y \neq Y' | Y' = n) = \max\{0, 1 - p_n/q_n\}$, which tends to 0 as $n \to \infty$. This yields (8), with the same argument giving (9).

We now define $Z' = \sum_{i=1}^{Y'} X_i$, on $(\Omega, \mathbb{P})$, and claim that

$$\lim_{u\to\infty} \frac{\mathbb{P}(Z' > u)}{\mathbb{P}(Z > u)} = 1, \tag{10}$$

so that the third statement of the lemma follows from the second applied to Z'. To prove (10), we construct under the measure $\mathbb{Q}$ the sequence $\{X_i : i \in \mathbb{N}\}$ having the same law as under $\mathbb{P}$, and doing so independently of Y and Y'. In this way, $\mathbb{Q}$ provides a coupling of Z and Z'. It suffices for (10) to show that

$$\lim_{u\to\infty} \mathbb{Q}\Big(Z' > u \Big| Z > u\Big) = 1 \tag{11}$$

and

$$\lim_{u\to\infty} \mathbb{Q}\Big(Z > u \Big| Z' > u\Big) = 1. \tag{12}$$

Given $Z \geq u$, it suffices for $Z' \geq u$ that $Y = Y'$. The conditional distribution of Y, given that $Z \geq u$, is supported on $[u/Q, \infty)$, where recall that $Q = \sup \operatorname{supp}(\mu)$. Moreover, given Y, the event that $Y = Y'$ is conditionally independent of Z. Hence, $\mathbb{Q}(Y = Y' | Z \geq u) \to 1$ as $u \to \infty$ is a consequence of (9). This proves (12). Likewise, (11) is derived by means of (8). $\square$

We now present Theorem 5, which states that $\omega(v)$ has a power-law tail for a vertex $v \in V(T)$ of maximal depth. The main tool is Lemma 1. We will need the vertex v to be chosen independently of the biases $\{\beta_e : e \in E(T)\}$. For this theorem, any algorithm to do so would be satisfactory: for instance, we could choose v uniformly at random among those vertices of maximal depth, independently of the biases. For definiteness, we specify one such algorithm: indeed, we have already done so, in specifying the vertex v_{base} in Definition 8 by means of the classical formal coding of the Galton-Watson tree (see [23] Section 1.1, and Appendix A).

Theorem 5 *Assume Hypotheses 1 and 2. The weight of the base vertex satisfies*

$$\mathbb{P}_{h,\nu}\Big(\omega(v_{\text{base}}) > u\Big) \sim d_2 u^{-\chi},$$

where χ is given in (2) and where

$$d_2 = \frac{\alpha}{\chi m_h \int_1^\infty y^\chi \log(y) d\nu(y)}, \tag{13}$$

with the constant $\alpha \in (0,\infty)$ being given by the limit

$$\alpha := \lim_{n\to\infty} m_h^{-n} \mathbb{P}_h\Big(D(T) = n\Big) \in (0,\infty). \tag{14}$$

The fact that the distribution of the depth is asymptotically geometric, i.e., that the limit (14) exists, is classical under Hypothesis 1. See [20].

Proof. We firstly note that the existence of the limit (14) is classical under Hypothesis 1. See [20].

The vertex v_{base} is at distance $D(T)$ from the root ϕ. Under $\mathbb{P}_{h,\nu}$, the biases attached to the path from ϕ to v_{base} are independent samples of the law ν. Hence, under $\mathbb{P}_{h,\nu}$,

$$\log \omega(v_{\text{base}}) = \sum_{i=1}^{D(T)} X_i, \tag{15}$$

where $\{X_i : 1 \le i \le D(T)\}$ is a sequence of independent random variables, each having the law $\nu \circ \log^{-1}$. Note that this sum satisfies the hypotheses of the third part of Lemma 1, due to $D(T)$ being independent of $\{X_i : i \in D(T)\}$, and (14). Applying this result, we find that

$$\mathbb{P}_{h,\nu}\Big(\log \omega(v_{\text{base}}) > u\Big) \sim d_2 e^{-\chi u}. \tag{16}$$

where $\chi > 0$ satisifes (2) and where d_2 has the form given in (13). $\square$

The quantity $\omega(v_{\text{base}})$ is a first approximation to $\omega(T)$, for a tree sampled under the measure $\mathbb{P}_{h,\nu,u}$. This approximation is never accurate up to leading order, although it is plausible that, for a typical sample of this measure, $\omega(T)$ is approximated up to a small multiplicative correction by a sum of $\omega(v)$ over vertices v ranging over several generations near the end of the trap. We will introduce such an approximation, and analyze it, also by means of the defective renewal theorem. We now define a decomposition of a tree, in order to make this approximation.

3 Bare trees and approximation of the total weight

3.1 Approximating $\omega(T)$ using a splitting of the tree

We consider firstly a simple splitting of the tree T and a convenient approximation of its total weight $\omega(T)$.

Definition 12

1. *We record the vertices in $P_{\phi,v_{\text{base}}}$ in the form $[\phi = \psi_0, \psi_1, \ldots, \psi_{D(T)} = v_{\text{base}}]$.*
2. *For $0 \leq i \leq D(T)$, let J_i denote the connected component containing ψ_i of the graph with vertex set $V(T)$ and edge-set $E(T) \setminus E(P_{\phi,v_{\text{base}}})$. We will call J_i the i-th outgrowth of T. Note that $J_{D(T)}$ is the singleton graph with vertex v_{base}.*
3. *For any integer k, define $w_k(T) = \sum_{i=0}^{k} \sum_{v \in V(J_{D(T)-i})} \omega(v)$.*

The quantity $w_k(T)$ will provide a better approximation to $\omega(T)$ than that which $\omega(v_{\text{base}})$ offers. Before explaining how this is so, we need a little more notation.

Definition 13 *Let $k \in \mathbb{N}$. Let T be a weighted tree for which $D(T) \geq k$. Set E_k to be the weighted descendent tree $T_{\psi_{D(T)-k}}$, or, equivalently, the weighted subgraph induced by the set of vertices $v \in \cup_{i=D(T)-k}^{D(T)} V(J_i)$. Note that $w_k(T) = \sum_{v \in V(E_k)} \omega(v)$.*

We also set E_k^* equal to the weighted subtree of T induced by the vertex set given by removing the strict descendents of $\psi_{D(T)-k}$ from $V(T)$. Note that $E(E_k) \cup E(E_k^*)$ is a partition of $E(T)$.

We extend the notation $\omega(v)$ for the weight of a vertex $v \in V(T)$ in a tree T in the following way.

Definition 14 *Let T be a weighted tree, and let $u, v \in V(T)$, with v being a descendent of u. We write $\omega_u(v)$ for the product of the edge-weights in the path $P_{u,v}$. Note that, for any $u \in V(T)$, we have that $\omega_\phi(u) = \omega(u)$.*

The quantity $w_k(T)$ has a convenient representation. For any $v \in V(E_k)$, we may write $P_{\phi,v}$ as the concatenation of $P_{\phi,\psi_{D(T)-k}}$ and $P_{\psi_{D(T)-k},v}$. By so doing, we obtain $\omega(v) = \omega(\psi_{D(T)-k})\omega_{\psi_{D(T)-k}}(v)$. By summing this formula over $v \in V(E_k)$, we arrive at

$$w_k(T) = u_k(T)v_k(T). \tag{17}$$

Here, $u_k(T) := \omega(\psi_{D(T)-k})$ and $v_k(T) := \sum_{v \in V(E_k)} \omega_{\psi_{D(T)-k}}(v)$. This product formula will be essential for our analysis of the approximation $w_k(T)$.

3.2 Outgrowths in a high-weight tree are small

To obtain Theorem 2, as well as deriving the asymptotic behaviour of $w_k(T)$, we must show that $w_k(T)$ is a good enough approximation to $\omega(T)$. To do so, we need to know that the, outgrowths in a typical sample of $\mathbb{P}_{h,\nu}$ are not too large, uniformly under conditioning on the sample being a tree of any high weight. This estimate is provided by the next theorem, which is stated using the notation introduced in Definition 7.

Theorem 6 *For the statement, we take $J_i = \emptyset$ if $i > D(T)$ (for any tree T). Let $\{h_i : i \in \mathbb{N}\}$ satisfy Hypothesis 1, and let ν be an edge-weight law with compact support in $(1, \infty)$. There exists $c > 0$ such that, for all $u > 0$ and $i \in \mathbb{N}$,*

$$\mathbb{P}_{h,\nu,u}\Big(|V(J_i)| \geq k\Big) \leq \exp\{-ck\},$$

for each $k \in \mathbb{N}$.

While the statement of Theorem 6 is natural enough, its proof is a little technical. We defer it to Appendix B.

We also record the following property of the decomposition.

Lemma 2 *The law of the tree E_k under $\mathbb{P}_{h,\nu}(\cdot|D(T) \geq k)$ coincides with that of T under $\mathbb{P}_{h,\nu}(\cdot|D(T) = k)$. Moreover, this statement holds if $\mathbb{P}_{h,\nu}(\cdot|D(T) \geq k)$ is conditioned on any admissible choice of E_k^*.*

Proof. Suppose that a sample T of $\mathbb{P}_{h,\nu}(\,\cdot\,|D(T) \geq k)$ is further conditioned on an admissible choice of E_k^*. (We write $\tilde{\mathbb{P}}$ for the conditioned law.) The tree E_k, which, by definition, is given by the descendent tree $T_{\psi_{D(T)-k}}$, must have $D(E_k) = k$, by the definition of $\psi_{D(T)-k}$. Moreover, E_k may take the value of any such tree. To see this, note firstly that there must exist a tree T' with $D(T') = k$ in the support of the random variable E_k under $\tilde{\mathbb{P}}$, since the choice of E_k^* under which we condition to obtain $\tilde{\mathbb{P}}$ is an admissible one; and note further that the tree obtained by making the choice $E_k = T'$ has $v_{\text{base}} \in V(E_k)$. This forces $\psi_{D(T)-k}$ to be lexicographically smaller than any vertex $v \in V(T) \setminus V(E_k)$ for which $d(\phi, v) = D(T)$. This means that, for an arbitrary choice of $E_k = T''$, with T'' a tree for which $D(T'') = k$, we necessarily have $v_{\text{base}} \in V(E_k)$, so that, indeed, E_k may take the value T''.

Hence, under $\tilde{\mathbb{P}}$, E_k has the law of $\mathbb{P}_{h,\nu}$ subject merely to the condition that $D(E_k) = k$. $\square$

3.3 Bare trees

The following definition and lemma reformulate Theorem 6 in a manner convenient to applications.

Definition 15 *For $B > 0$, we say that a weighted tree T is B-bare if $|V(J_i)| \leq B \log\log \omega(T)$ for each $i \in \{0, \ldots, D(T)-1\}$.*

Lemma 3 *Let $c > 0$ denote the constant that appears in Theorem 6. Fix $B > 2/c$. Then*

$$\mathbb{P}_{h,\nu,u}\Big(T \text{ is } B\text{-bare}\Big) \geq 1 - (\log u)^{-cB/2},$$

for sufficiently high u.

Proof. Note that $\omega(T) \geq \omega(v_{\text{base}}) \geq q^{D(T)}$. Hence, by (14),

$$\mathbb{P}_{h,\nu}\Big(\omega(T) > u\Big) \geq \mathbb{P}_{h,\nu}\Big(D(T) > \frac{\log u}{\log q}\Big) \sim c_4 u^{-\frac{\log\left(m_h^{-1}\right)}{\log q}}, \tag{18}$$

for some $c_4 \in (0, \infty)$.

By (14) once more, and (18), we see that, for any $C > \frac{1}{\log q}$ and for all sufficiently high u,

$$\begin{aligned} \mathbb{P}_{h,\nu,u}\Big(D(T) > C\log u\Big) &\leq \frac{\mathbb{P}_h(|D(T)| \geq C \log u)}{\mathbb{P}_{h,\nu}(\omega(T) > u)} \\ \leq \; & 2u^{-\left(1/\left(\log q\right)-C\right)\log\left(m_h^{-1}\right)}. \end{aligned}$$

We find then that

$$\begin{aligned} & \mathbb{P}_{h,\nu,u}\Big(T \text{ is not } B\text{-bare}\Big) \\ \leq \; & \mathbb{P}_{h,\nu,u}\Big(\max_{i\in\{0,\ldots,D(T)-1\}} |V(J_i)| > B\log\log u\Big) \\ \leq \; & C(\log u)\exp\{-cB\log\log u\} + \mathbb{P}_{h,\nu,u}(D(T) > C\log u) \\ \leq \; & C(\log u)^{1-cB} + 2u^{-\left(1/\left(\log q\right)-C\right)\log\left(m_h^{-1}\right)}, \end{aligned}$$

the second inequality by virtue of Theorem 6. Note that $cB > 2$ implies that $1 - cB \leq -(c/2)B$. This yields the result. □

We also record some properties of a B-bare tree.

Lemma 4 *Let T be a B-bare tree. There exists a constant $C_1 > 0$ such that*

$$\omega(v_{\text{base}}) \geq \frac{\omega(T)}{(\log \omega(T))^{2B \log Q}}.$$

and

$$D(T) \geq \frac{\log \omega(T)}{2 \log Q},$$

provided that $\omega(T) \geq C_1$.

Proof. Let $v \in V(J_k)$ for some $k \in \{0, \ldots, D(T)-1\}$. Note that $\psi_k \in V(T)$ is the latest common ancestor of v and v_{base}. Recalling that $d(\cdot,\cdot)$ denotes the graphical distance on $V(T)$, note that

$$\omega(v) \leq Q^{d\left(\psi_k, v\right)} \omega(\psi_k) \leq Q^{B \log \log \omega(T)} \omega(\psi_k),$$

where we used that T is B-bare in the second inequality, while

$$\omega(v_{\text{base}}) \geq q^{d\left(\psi_k, v_{\text{base}}\right)} \omega(\psi_k) = q^{D(T)-k} \omega(\psi_k).$$

Thus,

$$\omega(v) \leq Q^{B \log \log \omega(T)} q^{-\left(D(T)-k\right)} \omega(v_{\text{base}}).$$

We find that

$$\begin{aligned} \omega(T) &= \sum_{v \in V(T)} \omega(v) = \sum_{j=0}^{D(T)} \sum_{v \in V\left(J_{D(T)-j}\right)} \omega(v) \\ \leq \quad & \omega(v_{\text{base}}) Q^{B \log \log \omega(T)} \sum_{k=0}^{D(T)} q^{-k} |V(J_{D(T)-k})| \leq \omega(v_{\text{base}}) Q^{2B \log \log \omega(T)}, \end{aligned}$$

the second inequality by T being B-bare and $\omega(T) > C_1$. Thus, the first statement of the lemma. Using $\omega(v_{\text{base}}) \leq Q^{D(T)}$, we obtain the second. □

4 The proof of Theorem 2

To prove Theorem 2, we will use the approximation of the weight $\omega(T)$ provided by $w_k(T)$ that was introduced in Definition 12. We will analyze this quantity by means of the product form (17). We will prove that, under the law $\mathbb{P}_{h,\nu}(\cdot | D(T) \geq k)$, the two constituents $u_k(T)$ and $v_k(T)$ are independent, and $u_k(T)$ is typically the larger. As we will shortly explain, the asymptotic decay of $u_k(T)$ may be obtained by the defective renewal theory with which

we obtained that of $\omega(v_{\text{base}})$ under $\mathbb{P}_{h,\nu}$ in the proof of Theorem 5. We will see that the term $v_k(T)$ modifies the determined decay rate only by a constant. Indeed, this is reflected in the statement of Theorem 2, where the constant d_1 is expressed as a limit of a certain average of vertex weights near the end of the tree.

We begin by computing the tail of the main term $u_k(T)$ by another use of defective renewal theory.

Lemma 5 *There exists a sequence of finite and positive constants* $\{d(k) : k \in \mathbb{N}\}$ *such that, for each* $k \in \mathbb{N}$,

$$\mathbb{P}_{h,\nu}\Big(u_k(T) \geq u \Big| D(T) \geq k\Big) \sim d(k)u^{-\chi}.$$

We have that

$$\lim_{k\to\infty} d(k) = \frac{1 - m_h}{\chi m_h \int_1^\infty y^\chi \log(y) d\nu(y)}.$$

Proof. Analogously to (15), we have that, under $\mathbb{P}_{h,\nu}(\,\cdot\,|D(T) \geq k)$,

$$\log u_k(T) = \sum_{i=1}^{D(T)-k} X_i,$$

where $\{X_i : i \in \mathbb{N}\}$ is a sequence of independent random variables, each having the law $\nu \circ \log^{-1}$. The depth $D(T)$ is independent of $\{X_i : i \in \mathbb{N}\}$ and, for every $k \in \mathbb{N}$, the following limit exists by (14):

$$\lim_{n\to\infty} m_h^{-n} \mathbb{P}_h\Big(D(T) = n + k \Big| D(T) \geq k\Big) =: \alpha_k. \tag{19}$$

as $n \to \infty$. It follows directly from (19) that

$$\lim_{k\to\infty} \alpha_k = 1 - m_h. \tag{20}$$

By the third part of Lemma 1, (and analogously to (16)),

$$\mathbb{P}_{h,\nu}\Big(\log u_k(T) > u \Big| D(T) \geq k\Big) \sim d(k)e^{-\chi u},$$

as $u \to \infty$, with $d(k) = \frac{\alpha_k}{\chi m_h \int_1^\infty y^\chi \log(y) d\nu(y)}$.

The stated convergence of $\{d(k) : k \in \mathbb{N}\}$ follows from (20). $\square$

We will now show that the two terms in the product decomposition (17) for $w_k(T)$ are independent, and that $u_k(T)$ is the dominant one:

Lemma 6 *Under the measure $\mathbb{P}_{h,\nu}(\cdot|D(T) \geq k)$, the random variables $u_k(T)$ and $v_k(T)$ are independent.*

Proof. On the space $\mathbb{P}_{h,\nu}(\,\cdot\,|D(T) \geq k)$, the random variables $u_k(T)$ and $v_k(T)$ are respectively measurable with respect to the labelling with weights of the edge-disjoint trees E_k^* and E_k that were introduced in Definition 13. The statement thus follows from Lemma 2. □

Lemma 7 *There exist $c > 0$ and $u_0 : \mathbb{N} \to (0, \infty)$ such that, for any $k \in \mathbb{N}$, and for $u > u_0(k)$,*

$$\mathbb{P}_{h,\nu}\Big(v_k(T) > u\Big|D(T) \geq k\Big) \leq \exp\{-cu\}\mathbb{P}_{h,\nu}\Big(u_k(T) > u\Big|D(T) \geq k\Big). \quad (21)$$

Proof. Note that

$$v_k(T) \leq |V(E_k)|Q^k. \quad (22)$$

Note further that

$$\begin{aligned} &\mathbb{P}_{h,\nu}\Big(|V(E_k)| \geq v\Big|D(T) \geq k\Big) = \mathbb{P}_{h,\nu}\Big(|V(T)| \geq v\Big|D(T) = k\Big) \\ \leq\ & \frac{\mathbb{P}_{h,\nu}\Big(|V(T)| \geq v\Big)}{\mathbb{P}_{h,\nu}\Big(D(T) = k\Big)} \leq C_k \exp\{-cv\} \end{aligned} \quad (23)$$

Here, the equality is due to Lemma 2, and the second inequality, to the following Lemma 8.

By (22) and (23), $\mathbb{P}_{h,\nu}(v_k(T) \geq u|D(T) \geq k) \leq \exp\{-cQ^{-k}u\}$ for sufficiently high u. By Lemma 5, we obtain the statement. □

Lemma 8 *There exists a constant $c \in (0, 1)$ such that*

$$\mathbb{P}_h\Big(|V(T)| \geq n\Big) \leq c^n$$

for each $n \in \mathbb{N}$.

Proof. It is a classical fact (see [23], Corollary 1.6) that the random variable $|V(T)|$ has the same distribution as the hitting time $U = \inf\{n \leq 1, S_n = -1\}$, where $S_n = \sum_{i=1}^n (M_i - 1)$ and the $\{M_i : i \in \mathbb{N}\}$ are independent and identically distributed random variables with distribution h. The distribution of the increments M_i has a negative mean and an exponential tail; thus, by a classical argument, the hitting time U also has an exponential tail. □

We are now ready to prove a precise decay rate for $w_k(T)$ under $\mathbb{P}_{h,\nu}$.

Lemma 9 *For each $k \in \mathbb{N}$,*

$$\begin{aligned} & \mathbb{P}_{h,\nu}\Big(\Big\{w_k(T) > u\Big\} \cap \Big\{D(T) \geq k\Big\}\Big) \\ \sim \quad & d(k)\mathbb{E}_{h,\nu}\Big(\omega(T)^{\chi}\Big|D(T) = k\Big)\mathbb{P}_{h,\nu}\Big(D(T) \geq k\Big)u^{-\chi}, \end{aligned}$$

where the constants $\{d(k) : k \in \mathbb{N}\}$ appear in the statement of Lemma 5.

Proof. It suffices to show that, for each $k \in \mathbb{N}$,

$$\mathbb{P}_{h,\nu}\Big(w_k(T) > u\Big|D(T) \geq k\Big) \sim d(k)\mathbb{E}_{h,\nu}\Big((v_k(T))^{\chi}\Big|D(T) \geq k\Big)u^{-\chi}, \tag{24}$$

since Lemma 2 demonstrates that the expectation on the right-hand-side is equal to $\mathbb{E}_{h,\nu}(\omega(T)^{\chi}|D(T) = k)$.

By Lemmas 5 and 6, (24) follows from the next lemma. □

Lemma 10 *Let U and V be independent random variables on a probability space $(\Omega, \mathbb{P})$ such that*

$$\mathbb{P}(U > u) \sim cu^{-\kappa} \tag{25}$$

for some $c > 0$ and $\kappa > 0$, and, for some $\eta > 0$,

$$\mathbb{P}(V > u) \leq u^{-\eta}\mathbb{P}(U > u) \tag{26}$$

for u sufficiently high. Assume also that $V \geq 1$. Then

$$\mathbb{P}(UV > u) \sim c\mathbb{E}(V^{\kappa})u^{-\kappa}. \tag{27}$$

Proof.

Fix $\epsilon > 0$. We choose $u_0 = u_0(\epsilon)$ such that, for $u \geq u_0$,

$$\Big|\frac{\mathbb{P}(U > u)}{cu^{-\kappa}} - 1\Big| < \epsilon. \tag{28}$$

From (25), (26) and $V \geq 1$, it follows that, for $\delta > 0$, there exists $C = C(\delta) > 0$ with $C(\delta) \to \infty$ as $\delta \to 0$ and $u_1 = u_1(\delta) > 0$ such that $u > u_1$ implies that

$$\mathbb{P}(UV > u) \leq (1 + \delta)\mathbb{P}\Big(\Big\{UV > u\Big\} \cap \Big\{V \leq u/C\Big\}\Big). \tag{29}$$

Now, choose $\delta > 0$ such that $C(\delta) > u_0(\epsilon)$ and $\delta < \epsilon$.

Let g denote the distribution function of V. In what follows, each quantity E_i is an error term, in absolute value at most ϵ_1: for $u \geq \max\{u_0, u_1\}$,

$$\begin{aligned}
& \mathbb{P}\Big(UV > u\Big) \\
= \; & (1+E_1)\mathbb{P}\Big(\Big\{UV > u\Big\} \cap \Big\{V \leq \frac{u}{C(\delta)}\Big\}\Big) \\
= \; & (1+E_1)\int_0^{\frac{u}{C(\delta)}} \mathbb{P}\Big(U > \frac{u}{v}\Big) dg(v) \\
= \; & (1+E_1)(1+E_2)cu^{-\kappa}\int_0^{\frac{u}{C(\delta)}} v^{\kappa} dg(v)
\end{aligned}$$

where, in the third equality, we used $u/v > u_0$ for $v \leq u/C(\delta)$ (which is implied by $C(\delta) > u_0$), (28), and $\delta < \epsilon$. Noting that $\lim_{u\to\infty}\int_0^{\frac{u}{C(\delta)}} v^{\kappa} dg(v) = \mathbb{E}(V^{\kappa})$, and $\mathbb{E}(V^{\kappa}) < \infty$ by (25) and (26), we obtain (27). $\square$

The next step is to show that the approximation of $\omega(T)$ by $w_k(T)$ is sufficiently good.

Lemma 11 *There exists $\epsilon(k) \in (0,\infty)$ with $\epsilon(k) \to 0$ as $k \to \infty$ and $u_0(k) \in (0,\infty)$ such that, for $u \geq u_0(k)$,*

$$\begin{aligned}
& \mathbb{P}_{h,\nu}\Big(\Big\{w_k(T) > u\Big\} \cap \Big\{D(T) \geq k\Big\}\Big) \leq \mathbb{P}_{h,\nu}\Big(\omega(T) > u\Big) \\
\leq \; & \Big(1+\epsilon(k)\Big)\mathbb{P}_{h,\nu}\Big(\Big\{w_k(T) > u\Big\} \cap \Big\{D(T) \geq k\Big\}\Big).
\end{aligned}$$

Proof. The first inequality is implied by $w_k(T) \leq \omega(T)$. Regarding the second inequality, note that there exists $c > 0$ such that, for any $k \in \mathbb{N}$, and for $u > u_0(k)$,

$$\mathbb{P}_{h,\nu,u}\Big(D(T) \leq k\Big) \leq \exp\{-cQ^{-k}u\}. \tag{30}$$

Indeed, since $\omega(T) \leq |V(T)|Q^{D(T)}$, on $\{\omega(T) > u, D(T) \leq k\}$, we have that $|V(T)| \geq uQ^{-k}$. By means of Lemma 8, we have then that

$$\mathbb{P}_{h,\nu}(\omega(T) > u, D(T) \leq k) \leq \exp\{-cuQ^{-k}\}$$

However, by $\omega(T) \geq \omega(v_{\text{base}})$ and Theorem 5, we have that $\mathbb{P}_{h,\nu}(\omega(T) > u) \geq (d_2/2)u^{-\chi}$. Hence, we have (30).

In view of Lemma 9 and (30), it is enough for the second inequality in the statement to show that, for all $\epsilon > 0$, there exists a $k_0(\epsilon)$ such that for $k > k_0$,

$$\lim_{u\to\infty} \mathbb{P}_{h,\nu,u}\Big(w_k(T) > u(1-\epsilon)\Big) = 1. \tag{31}$$

It is sufficient for (31) that

$$\lim_{u\to\infty} \mathbb{P}_{h,\nu,u}\Big(\omega(T) - w_k(T) \le \epsilon\omega(T)\Big) = 1. \tag{32}$$

To show this, we begin by noting that $\omega(T)-w_k(T) = \sum_{i=0}^{D(T)-k-1}\sum_{v\in V(J_i)}\omega(v)$.

Recall Definition 14 and so note also that

$$\omega(T) \ge \omega(v_{\text{base}}) = \omega(\psi_{D(T)-k})\omega_{\psi_{D(T)-k}}(v_{\text{base}}) \ge u_k(T)q^k.$$

Hence,

$$\frac{\omega(T) - w_k(T)}{\omega(T)} \le q^{-k}u_k(T)^{-1}\sum_{i=0}^{D(T)-k-1}\sum_{v\in V(J_i)}\omega(v). \tag{33}$$

Let $j \in \{0,\dots,D(T)-k-1\}$ and $v \in V(J_{D(T)-k-1-j})$. We claim that

$$\omega(v) \le u_k(T)Q^{\left|V\left(J_{D(T)-k-1-j}\right)\right|}q^{-j-1}. \tag{34}$$

To prove this, note that $\omega(v) = \omega(\psi_{D(T)-k-1-j})\omega_{\psi_{D(T)-k-1-j}}(v)$, while

$$u_k(T) = \omega(\psi_{D(T)-k}) = \omega(\psi_{D(T)-k-1-j})\omega_{\psi_{D(T)-k-1-j}}(\psi_{D(T)-k}).$$

Hence,

$$\frac{\omega(v)}{u_k(T)} = \frac{\omega_{\psi_{D(T)-k-1-j}}(v)}{\omega_{\psi_{D(T)-k-1-j}}(\psi_{D(T)-k})}. \tag{35}$$

Now, $\omega_{\psi_{D(T)-k-1-j}}(v) \le Q^{\left|V\left(J_{D(T)-k-1-j}\right)\right|}$, and

$$\omega_{\psi_{D(T)-k-1-j}}(\psi_{D(T)-k}) \ge q^{j+1},$$

so that (34) follows from (35).

From (34), we find that

$$\frac{1}{u_k(T)}\sum_{i=0}^{D(T)-k-1}\sum_{v\in V(J_i)}\omega(v) \le \sum_{i=0}^{D(T)-k-1} q^{-1-i}(2Q)^{\left|V\left(J_{D(T)-k-1-i}\right)\right|}.$$

In seeking to verify (32), we require a bound on the upper tail under $\mathbb{P}_{h,\nu,u}$ of the right-hand-side of the preceding inequality, which amounts, in effect, to an assertion that a tree sampled under $\mathbb{P}_{h,\nu,u}$ is typically sparse at a short distance from its end. The next lemma provides such a bound. It does not follow from the statement of Theorem 6, since the indices $D(T), D(T)-1,\dots$ of the final outgrowths of T are random. However, it is, in essence, a byproduct of the method of proof of Theorem 6, and its proof appears at the end of Appendix B.

Lemma 12 *There exists a constant $c > 0$ such that, for all $\ell \in \mathbb{N}$,*

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}\Big(\sum_{i=0}^{D(T)-\ell} (2Q)^{\left|V\left(J_{D(T)-\ell-i}\right)\right|} q^{-1-i-\ell} \geq \exp\{-c\ell\}\Big) \leq \exp\{-c\ell\}.$$

Applying Lemma 12 with $\ell = k+1$ in the second inequality below, we obtain

$$\mathbb{P}_{h,\nu,u}\Big(\frac{1}{u_k(T)} \sum_{i=0}^{D(T)-k-1} \sum_{v \in V(J_i)} \omega(v) > \exp\{-ck\} q^k, D(T) \geq k\Big) \tag{36}$$

$$\leq \mathbb{P}_{h,\nu,u}\Big(\sum_{i=0}^{D(T)-k-1} q^{-i-k} (2Q)^{\left|V\left(J_{D(T)-k-1-i}\right)\right|} > \exp\{-ck\}\Big) \leq \exp\{-ck\}.$$

In light of (33) and (36),

$$\mathbb{P}_{h,\nu,u}\Big(\omega(T) - w_k(T) > \exp\{-ck\}\omega(T), D(T) \geq k\Big) \leq \exp\{-ck\}.$$

Using (30), and letting $u \to \infty$ and taking k large enough, we obtain (32). This completes the proof. □

The following result, whose proof is trivial, is required to obtain Theorem 2 from Lemmas 9 and 11.

Lemma 13 *Let $\gamma > 0$. Let $s : [0, \infty) \to [0, \infty)$, $\{s_\epsilon : [0, \infty) \to [0, \infty), \epsilon > 0\}$ and the collection $\{c_\epsilon : \epsilon > 0\}$ of constants be such that, for all $\epsilon > 0$,*

$$s_\epsilon(u) \sim c_\epsilon u^{-\gamma}$$

and

$$s_\epsilon(u) \leq s(u) \leq (1+\epsilon) s_\epsilon(u)$$

for all $u \geq u_0(\epsilon)$ sufficiently high. Then $c = \lim_{\epsilon \downarrow 0} c_\epsilon$ exists and

$$s(u) \sim c u^{-\gamma}.$$

Proof of Theorem 2. Invoking Lemmas 9 and 11 to show that the hypotheses of Lemma 13 are satisfied, we see that

$$\lim_{k\to\infty} d(k) \mathbb{E}_{h,\nu}\Big(\omega(T)^\chi \Big| D(T) = k\Big) \mathbb{P}_h\Big(D(T) \geq k\Big)$$

exists. Recalling further the value of the limit $\lim_{k\to\infty} d(k)$ from Lemma 5, the application of Lemma 13 gives that

$$
\begin{aligned}
& u^\chi \mathbb{P}_{h,\nu}\Big(\omega(T) > u\Big) \\
\sim \quad & \frac{1-m_h}{\chi m_h \int_1^\infty y^\chi \log(y) d\nu(y)} \lim_{k\to\infty} \mathbb{E}_{h,\nu}\Big(\omega(T)^\chi \Big| D(T)=k\Big) \mathbb{P}_h\Big(D(T)\geq k\Big) \\
= \quad & \frac{1}{\chi m_h}\Big(\int_1^\infty y^\chi \log(y) d\nu(y)\Big)^{-1} \lim_{k\to\infty} \mathbb{E}_{h,\nu}\Big(\omega(T)^\chi 1_{D(T)=k}\Big)
\end{aligned}
$$

We thus obtain the formula (6) for the constant d_1 in the statement of the theorem. □

5 The proof of Theorems 3 and 4

We begin by showing that the return time of a walk from the base of a typical tree to its root is well approximated by an exponential random variable.

Lemma 14 *For any constant $B > 0$, there exist $v_0 > 0$ and $c > 0$ such that the following holds. Let T be a B-bare tree satisfying $\omega(T) > v_0$. There exists a random variable R having an exponential distribution under the law $\mathbb{P}^{v_{\text{base}}}_{T,\beta}$, such that the difference $\mathcal{E}_0 = H_\phi - R$ satisfies, for $u > (\mathbb{E}^{v_{\text{base}}}_{T,\beta} H_\phi)^{1/2}$,*

$$
\mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|\mathcal{E}_0| > u\Big) \leq \exp\Big\{ - cu^{1/40} (\mathbb{E}^{v_{\text{base}}}_{T,\beta} H_\phi)^{-1/80} \Big\}. \tag{37}
$$

Preparing for the proof of Lemma 14, we decompose the duration H_ϕ under $\mathbb{P}^{v_{\text{base}}}_{T,\beta}$ into a large and geometrically distributed number of typically short excursions from v_{base}, whose total duration is close to exponential, and a further brief interval of passage from v_{base} to ϕ. That is, we write the last-exit decompoosition

$$
H_\phi = \sum_{i=1}^{S} N_i \, + \, F, \tag{38}
$$

where S is the number of visits to v_{base} at positive times before H_ϕ, (so that S is a geometric random variable), and $\sum_{i=1}^{j} N_i$ is the time of the j-th such visit, for $j \in \{1, \ldots, S\}$; and F is the duration between the last visit of X to v_{base} before time H_ϕ and time H_ϕ itself.

Before proceeding, we require

Definition 16 *Let $\mathcal{P} = V(P_{\phi,v_{\text{base}}})$ denote the vertex-set of the path from the root to the base, as introduced in the decomposition in Definition 12. Under*

$\mathbb{P}_{T,\beta}$, *let* $\{T_i^{\mathcal{P}} : i \in \mathbb{N}\}$ *denote the successive distinct returns of* X *to* $\mathcal{P}$*: that is,* $T_1^{\mathcal{P}} = \inf\{i \in \mathbb{N} : X_i \in \mathcal{P}\}$*, with* $T_{n+1}^{\mathcal{P}} = \inf\{i > T_n^{\mathcal{P}} : X_i \in \mathcal{P}, X_i \neq X_{T_n^{\mathcal{P}}}\}$ *for* $n \geq 1$*. We abbreviate* $T_i = T_i^{\mathcal{P}}$*, and set* $X_{\mathcal{P}} : \mathbb{N} \to V(T)$ *by* $X_{\mathcal{P}}(i) = X(T_i)$*.*

A few observations are useful. Recall that we write $|v| = d(\phi, v)$ for $v \in V(T)$.

Lemma 15 *Let* T *be a* B*-bare tree.*

1. *The process* $\{|X_{\mathcal{P}}(i)| : i \geq 0\}$ *under* $\mathbb{P}_{T,\beta}^{v_{\text{base}}}$ *is a Markov chain. Under* $\mathbb{P}_{T,\beta}^{v_{\text{base}}}(\,\cdot\,|H_{v_{\text{base}}} < H_\phi)$*, the process* $\{|X_{\mathcal{P}}(i)| : 0 \leq i \leq H_{v_{\text{base}}}^{X_{\mathcal{P}}}\}$*, until the first return to* v_{base} *by* $X_{\mathcal{P}}$*, is a stopped Markov chain.*

2. *There exists* $c \in (0, 1/2)$ *such that, for all* $i \in \{1, \ldots, D(T) - 1\}$*,*

$$\mathbb{P}_{T,\beta}\Big(|X_{\mathcal{P}}(n+1)| = i + 1 \Big| |X_{\mathcal{P}}(n)| = i\Big) \geq 1/2 + c,$$

   where $n \in \mathbb{N}$ *is arbitrary. The same statement holds under the law* $\mathbb{P}_{T,\beta}^{v_{\text{base}}}(\,\cdot\,|n < H_{v_{\text{base}}}^{\mathcal{P}} < H_\phi^{\mathcal{P}})$*.*

3. *there exists* $c > 0$ *and* $D_1 > 0$ *such that, for any* $B > 0$*, and for any* B*-bare tree* T*,*

$$\mathbb{P}_{T,\beta}\Big(T_{i+1} \geq (\log \omega(T))^{D_1 B} \Big| |X_{\mathcal{P}}(i)| = j\Big) \leq \exp\{-ck\},$$

   for all $i \in \mathbb{N}$*,* $j \in \{1, \ldots, D(T) - 1\}$ *and* $k \in \mathbb{N}$*. This statement also holds under* $\mathbb{P}_{T,\beta}^{v_{\text{base}}}(\,\cdot\,|n < H_{v_{\text{base}}}^{\mathcal{P}} < H_\phi^{\mathcal{P}})$*.*

Proof. The first two assertions are trivially verified. Regarding the third, note that, if $v \in V(T) \setminus \mathcal{P}$, then the conditional distribution of $X(n+1)$ given $X(n) = v$ is the same, under the two laws $\mathbb{P}_{T,\beta}$ and $\mathbb{P}_{T,\beta}^{v_{\text{base}}}(\,\cdot\,|n < H_{v_{\text{base}}} < H_\phi)$. It clearly suffices then to show that there exists $D_1 > 0$ such that, for any $j \in \{1, \ldots, D(T) - 1\}$ and $v \in V(J_i)$, the hitting time under $\mathbb{P}_{T,\beta}$ of X on $\mathcal{P} \setminus \{\psi_i\}$ given $X(0) = v$ is at most $B \log\log \omega(T)$ with probability at least $(\log \omega(T))^{-D_1 B}$.

To see this, label the vertices of the path $P_{v,\psi_i} = (v = \varepsilon_0, \ldots, \varepsilon_{d(v,\psi_i)} = \psi_i)$. Let m_j denote the number of offspring of ε_j. Note that

$$\sum_{j=0}^{d(v,\psi_i)} m_j \leq |V(J_i)|, \tag{39}$$

since all offspring of vertices in P_{v,ψ_i} belong to $V(J_i) \cup \{\psi_{i+1}\} \setminus \{\psi_i\}$.

Under $\mathbb{P}_{T,\beta}$, given $X(n) = v$, the walk X will move along P_{v,ψ_i} to ψ_i in successive steps from time n, and then make a jump to an adjacent vertex in $\mathcal{P}$, with probability at least

$$\prod_{j=0}^{d(v,\psi_i)} \frac{1}{1+m_iQ} \geq \prod_{j=0}^{d(v,\psi_i)} (2Q)^{-2m_j} \geq (2Q)^{-2|V(J_i)|},$$

where (39) was used in the second inequality. With the choice $D_1 = 2\log(2Q)$, the statement follows, then, from T being B-bare. $\square$

The following lemma, treating the law of an excursion from v_{base}, follows directly from Lemma 15. Its proof is left to the reader. Recall that N_1 is the first return time to v_{base}.

Lemma 16 *Let T be a B-bare tree. Let $D_1 > 0$ be as in Lemma 15(iii). There exists $c > 0$ such that, for all $v > 0$,*

$$\mathbb{P}_{T,\beta}^{v_{\text{base}}}\Big(N_1 > v \Big| S \geq 1\Big) \leq \exp\Big\{ - \frac{cv}{(\log\omega(T))^{D_1B}}\Big\}.$$

The following lemma has a similar proof that is also left to the reader.

Lemma 17 *Let T be a B-bare tree. There exists a constant $C > 0$ such that*

$$\mathbb{E}_{T,\beta}^{\phi}(H_{v_{\text{base}}}) \leq CD(T)(\log\omega(T))^{D_1B}.$$

We now bound the error in estimating the total time of excursions from v_{base} by using the mean excursion time.

Lemma 18 *Let T be a B-bare tree. Write $\mu = \mathbb{E}_{T,\beta}^{v_{\text{base}}}(S)$. Set $L = \sum_{i=1}^{S} N_i - S\mathbb{E}N_1$. There exists $C_2 > 0$ such that, if $\omega(T) > C_2$ and $u \geq \mu^{1/2}$,*

$$\mathbb{P}_{T,\beta}^{v_{\text{base}}}\Big(|L| \geq u\Big) \leq \exp\Big\{ - u^{1/40}\mu^{-1/80}\Big\}.$$

Proof. It is easy to see that $\mu \geq q^{D(T)-1}$. Thus, Lemma 4(ii) permits us to assume that $\mu \geq C$, where $C > 0$ is an arbitrary constant, by increasing the constant $C_2 > 0$ as necessary.

Lemma 16 implies that, for some $C > 0$, under $\mathbb{P}_{T,\beta}^{v_{\text{base}}}(\cdot|S = 1)$, $\text{Var}(N_1) \leq C(\log\omega(T))^{2D_1B}$. Hence Theorem 3.7.1 of [10] yields the following moderate deviations estimate. For $0 \leq \epsilon < 1/2$, there exists $c > 0$ such that, for all $n \in \mathbb{N}$,

$$\mathbb{P}_{T,\beta}^{v_{\text{base}}}\Big(\Big|\sum_{i=1}^{n} N_i - n\mathbb{E}N_1\Big| \geq (\log\omega(T))^{D_1B} n^{1/2+\epsilon}\Big| S = n\Big) \leq \exp\{-cn^{2\epsilon}\}. \quad (40)$$

Note that, for any $k \in \mathbb{N}$,

$$\begin{aligned} & \mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|L| \geq k\Big) \\ = & \sum_{i=1}^{\mu^{1+\epsilon}} \mathbb{P}^{v_{\text{base}}}_{T,\beta}(|L| \geq k | S = i)\mathbb{P}(S = i) \, + \, \mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(S > \mu^{1+\epsilon}\Big). \end{aligned} \tag{41}$$

The random variable S being geometric, we have that there exists $c > 0$ such that, for any $\epsilon > 0$,

$$\mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(S > \mu^{1+\epsilon}\Big) \leq \exp\{ - c\mu^{\epsilon}\}. \tag{42}$$

For the remainder of the proof, we write $\mathbb{P}$ for a measure under which $\{N_i : i \in \mathbb{N}\}$ is an independent and identically distributed sequence of random variables, each of which has the law of N_1 under $\mathbb{P}^{v_{\text{base}}}_{T,\beta}(\,\cdot\,|S \geq 1)$.

Note that, for $\mu^{\frac{1}{20}} \leq i < \mu^{1+\epsilon}$,

$$\begin{aligned} & \mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|L| \geq (\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big| S = i\Big) \\ = & \mathbb{P}\Big(\Big|\sum_{j=1}^{i} N_j \, - \, i\mathbb{E}N_1\Big| \geq (\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big) \\ \leq & \mathbb{P}\Big(\Big|\sum_{j=1}^{i} N_j \, - \, i\mathbb{E}N_1\Big| \geq (\log\omega(T))^{D_1 B} i^{1/2+\epsilon}\Big) \\ \leq & \exp\Big\{ - ci^{2\epsilon}\Big\} \leq \exp\Big\{ - c\mu^{\frac{\epsilon}{10}}\Big\}, \end{aligned} \tag{43}$$

the second inequality by (40) and the third by $i \geq \mu^{1/20}$. Note also that, for $i < \mu^{\frac{1}{20}}$,

$$\begin{aligned} & \mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|L| \geq (\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big| S = i\Big) \\ = & \mathbb{P}\Big(\Big|\sum_{j=1}^{i} N_j \, - \, i\mathbb{E}N_1\Big| \geq (\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big) \\ \leq & \mathbb{P}\Big(\sum_{j=1}^{i} N_j \geq 2^{-1}(\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big) \\ \leq & \mu^{1/20}\exp\Big\{ - c\mu^{(1+\epsilon)(1/2+\epsilon)-1/20}\Big\}, \end{aligned} \tag{44}$$

where, in the first inequality, we used

$$i\mathbb{E}N_1 \leq C\mu^{1/20}(\log\omega(T))^{D_1 B} \leq 2^{-1}(\log\omega(T))^{D_1 B}\mu^{(1+\epsilon)(1/2+\epsilon)},$$

(which follows from Lemma 16 and the assumed $\mu \geq C$). The second inequality follows because one among the N_i in question exceeds the average requirement, along with Lemma 16 and $i \leq \mu^{1/20}$.

By (41), (43), (44) and (42),

$$\begin{aligned} & \mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|L| \geq (\log\omega(T))^{BD_1}\mu^{(1+\epsilon)(1/2+\epsilon)}\Big) \\ \leq\ & \mu^{1/10}\exp\Big\{-c\mu^{(1+\epsilon)(1/2+\epsilon)-1/20}\Big\} + \mu^{1+\epsilon}\exp\{-c\mu^{\epsilon/10}\} + \exp\{-c\mu^{\epsilon}\} \\ \leq\ & \exp\{-\mu^{\epsilon/20}\}, \end{aligned}$$

the second inequality due to $\mu \geq C$.

It follows from $\mu^{\epsilon/4} \geq (\log\omega(T))^{BD_1}$, (which is implied by $\mu \geq q^{D(T)-1}$ and Lemma 4(ii)), that

$$\mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|L| \geq \mu^{1/2+2\epsilon}\Big) \leq \exp\Big\{-\mu^{\frac{\epsilon}{20}}\Big\}.$$

Recalling that $u \geq \mu^{1/2}$, we may set $\epsilon > 0$ so that $\mu^{2\epsilon} = u\mu^{-1/2}$. In this way, we obtain the statement of the lemma. □

The geometric random variable S must now be compared with an exponential random variable.

Lemma 19 *Any geometric random variable may be coupled to an exponential random variable E in such a way that $|G - E| \leq 1$ almost surely.*

Proof. To the geometric random variable G such that $\mathbb{P}(G = i) = p^i(1-p)$ for $i \geq 0$, the exponential random variable E satisfying $\mathbb{P}(E > u) = \exp\{-\log(p^{-1})u\}$ for $u > 0$, may be coupled by setting $G = \lfloor E \rfloor$. □

Proof of Lemma 14. By Lemma 19, we may construct under $\mathbb{P}^{v_{\text{base}}}_{T,\beta}$ an exponential random variable E such that $|S - E| \leq 1$. Set $R = E\mathbb{E}^{v_{\text{base}}}_{T,\beta}(N_1|S \geq 1)$. Then R has an exponential distribution, and

$$\mathcal{E}_0 := \sum_{i=1}^{S} N_i - R$$

satisfies

$$\mathbb{P}^{v_{\text{base}}}_{T,\beta}\Big(|\mathcal{E}_0| \geq u + C(\log\omega(T))^{D_1 B}\Big) \leq \exp\Big\{-u^{1/40}\mu^{-1/80}\Big\}, \tag{45}$$

by means of $\mathbb{E}^{v_{\text{base}}}_{T,\beta}(N_1|S \geq 1) \leq C(\log\omega(T))^{BD_1}$ (which follows from Lemma 16), and Lemma 18.

Noting that $S \le H_\phi$ on $\mathbb{P}_{T,\beta}^{v_{\rm base}}$, we obtain

$$\mathbb{P}_{T,\beta}\Big(|\mathcal{E}_0| \ge u\Big) \le \exp\Big\{ - cu^{1/40}(\mathbb{E}_{T,\beta}H_\phi)^{-1/80}\Big\}$$

for $u \ge C(\log\omega(T))^{D_1B}$, whence, the statement of the lemma. $\square$

We now present a result that combines the content of Lemma 14 with an asymptotic expression for the mean return time to the root under $\mathbb{P}_{T,\beta,\mathcal{DE}}^{\phi}$. The quantity ω_* from Definition 11 appears in the statement. We mention that the the notation of Definition 14 may be employed to give a succinct expression for this quantity: $\omega_* = \sum_{v\in V(T_{v_{\rm child}})}\omega_{v_{\rm child}}(v)$.

Proposition 1 *Let $B \in (0,\infty)$ be an arbitrary constant. There exists $v_0 \in (0,\infty)$ and $C_3 > 0$, such that the following holds. Let T denote a B-bare weighted tree such that $\omega_* > v_0$.*

The distribution of H_ϕ under $\mathbb{P}_{T,\beta}^{v_{\rm child}}(\,\cdot\,|\mathcal{DE})$ is such that we may construct on this space an exponential random variable E with $\mathbb{E}_{T,\beta}^{v_{\rm child}}(E|\mathcal{DE}) = 2\omega_*/p_{\rm de}$ and with

$$\mathbb{E}_{T,\beta}^{v_{\rm child}}\Big(|H_\phi - E|\Big|\mathcal{DE}\Big) \le C_3\omega_*^{1/2}. \tag{46}$$

We further have that

$$\mathbb{E}_{T,\beta}^{v_{\rm child}}\Big(H_\phi\Big|\mathcal{DE}^c\Big) \le C\Big(\log\omega_*\Big)^{BD_1}. \tag{47}$$

Proof. We begin by showing (47). Note that, under $\mathbb{P}_{T,\beta}^{v_{\rm child}}(\,\cdot\,|\mathcal{DE}^c)$, the process $|X_P|$ (which is specified in Definition 16), has the conditional distribution of $|X_P|$ under $\mathbb{P}_{T,\beta}^{v_{\rm child}}$ given that this process hits zero before $D(T)$. The law of $|X_P|$ under $\mathbb{P}_{T,\beta}^{v_{\rm child}}$ having a uniform bias to the right (except for a reflection at $D(T)$), we see that this conditional distribution has a uniform bias to the left, so that the hitting time of 0 by this process has a finite mean, uniformly in T. Note further that the bound in Lemma 15(iii) remains valid under the law $\mathbb{P}_{T,\beta}^{v_{\rm child}}(\,\cdot\,|\mathcal{DE}^c)$, since the walk X under this measure has the unconditioned jump distribution at points in $V(T)\setminus P$. In this way, we conclude that (47) holds.

To prove (46), we now argue that

$$\mathbb{E}_{T,\beta}^{v_{\rm child}}\Big(H_{v_{\rm base}}\Big|\mathcal{DE}\Big) \le C\Big(\log\omega_*\Big)^{D_1B+1}. \tag{48}$$

Indeed, using the notation (38), under $\mathbb{P}_{T,\beta}^{v_{\rm child}}$, we have that $\mathcal{DE} = \{S \ge 1\}$. By Lemma 17, $E_{T,\beta}^{v_{\rm child}}(H_{v_{\rm base}}) \le CD(T)(\log\omega_*)^{D_1B}$. We now obtain (48) from $q^{D(T_0)-1} \le \omega_*$.

By (7), we have that

$$\begin{aligned}
& 2\omega_* - 1 = \mathbb{E}^{v_{\text{child}}}_{T,\beta}(H_\phi) \\
= \;& \mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi\Big|\mathcal{DE}\Big)\mathbb{P}^{v_{\text{child}}}_{T,\beta}\Big(\mathcal{DE}\Big) + \mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi\Big|\mathcal{DE}^c\Big)\mathbb{P}^{v_{\text{child}}}_{T,\beta}\Big(\mathcal{DE}^c\Big) \\
= \;& p_{\text{de}}\Big(\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi - H_{v_{\text{base}}}\Big|\mathcal{DE}\Big) + \mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_{v_{\text{base}}}\Big|\mathcal{DE}\Big)\Big) \\
& \qquad + (1-p_{\text{de}})\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi\Big|\mathcal{DE}^c\Big).
\end{aligned}$$

Hence,

$$\begin{aligned}
& \Big|\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi - H_{v_{\text{base}}}\Big|\mathcal{DE}\Big) \;-\; \frac{2\omega_*-1}{p_{\text{de}}}\Big| \qquad (49) \\
\leq \;& \mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_{v_{\text{base}}}\Big|\mathcal{DE}\Big) \;+\; (p_{\text{de}}^{-1}-1)\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(H_\phi\Big|\mathcal{DE}^c\Big) \leq C(\log\omega_*)^{BD_1+1},
\end{aligned}$$

where, in the second inequality, we used (48) and (47), as well as that p_{de} is bounded away from zero uniformly, (as shown by the readily verified inequality, $p_{\text{de}} \geq 1 - q^{-1}$).

By Lemma 14, under $\mathbb{P}^{v_{\text{child}}}_{T,\beta}(\,\cdot\,|\mathcal{DE})$, we may construct an exponentially distributed random variable E_0 in such a way that $H_\phi = H_{v_{\text{base}}} + E_0 + \mathcal{E}_1$, where $\mathcal{E}_1$ satisfies the bound

$$\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(|\mathcal{E}_1|\Big|\mathcal{DE}\Big) \leq C\mathbb{E}^{v_{\text{base}}}_{T,\beta}(H_\phi)^{1/2}. \qquad (50)$$

By (49) and (50),

$$\begin{aligned}
\Big|\mathbb{E}^{v_{\text{child}}}_{T,\beta}(E_0|\mathcal{DE}) - 2p_{\text{de}}^{-1}\omega_*\Big| \;&\leq\; C(\log\omega_*)^{BD_1+1} + \mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(|\mathcal{E}_1|\Big|\mathcal{DE}\Big) \\
&\leq\; C(\log\omega_*)^{BD_1+1} + C(\mathbb{E}^{v_{\text{base}}}_{T,\beta}H_\phi)^{1/2}. \qquad (51)
\end{aligned}$$

From $p_{\text{de}}\mathbb{E}^{v_{\text{base}}}_{T,\beta}(H_\phi) \leq \mathbb{E}^{v_{\text{child}}}_{T,\beta}(H_\phi)$, $p_{\text{de}} \geq 1 - q^{-1}$ and (7), we see that

$$\mathbb{E}^{v_{\text{base}}}_{T,\beta}(H_\phi) \leq C\omega_*. \qquad (52)$$

This implies that

$$\Big|\mathbb{E}^{v_{\text{base}}}_{T,\beta}(E_0|\mathcal{DE}) - 2p_{\text{de}}^{-1}\omega_*\Big| \leq C\omega_*^{1/2}, \qquad (53)$$

since ω_* is assumed to exceed a large constant v_0. Now, we let

$$E = \frac{2\omega_*}{p_{\text{de}}}\frac{E_0}{\mathbb{E}^{v_{\text{base}}}_{T,\beta}(E_0|\mathcal{DE})},$$

so that E under $\mathbb{P}^{v_{\text{child}}}_{T,\beta}(\cdot|\mathcal{DE})$ is an exponential random variable with $\mathbb{E}^{v_{\text{child}}}_{T,\beta}(E|\mathcal{DE}) = 2p_{\text{de}}^{-1}\omega_*$. Thus, under $\mathcal{DE}$, we have that

$$H_\phi = E + \mathcal{E}_2, \tag{54}$$

where $\mathcal{E}_2 = H_{v_{\text{base}}} + (E_0 - E) + \mathcal{E}_1$ satisfies

$$\mathbb{E}^{v_{\text{child}}}_{T,\beta}\Big(|\mathcal{E}_2|\Big|\mathcal{DE}\Big) \le C(\log\omega_*)^{BD_1+1} + C\omega_*^{1/2} + C\omega_*^{1/2} \le C\omega_*^{1/2},$$

the first inequality by (48), (53), (50) and (52), and the second by $\omega_* > v_0$. We have obtained (46). □

Proof of Theorems 3 and 4. Note that Theorem 3 follows directly from the two statements in Theorem 4.

Note that the distribution of H_ϕ under $\mathbb{P}^{v_{\text{base}}}_{T,\beta}$ coincides with that of $H_\phi - H_{v_{\text{base}}}$ under $\mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}$. For this reason, the first statement of Theorem 4 follows from (49) and Lemma 3.

We begin deriving the second statement of Theorem 4. Let T be any weighted tree. Clearly, under the conditional measure $\mathbb{P}^{\phi}_{T,\beta,\mathcal{DE}}$, we have that $X_1 = v_{\text{child}}$ almost surely, so that the distribution of H_ϕ under $\mathbb{P}^{\phi}_{T,\beta,\mathcal{DE}}$ coincides with that of $1+H_\phi$ under $\mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}$. For the second statement of Theorem 4, it thus suffices to show that

$$\Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}\Big)\Big(\frac{H_\phi}{2\omega_* p_{\text{de}}^{-1}} > t\Big) \to \exp\{-t\} \tag{55}$$

as $u \to \infty$. To demonstrate this, let A denote the event that T is B-bare and that $\omega_* > u^{\frac{\log q}{4\log Q}}$. Note that, by $\omega_* \ge q^{D(T)-1}$ and Lemma 4, we have that a B-bare tree T for which $\omega(T) \ge u$ satisfies $\omega_* > u^{\frac{\log q}{4\log Q}}$. By this and Lemma 3, we find that

$$\mathbb{P}_{h,\nu,u}(A) \to 1 \tag{56}$$

as $u \to \infty$. Under the law $\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}(\,\cdot\,|A)$, we construct a random variable E that, conditioning on the tree, enjoys the properties given in Proposition 1. Fixing $\epsilon > 0$, we have that

$$\Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}\Big)\Big(H_\phi > 2\omega_* p_{\text{de}}^{-1} t\Big) \le A_1 + A_2 + A_3, \tag{57}$$

where

$$A_1 = \Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}\Big)\Big(E > 2\omega_* p_{\text{de}}^{-1} t(1-\epsilon), A\Big),$$

$$A_2 = \Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\text{child}}}_{T,\beta,\mathcal{DE}}\Big)\Big(|H_\phi - E| > 2\omega_* p_{\text{de}}^{-1} t\epsilon, A\Big)$$

and $A_3 = \mathbb{P}_{h,\nu,u}(A^c)$. Note that (56) says that $A_3 \to 0$ as $u \to \infty$. Recalling that, conditionally on T, E is an exponential random variable of mean $2\omega_*/p_{\mathrm{de}}$, (56) yields that

$$\lim_{u\to\infty} A_1 = \exp\{-t(1-\epsilon)\}.$$

To bound A_2, fix a B-bare tree for which $\omega_* > u^{\frac{\log q}{4\log Q}}$. By Proposition 1 and Markov's inequality, we have that

$$\mathbb{P}^{v_{\mathrm{child}}}_{T,\beta,\mathcal{DE}}\Big(|H_\phi - E| \geq 2\omega_* p_{\mathrm{de}}^{-1}\epsilon t\Big) \leq C_3(\epsilon t)^{-1}\omega_*^{-1/2} \leq C_3(\epsilon t)^{-1} u^{-\frac{\log q}{8\log Q}}.$$

Thus, $A_2 \leq C_3(\epsilon t)^{-1} u^{-\frac{\log q}{8\log Q}}$. Substituting the obtained bounds into (57), and taking limits $u \to \infty$ followed by $\epsilon \to 0$, yields

$$\limsup_{u\to\infty} \Big(\mathbb{P}_{h,\nu,u} \times \mathbb{P}^{v_{\mathrm{child}}}_{T,\beta,\mathcal{DE}}\Big)\Big(H_\phi > 2\omega_* p_{\mathrm{de}}^{-1} t\Big) \leq \exp\{-t\}.$$

The complementary lower bound has a verbatim proof. In this way, we obtain (55). This completes the proof of Theorem 4. □

Appendices

A Coding of trees

We provide here a precise formulation of the notion of a finite rooted tree and the Galton-Watson distribution on them. This treatment has been taken essentially from [23].

We begin by defining a space of labels

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N} = \{1, 2, \dots\}$ and, by convention, $\mathbb{N}^0 = \{\emptyset\}$. Each element $u = (u_1, \dots, u_n) \in \mathcal{U}$ is thus a finite sequence of natural numbers. We denote by $|u| = n$ the generation of u. We define the concatenation uv of $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_m)$ by means of $uv = (u_1, \dots, u_n, v_1, \dots, v_m)$.

The mapping $\pi : \mathcal{U} \setminus \{\emptyset\} \to \mathcal{U}$, given by $\pi(u_1 \dots u_n) = u_1 \dots u_{n-1}$ associates to each individual its parent.

A finite rooted ordered tree T has vertex set given by a finite subset $\mathcal{U}$ such that

1. $\emptyset \in T$.

2. $u \in T \setminus \{\emptyset\}$ implies that $\pi(u) \in T$.

3. For every $u \in T$, there exists an integer $k_u(T) \geq 0$ such that, for every $j \in \mathbb{N}$, $uj \in T$ if and only if $1 \leq j \leq k_u(T)$.

The number $k_u(t)$ is interpreted as the number of offspring of u in T.

The edges of such a tree are the (unoriented) edges connecting each of its elements u (except $\emptyset$) to its parent $\pi(u)$. Such a tree T is weighted when we associate to it a function $\beta : E(T) \to (0, \infty)$.

The formal definition of the Galton-Watson law is now given.

Definition 17 *Let $h = \{h_i : i \in \mathbb{N}\}$ satsfy $\sum_{i=1}^{\infty} ih_i \leq 1$. Define a family of independent and identically distributed random variables $\{K_u : u \in \mathcal{U}\}$, each having the law h. Write θ for the subset of $\mathcal{U}$ given by*

$$\theta = \Big\{u = (u_1 \dots u_n) \in \mathcal{U} : u_j \leq K_{u_1 \dots u_{j-1}} \textit{ for each } 1 \leq j \leq n\Big\}.$$

Then θ is the vertex set of the Galton-Watson tree sampled according to $\mathbb{P}_h$.

With $Q \geq q > 1$, and for a law ν supported in $[q, Q]$, the law $\mathbb{P}_{h,\nu}$ is defined by introducing an independent collection $\{\beta_u : u \in \mathcal{U} \setminus \{\emptyset\}\}$ of ν-distributed random variables, and setting $\beta_e = \beta_u$, for each edge $e = (u, \pi(u))$, with $u \in \theta \setminus \{\emptyset\}$.

Note that $\mathbb{P}_{h,\nu}$ is a law on finite rooted weighted ordered trees. In the next appendix, we will wish to work with the unordered variant of this object. We now formally define this.

Definition 18 *Let T and T' be two finite rooted weighted ordered trees. We say that T and T' are isomorphic if there exists a bijection $\psi : V(T) \to V(T')$ with the properties that*

1. *$\psi(\emptyset) = \emptyset$.*

2. *$\psi \circ \pi = \pi \circ \psi$ on $V(T)$.*

3. *$\beta(\psi(e)) = \beta(e)$ for each $e \in E(T)$.*

In 3, we used 2 to extend the definition of ψ to $E(T)$: indeed, each edge $e \in E(T)$ taking the form $e = (\pi(u), u)$ for some $u \in V(T) \setminus \{\emptyset\}$, we set $\psi(e)$ to be the edge $(\psi \circ \pi(u), \psi(u))$.

A rooted weighted unordered tree is an isomorphism class of such trees. An unordered weighted Galton-Watson tree is the law of the isomorphism class of a sample of $\mathbb{P}_{h,\nu}$.

After Definition 7, we introduced the joint tree and walk measure $\mathbb{P}_{h,\nu} \times \mathbb{P}^v_{T,\beta}$. The definition was left imprecise, in the sense that a *selection rule* was required to specify the vertex $v \in V(T)$. We are now able to clarify this point formally.

Definition 19 *Let Θ denote the set of finite rooted weighted ordered trees. A selection rule is a map $s : \Theta \to \mathcal{U}$ with the property that $s(T) \in V(T)$ for all $T \in \Theta$. In the definition of $\mathbb{P}_{h,\nu} \times \mathbb{P}^v_{T,\beta}$, v is a selection rule.*

B The proof of Theorem 6

We aim to show that, for any $u > 0$, under the law $\mathbb{P}_{h,\nu,u}$, (which we recall denotes $\mathbb{P}_{h,\nu}(\,\cdot\,|\omega(T) > u)$), a long outgrowth is unlikely.

B.1 Surgery using the FSO-decomposition

The plan is to argue that, under the measure $\mathbb{P}_{h,\nu}$, such an outgrowth is not the most efficient means of securing the condition $\omega(T) > u$ of high weight. We will demonstrate that, from a high-weight tree with a long outgrowth, the outgrowth may be removed, and the tree lengthened a little, in such a way that the surgically altered tree has at least the weight of the original tree, with the altered tree being demonstrably more probable under $\mathbb{P}_{h,\nu}$ than the original one.

The proof that we are explaining being a little involved, we prefer to give it firstly under two assumptions, that serve to remove some distracting details in the argument.

Hypothesis 3 *The edge-weight law ν has no atoms: that is, for all $x \in (1, \infty)$, $\nu(\{x\}) = 0$. The offspring distribution $h = \{h_i : i \in \mathbb{N}\}$ satisfies $h_1 > 0$.*

The first hypothesis is being used because it ensures that the weights attached to the vertices in any finite tree are distinct. The second allows the presence in a tree of finite paths each of whose vertices has a single offspring. Regarding hypotheses, we also mention that we will make the harmless assumption that $h_i > 0$ for some $i \geq 2$: for, were this to fail, all of our trees would be finite paths, so that Theorem 6.

To define and analyse the surgical procedure, we introduce a new decomposition of a tree.

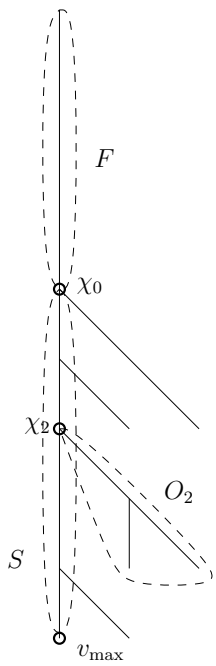

Figure 1: The FSO-decomposition is illustrated.

Definition 20 *Let T denote a weighted tree such that the maximal value of $\omega(v)$ among $v \in V(T)$ is assumed by a unique vertex that we will denote by $v_{\max}$. Recall that $\phi = \phi_T$ denotes the root of T.*

Let the first branch point v_{fbp} denote the first vertex on the path $P_{\phi,v_{\max}}$ from ϕ to $v_{\max}$ having at least two offspring. (We set $v_{\text{fbp}} = v_{\max}$ if there is no such vertex in $P_\phi, v_{\max}$.) We define the foundation F to be $P_{\phi,v_{\text{fbp}}}$. We define the spine S to be $P_{v_{\text{fbp}},v_{\max}}$.

Writing $s = |V(S)| - 1$, we further label the successive vertices in S as $v_{\text{fbp}} = \chi_0, \chi_1, \ldots, \chi_s = v_{\max}$. For $i \in \{0, \ldots, s\}$, we let O_i, the i-th offshoot of T, denote the connected component containing χ_i in the graph with vertex set $V(T)$ and edge-set $E(T) \setminus E(S)$. We also set $O_i = \emptyset$ if $i > s$. (Note that O_s is equal to the singleton graph with vertex $v_{\max}$.)

In this way, the edges of T are partitioned into the foundation F, the spine S and the various offshoots O_j, $0 \leq j \leq s-1$. We refer to the decomposition as the FSO*-decomposition.*

To obtain Theorem 6, our plan is to show that any given offshoot O_i is likely to be small under $\mathbb{P}_{h,\nu,u}$. A short additional argument will then yield the same conclusion for any given outgrowth J_i, as desired. (The offshoots and the outgrowths are each the subtrees hanging off the path connecting the

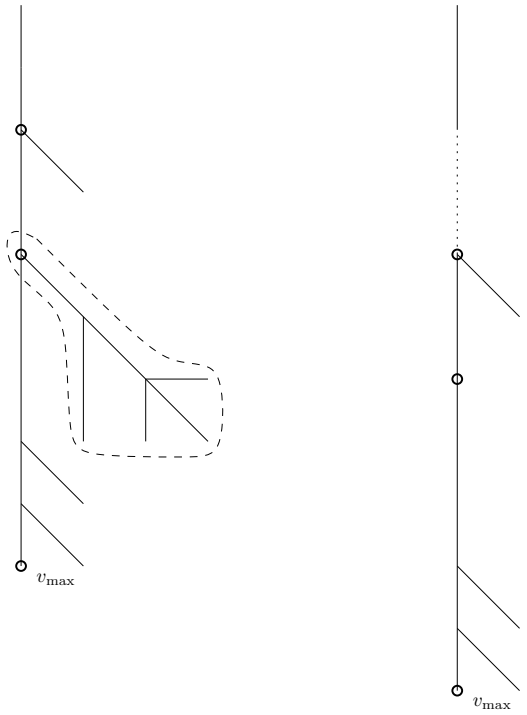

Figure 2: Before and after: a large offshoot is removed. The only trace that the procedure leaves is the unobtrusive short dotted segment shown on the right.

root to a vertex near the base of the tree, $v_{\max}$ or v_{base}. Except for a few terms at the extreme ends, the list of outgrowths coincides with that of the offshoots up to a small additive random shift in the index. In other words, the offshoots are a convenient technical tool, and it is only a small step to learn that the outgrowths are small once we know that the offshoots are.)

To show that the offshoots are small, we will fix $i \in \mathbb{N}$, and condition a sample T of $\mathbb{P}_{h,\nu}$ on all elements in its FSO-decomposition except for F and O_i, and then show that, for this law, there is a more probable means to achieve $\omega(T) > u$ than by insisting that O_i be large. This means is to insist instead that F be long (but with a length that is much shorter than the size of O_i under the comparison, so that this outcome is the more probable).

Figure 2 illustrates the surgical procedure that is the main tool in the proof of Theorem 6. The left-hand figure depicts a tree sampled under $\mathbb{P}_{h,\nu,u}$ that happens to contain a large offshoot. We now describe the surgery in a little more detail: the formal argument corresponding to this sketch will appear in the proof of Lemma 22. We will associate to the tree on the left-

hand-side, another one, shown on the right, that still satisfies the condition that the tree weight be at least u, but which is more probable under $\mathbb{P}_{h,\nu}$ than is the original tree. This latter tree will be obtained from the first by some surgery. The offending offshoot will be trimmed from the tree. This trimming necessarily entails some loss in the weight of the tree. This weight must be regained, if the outcome is to have weight at least u. The means of gaining weight is to add a few edges into the foundation of the tree. The number of edges that we need to add is only logarithmic in the size of the removed offshoot, so that the cost of this alteration to the tree is negligible compared to the cost of the original offshoot. It is in order to ensure that only a short lengthening of the foundation is sufficient to regain the lost weight that we have chosen to use the FSO-decomposition, defined in terms of the maximum weight vertex, $v_{\max}$. Indeed, the removed offshoot necessarily does not contain $v_{\max}$, so that, if this offshoot has a certain size k, the weight of the tree formed by its removal is at least $1/(k+1)$ that of the original weight. The fact that this trimmed tree necessarily has a significant weight is then vital in arguing that a slight extension to the foundation of this tree yields one with at least the weight of the original tree.

This , then, is the plan in outline. To implement it, we need a little more notation. Firstly, recall that, in Appendix A, we defined the weighted Galton-Watson law $\mathbb{P}_{h,\nu}$ as a measure on ordered weighted trees. As we will shortly explain, the ordering information specified by this definition is inconvenient for our present method of the proof. As such, we wish to consider unordered trees instead. We henceforth abuse notation and write $\mathbb{P}_{h,\nu}$ for the law on unordered weighted trees given in Definition 18.

Definition 21 *Let $s \in \mathbb{N}$ and $i \in \{0, \ldots, s-1\}$. Let S' denote a weighted path of length $|E(S')| = s$, and let O'_j, $j \in \{0, \ldots, s-1\} \setminus \{i\}$ denote a collection of (rooted) weighted trees, for which there exists a (rooted) weighted tree T whose* FSO*-decomposition satisfies $S = S'$, and $O_j = O'_j$, for such j. We will refer to the collection of data $(s, i, S', \{O'_j : j \neq i\})$ as an i-absent foliage. For brevity, we write $\mathbb{P}^{(i)}_{h,\nu}$ for the conditional distribution of $\mathbb{P}_{h,\nu}$ given that the* FSO*-decomposition of T satisfies $S = S'$ and $O_j = O'_j$ for such j. Note that implicit in the notation $\mathbb{P}^{(i)}_{h,\nu}$ is a given i-absent foliage.*

The deduction of our use of surgery, that rules out long offshoots in the FSO-decomposition, is now stated. The remainder of this subsection is devoted to establishing the result.

Proposition 2 *Assume Hypothesis 1, and that ν has support in $[q, Q] \subseteq (1, \infty)$. There exists $c > 0$ such that, for any $u > 0$, for all $i, k \in \mathbb{N}$, and for*

all i-absent foliages,

$$\mathbb{P}^{(i)}_{h,\nu}\Big(|V(O_i)| = k\Big|\omega(T) > u\Big) \leq \exp\{-ck\}.$$

Remark. From the form of this lemma, we may explain why we work with unordered trees. The statement would be untrue in the other case. For an ordered tree, each edge in the spine of the tree's FSO-decomposition would carry with it a positive integer indicating its order among the edges emanating from its parent vertex. If this number were high, then all the edges with a lower index would necessarily live in a single offshoot. This would mean that certain offshoots would necessarily be large, so that the statement of Proposition 2 would be false for ordered trees.

To prove Proposition 2, we fix $k \in \mathbb{N}$, and aim to show that, for a given i-absent foliage, it is probabilistically cheaper to obtain the condition that $\omega(T) > u$ by sampling a slightly long foundation (of length about $\log k$) than by sampling a very large offshoot O_i (of size k). In seeking to realize this aim, note the following difficulty. If $u > 0$ is large, and the spine and offshoots fixed in the i-absent foliage are small, then it will be necessary for $\omega(T) > u$ that F be long (having a length of order $\log u$) if the offshoot O_i has a size of at most k. In considering a lengthening of F as an alternative means to $|V(O_i)| = k$ for attaining $\omega(T) > u$, we are referring to the additional length of F in excess of this necessary length.

We now prepare to give a precise definition of a splitting of F into an initial "necessary" subpath and a final "optional" subpath.

Definition 22 *Note that, under the conditional law $\mathbb{P}^{(i)}_{h,\nu}$, the tree T is a random variable specified by the data F and O_i. We will sometimes denote this dependence explicitly: $T = T(F, O_i)$. Note further that, under this law, the descendent tree $T_{v_{\text{fbp}}}$ is a random variable specified by O_i. We will sometimes write $T_{v_{\text{fbp}}} = T_{v_{\text{fbp}}}(O_i)$.*

Definition 23 *Given a weighted tree S and a subset $A \subseteq V(S)$, we set $\omega^S(A) = \sum_{v \in V(S)} \omega_{\phi(S)}(v)$, where the summand is written in the notation of Definition 14. We will make choices of S and A in terms of the notation in Definition 22 that specifies tree-valued random variables under the law $\mathbb{P}^{(i)}_{h,\nu}$. For example, given the data implicit in $\mathbb{P}^{(i)}_{h,\nu}$, the quantity $\omega^{T_{v_{\text{fbp}}}(\emptyset)}\Big(T_{v_{\text{fbp}}}(\emptyset)\Big)$ denotes $\sum_{u \in V(T_{v_{\text{fbp}}})} \omega_{v_{\text{fbp}}}(u)$ when the choice $O_i = \emptyset$ is made. That is,*

$$\omega^{T_{v_{\text{fbp}}}(\emptyset)}\Big(T_{v_{\text{fbp}}}(\emptyset)\Big) = \sum_{j=0}^{s} \omega_{v_{\text{fbp}}}(\chi_j) + \sum_{j=0, j\neq i}^{s} \sum_{v \in V(O_j), v \neq \chi_j} \omega_{v_{\text{fbp}}}(v).$$

At present, we seek to prove Proposition 2 under the assumption of Hypothesis 3. In the ensuing lemmas, we assume this hypothesis, as well as the other hypotheses of Proposition 2. When we are done, we will revisit these arguments, to remove the need for Hypothesis 3.

Lemma 20 *Let T be a weighted tree whose* FSO*-decomposition has a given i-absent foliage S', O'_j, $j \in \{0, \ldots, s-1\} \setminus \{i\}$ for $s > i \geq 0$. Suppose further that $\omega(T) > u$ and that $|V(O_i)| \leq k$. Then T has the property that the inequality*

$$\frac{q}{q-1}\Big(\omega^{T_{v_{\mathrm{fbp}}}(\emptyset)}\Big(T_{v_{\mathrm{fbp}}}(\emptyset)\Big) + \omega_{v_{\mathrm{fbp}}}(\chi_i)\sum_{\ell=1}^{k} Q^{\ell}\Big)\omega(v) > u. \tag{58}$$

is satisfied by the choice $v = v_{\mathrm{fbp}}$. Note that $T_{v_{\mathrm{fbp}}}(\emptyset)$ denotes the subtree of T induced by the union of the vertices of S and of the offshoots excluding O_i.

The following definition is convenient.

Definition 24 *For $k \in \mathbb{N}$, let L_k denote the weighted tree consisting of k consecutive edges, each having bias Q. Let ℓ_k denote the weighted tree consisting of k consecutive edges, each having bias q.*

Proof of Lemma 20. Note firstly that, for such a tree T,

$$\omega(T) = \sum_{v \in V(F), v \neq v_{\mathrm{fbp}}} \omega(v) + \sum_{v \in V(T_{v_{\mathrm{fbp}}})} \omega(v).$$

Note that, for $u \in V(F)$, $\omega(u) \leq q^{-d(u, v_{\mathrm{fbp}})}\omega(v_{\mathrm{fbp}})$, since all edge have bias are at least q. Hence,

$$\sum_{v \in V(F), v \neq v_{\mathrm{fbp}}} \omega(v) \leq \frac{1}{q-1}\omega(v_{\mathrm{fbp}}) \leq \frac{1}{q-1}\sum_{v \in V(T_{v_{\mathrm{fbp}}})} \omega(v),$$

whence

$$\omega(T) \leq \Big(1 + \frac{1}{q-1}\Big)\sum_{v \in V(T_{v_{\mathrm{fbp}}})} \omega(v).$$

For $v \in V(T_{v_{\mathrm{fbp}}})$, $\omega(v) = \omega(v_{\mathrm{fbp}})\omega^{T_{v_{\mathrm{fbp}}}}(v)$, implying that

$$\omega(T) \leq \frac{q}{q-1}\omega(v_{\mathrm{fbp}})\omega^{T_{v_{\mathrm{fbp}}}}(T_{v_{\mathrm{fbp}}}). \tag{59}$$

Note that the tree L_k is the $(k+1)$-vertex tree of maximal weight. Adopting the notation of Definition 22, note thus that, among choices of $O_i = O$ for which $|V(O)| \leq k$, we have that

$$\omega^{T_{v_{\text{fbp}}}(O)}\Big(T_{v_{\text{fbp}}}(O)\Big) \leq \omega^{T_{v_{\text{fbp}}}(\emptyset)}\Big(T_{v_{\text{fbp}}}(\emptyset)\Big) + \omega^{T_{v_{\text{fbp}}}(\emptyset)}(\chi_i)\sum_{\ell=1}^{k} Q^{\ell}. \tag{60}$$

From (59), (60) and the assumption that $|V(O_i)| \leq k$, we obtain the statement of the lemma. □

Lemma 20 permits the following definition.

Definition 25 *Let T be a weighted tree satisfying the hypotheses of Lemma 20. Let v_{ess} denote the vertex v in F closest to ϕ among those satisfying (58). We refer to v_{ess} as the essential vertex. We call the path $P_{\phi,v_{\text{ess}}} \subseteq F$, the essential foundation EF. We call the path $P_{v_{\text{ess}},v_{\text{fbp}}} \subseteq F$, the optional foundation OF.*

We are aiming to prove Proposition 2, which concerns the conditional law $\mathbb{P}^{(i)}_{h,\nu}$, in which we condition on a given i-absent foliage. To do so, it turns out to be better to condition also on the essential foundation, since what is then left undetermined is the i-th offshoot and the optional foundation, and it is in terms of the length of the optional foundation that we will shortly phrase a convenient sufficient condition for $\omega(T) > u$. We now extend the $\mathbb{P}^{(i)}_{h,\nu}$ notation to incorporate this additional conditioning.

Definition 26 *Suppose given an i-absent foliage. Let $u > 0$. Under the law $\mathbb{P}^{(i)}_{h,\nu}$, we write* ESS *for the event that v_{ess} exists. (Note that this event depends implicitly on the i-absent foliage, on $k \in \mathbb{N}$ and on u.) Let EF' denote a fixed weighted tree (which is necessarily has the form of a succession of edges) which is a possible choice for EF under the law $\mathbb{P}^{(i)}_{h,\nu}(\,\cdot\,|\text{ESS})$. We call the data consisting of the given i-absent foliage, and EF', an i-absent verdant foliage. We adopt the notation $\mathbb{P}^{(i),+}_{h,\nu} = \mathbb{P}^{(i)}_{h,\nu}(\,\cdot\,|\text{ESS}, EF = EF')$. Note that implicit in the definition of $\mathbb{P}^{(i),+}_{h,\nu}$ is an i-absent verdant foliage, which itself depends implicitly on $k \in \mathbb{N}$ and $u > 0$.*

To summarise, for a sample T of $\mathbb{P}_{h,\nu}$, given the spine S and the offshoots O_j, and with O_i unspecified except for the condition that $|V(O_i)| \leq k$, it is required, for $\omega(T) > u$, that a certain initial segment of the foundation F exist. This segment is EF. What is left of F after EF is OF. By further conditioning T on EF, (that is, by using the law $\mathbb{P}^{(i),+}_{h,\nu}$), we have arrived at a convenient framework in which to pose the question 'how may $\omega(T) > u$ occur?': we will show that, under $\mathbb{P}^{(i),+}_{h,\nu}$, the optional foundation

having length $C \log k$, with C a large constant, is enough to ensure $\omega(T) > u$, and that the conditional probability of this outcome has polynomial decay in k. The $\mathbb{P}_{h,\nu}^{(i),+}$-probability of $|V(O_i)| = k$ will be shown to have exponential decay in k. Thus, a logarithmically long optional foundation offers a much more probable alternative to a k-sized i-th offshoot for realizing the event $\omega(T) > u$.

Lemma 21 *There exists a constant $c > 0$ such that, for each $\ell \in \mathbb{N}$,*

$$\mathbb{P}_{h,\nu}^{(i),+}\Big(|V(O_i)| = \ell\Big) \leq \exp\{-c\ell\}.$$

There exists a constant $C > 0$ such that

$$\mathbb{P}_{h,\nu}^{(i),+}\Big(\omega(T) > u\Big) \geq k^{-C},$$

(where recall that k is specified by $\mathbb{P}_{h,\nu}^{(i),+}$). The constants c and C may be chosen uniformly in the data that specifies $\mathbb{P}_{h,\nu}^{(i),+}$.

We will firstly establish:

Lemma 22 *Let $i \in \mathbb{N}$. There exists $k_0 \in \mathbb{N}$ such that, for any i-absent verdant foliage for which the implied k satisfies $k \geq k_0$, and for any constant $C > 1/(\log q)$, the following holds. If a tree T having the given i-absent verdant foliage satisfies $|V(OF)| \geq C \log k$, then $\omega(T) > u$.*

Remark on notation. The tree T in the statement of Lemma 22 has the form of the concatenation of EF with OF and then with $T_{v_{\text{fbp}}}$. (By concatenation, we mean the operation under which the endpoint of one path is identified with the root of a second path or tree.) The unknown quantities are O_i, which determines part of $T_{v_{\text{fbp}}}$, and OF. We write $T_{v_{\text{fbp}}} = T_{v_{\text{fbp}}}(O_i)$ and $T = T(OF, O_i)$ to indicate this dependence. This use of $T(\cdot,\cdot)$ is in conflict with that of Definition 22, and replaces it from now on.

Note that, for a given tree O, the tree $T_{v_{\text{fbp}}}(O)$ may be obtained from $T_{v_{\text{fbp}}}(\emptyset)$ by identifying $\chi_i \in V(T_{v_{\text{fbp}}}(\emptyset))$ with the root ϕ_O of O. For some choices of O, the quantity $T_{v_{\text{fbp}}}(O)$ may be ill-defined, however, since the choice $O_i = O$ may be incompatible with the given i-absent foliage. (It may be impossible for O to play the role of O_i, since this may create a vertex whose weight exceeds that of the maximum vertex of T.) In the ensuing proof, we find it convenient to work with such trees as $T_{v_{\text{fbp}}}(L_k)$, even though they may technically be illegimate. In such cases, the definition of $T_{v_{\text{fbp}}}(L_k)$ is simply taken to be as in the first sentence of this paragraph. (Recall that

the tree L_k is introduced in Definition 24.)

Proof of Lemma 22. We claim that

$$\omega^{T(\emptyset,L_k)}\Big(T_{v_{\text{fbp}}}(L_k)\Big) \geq \frac{q-1}{q}u. \tag{61}$$

To prove this, note that the expression in the large brackets in (58) is equal to

$$\sum_{v\in V\big(T_{v_{\text{fbp}}}(L_k)\big)} \omega_{v_{\text{fbp}}}(v),$$

or, equivalently, $\omega^{T_{v_{\text{fbp}}}(L_k)}(T_{v_{\text{fbp}}}(L_k))$, the weight of the descendent tree $T_{v_{\text{fbp}}}$, rooted at v_{fbp}, when the choice $O_i = L_k$ is made.

Thus,

$$\omega^{T(\emptyset,L_k)}\Big(T_{v_{\text{fbp}}}(L_k)\Big) = \omega(v_{\text{ess}}) \sum_{v\in V\big(T_{v_{\text{fbp}}}(L_k)\big)} \omega_{v_{\text{fbp}}}(v) \geq \frac{q-1}{q}u,$$

which is (61).

The equality depends on $v_{\text{ess}} = v_{\text{fbp}}$, which is true because $OF = \emptyset$. The inequality holds because (58) is verified by $v = v_{\text{ess}}$, by the definition of v_{ess}.

We now claim that

$$\omega^{T_{v_{\text{fbp}}}(\emptyset)}\Big(T_{v_{\text{fbp}}}(\emptyset)\Big) \geq \frac{1}{k+1}\omega^{T_{v_{\text{fbp}}}(L_k)}\Big(T_{v_{\text{fbp}}}(L_k)\Big). \tag{62}$$

Indeed,

$$\omega^{T_{v_{\text{fbp}}}(L_k)}\Big(T_{v_{\text{fbp}}}(L_k)\Big) - \omega^{T_{v_{\text{fbp}}}(\emptyset)}\Big(T_{v_{\text{fbp}}}(\emptyset)\Big)$$

equals $\sum_{v\in V(S_i),v\neq\chi_i} \omega_{v_{\text{fbp}}}(v)$, given that the choice $S_i = L_k$ is made.

Recall that the offshoot O_s comprises one vertex, which is $v_{\max}$. Continuing to set $S_i = L_k$, we thus have that

$$\sum_{v\in V(S_i),v\neq\chi_i} \omega_{v_{\text{fbp}}}(v) \leq \frac{|V(S_i)|-1}{|V(S_i)|}\Big(\sum_{v\in V(S_i),v\neq\chi_i} \omega_{v_{\text{fbp}}}(v) + \omega_{v_{\text{fbp}}}(v_{\max})\Big)$$
$$\leq \frac{|V(S_i)|-1}{|V(S_i)|}\omega^{T_{v_{\text{fbp}}}(L_k)}\Big(T_{v_{\text{fbp}}}(L_k)\Big),$$

the first inequality since there are $|V(S_i)|$ summands appearing inside the bracket on its right-hand-side, of which $\omega_{v_{\text{fbp}}}(v_{\max})$ is the largest. It is in deriving this bound that we invoke the defining property of $v_{\max}$.

We find that

$$\begin{aligned} & \omega^{T_{v_{\mathrm{fbp}}}(L_k)}\Big(T_{v_{\mathrm{fbp}}}(L_k)\Big) - \omega^{T_{v_{\mathrm{fbp}}}(\emptyset)}\Big(T_{v_{\mathrm{fbp}}}(\emptyset)\Big) \\ \leq \ & \frac{k}{k+1}\omega^{T_{v_{\mathrm{fbp}}}(L_k)}\Big(T_{v_{\mathrm{fbp}}}(L_k)\Big), \end{aligned}$$

which yields (62).

We also claim that, for any $m \in \mathbb{N}$,

$$\omega\Big(T(\ell_m, \emptyset)\Big) > \frac{q^m}{k+1}\frac{q-1}{q}u. \tag{63}$$

To this end, note that

$$\omega\Big(T(\ell_m, \emptyset)\Big) > \omega^{T(\ell_m,\emptyset)}(v_{\mathrm{fbp}}) \sum_{v \in V\left(T_{v_{\mathrm{fbp}}}(\emptyset)\right)} \omega_{v_{\mathrm{fbp}}}(v).$$

Note that, in the case of the tree $T(\ell_m, \emptyset)$, $\omega^{T(\ell_m,\emptyset)}(v_{\mathrm{fbp}}) = \omega(v_{\mathrm{ess}})q^m$. The inequality (58) holds with $v = v_{\mathrm{ess}}$. Hence, by (62), we have (63):

$$\begin{aligned} & \omega\Big(T(\ell_m, \emptyset)\Big) > \omega^{T(\ell_m,\emptyset)}(v_{\mathrm{fbp}})\omega^{T_{v_{\mathrm{fbp}}}(\emptyset)}\Big(T_{v_{\mathrm{fbp}}}(\emptyset)\Big) \\ \geq \ & \frac{q^m}{k+1}\omega(v_{\mathrm{ess}})\omega^{T_{v_{\mathrm{fbp}}}(L_k)}\Big(T_{v_{\mathrm{fbp}}}(L_k)\Big) \geq \frac{q^m}{k+1}\frac{q-1}{q}u, \end{aligned}$$

the third inequality by

$$\omega(v_{\mathrm{ess}})\omega^{T_{v_{\mathrm{fbp}}}(L_k)}\Big(T_{v_{\mathrm{fbp}}}(L_k)\Big) = \omega^{T(\emptyset,L_k)}\Big(T_{v_{\mathrm{fbp}}}(L_k)\Big),$$

and by (61).

By (63), $C > 1/(\log q)$ and $k \geq k_0$, we have that

$$\omega\Big(T(\ell_{\lfloor C \log k \rfloor}, \emptyset)\Big) \geq u.$$

Noting that, for each $m \in \mathbb{N}$, $\omega(T(OF, \emptyset))$ is minimized among those OF satisfying $|V(OF)| = m + 1$ by the choice $OF = \ell_m$, we complete the proof of Lemma 22. □

In the preceding lemma, we established a sufficient condition for $\omega(T) > u$ in terms of the optional foundation OF. For the measures $\mathbb{P}^{(i),+}_{h,\nu}$, we now bound from below the probability that the condition is satisfied, and find an upper bound on the probability that $|V(S_i)| = k$.

Definition 27 *Set $q_i = \frac{h_{i+1}}{1-h_0}$ for $i \geq 0$. Note that q_i is the probability of $i+1$ offspring for the offspring distribution $\{h_i : i \in \mathbb{N}\}$ conditioned on there being at least one child. Let Θ denote the law on weighted trees in which the root has a $\{q_i : i \in \mathbb{N}\}$-distributed number of offspring, all other vertices having an independent $\{h_i : i \in \mathbb{N}\}$-distributed number of offspring, and where edge-weights are independently assigned according to the law ν.*

Lemma 23 *The distribution of OF under $\mathbb{P}_{h,\nu}^{(i),+}$ coincides with the distribution of F under $\mathbb{P}_{h,\nu}$. As such, for $k \geq 1$,*

$$\mathbb{P}_{h,\nu}^{(i),+}\Big(|V(OF)| = k\Big) = h_1^{k-1}(1-h_1). \tag{64}$$

The distribution of O_i under $\mathbb{P}_{h,\nu}^{(i),+}$ is given by a sample tree R from Θ given that

$$\max_{v \in V(R)} \omega^R(v) < \omega_{\chi_i}(v_{\max}). \tag{65}$$

There exists $c > 0$ such that, for $i \in \mathbb{N}$, for all i-absent verdant foliages, and for $\ell \in \mathbb{N}$,

$$\mathbb{P}_{h,\nu}^{(i),+}\Big(|V(O_i)| = \ell\Big) \leq \exp\{-c\ell\}. \tag{66}$$

Remark. Regarding (65), note that, under the law $\mathbb{P}_{h,\nu}^{(i),+}$, $\omega_{\chi_i}(v_{\max})$ is a deterministic quantity, since it is determined by the biases of the edges in the spine S, and these form part of the data of an i-absent verdant foliage.

Proof. Under the conditional law $\mathbb{P}_{h,\nu}^{(i),+}$, the descendent tree $T_{v_{\text{ess}}}$ has the form of the law $\mathbb{P}_{h,\nu}$ conditional on $T_{v_{\text{fbp}}}$ having a collection of offshoots that are consistent with the i-absent verdant foliage specified by the choice of $\mathbb{P}_{h,\nu}^{(i),+}$. In the latter law, the length $|E(F)|$ of the foundation F does not depend the form of the data in the conditioning, so that this length has the same law as under $\mathbb{P}_{h,\nu}$ conditioned on the sample T being a path. Hence, we obtain (64), since it is clearly true for $\mathbb{P}_{h,\nu}(\,\cdot\,|\,T$ is a path).

To obtain (65), note that, under $\mathbb{P}_{h,\nu}^{(i),+}$, at least one offspring of χ_i is known to exist, this being χ_{i+1}. Any other offspring form part of $V(O_i)$, by definition. Any form of O_i is admissible, provided that it is compatible with the given i-absent verdant foliage. The operative part of this requirement is that no vertex in O_i have a weight exceeding that of the maximum vertex $v_{\max}$, whose location and weight forms part of the definition of the i-absent verdant foliage. The condition that no such vertex exist in $V(O_i)$ is (65).

The assertion (66) arises as follows. It follows directly from Lemma 8 that

$$\Theta\Big(|V(R)| = \ell\Big) \leq \exp\{-c\ell\}$$

for all $\ell \in \mathbb{N}$, (where R denotes the sample of the measure Θ). Thus, (66) follows from

$$\Theta\Big(\max_{v \in V(R)} \omega^R(v) < \omega_{\chi_i}(v_{\max})\Big) \geq c. \tag{67}$$

We now explain why (67) holds. Writing $\{\phi\}$ for the singleton graph consisting only of a root, we see that the choice $R = \{\phi\}$ implies (65), since it reduces to the statement $1 < \omega_{\chi_i}(v_{\max})$, which is true because $\chi_i \neq v_{\max}$ is implied by the definition of an i-absent foliage (in the form of the assumption that $i < s$). However, $h_1 > 0$ implies that $\Theta(R = \{\phi\}) > 0$, whence (67).

Proof of Proposition 2. We have that

$$\begin{aligned} & \mathbb{P}^{(i)}_{h,\nu}\Big(\omega(T) > u, |V(O_i)| = k\Big) \\ = \; & \mathbb{P}^{(i)}_{h,\nu}\Big(\mathrm{ESS}\Big)\mathbb{P}^{(i)}_{h,\nu}\Big(\omega(T) > u, |V(O_i)| = k\Big|\mathrm{ESS}\Big) \\ \leq \; & \mathbb{P}^{(i)}_{h,\nu}\Big(\mathrm{ESS}\Big)\exp\{-ck\}. \end{aligned}$$

The inequality here follows by conditioning on the form of EF and then applying the final assertion of Lemma 23. On the other hand,

$$\begin{aligned} & \mathbb{P}^{(i)}_{h,\nu}\Big(\omega(T) > u\Big) \geq \mathbb{P}^{(i)}_{h,\nu}\Big(\mathrm{ESS}\Big)\mathbb{P}^{(i)}_{h,\nu}\Big(\omega(T) > u\Big|\mathrm{ESS}\Big) \\ \geq \; & \mathbb{P}^{(i)}_{h,\nu}\Big(\mathrm{ESS}\Big)\mathbb{P}^{(i),+}_{h,\nu}\Big(|V(OF)| \geq C\log k\Big) \geq \mathbb{P}^{(i)}_{h,\nu}\Big(\mathrm{ESS}\Big)k^{-2C\log\left(h_1^{-1}\right)}. \end{aligned}$$

In the use of the law $\mathbb{P}^{(i),+}_{h,\nu}$, an arbitrary form for the fixed essential foundation EF' may be taken. The inequalities are due to Lemma 22 and to Lemma 23(i). By the two preceding displayed statements,

$$\mathbb{P}^{(i)}_{h,\nu}\Big(|V(O_i)| = k\Big|\omega(T) > u\Big) \leq k^{2C\log\left(h_1^{-1}\right)}\exp\{-ck\}.$$

By decreasing the value of $c > 0$, we obtain the statement of the lemma (under Hypothesis 3).

It remains to explain how the argument for Proposition 2 may be modified so that Hypothesis 3 is not invoked.

Firstly, suppose that the offspring law satisfies $h_1 = 0$. Set $\kappa = \inf\{\ell \geq 2 : h_\ell > 0\}$. (Note that $\kappa < \infty$, since we made have made this assumption from the outset to avoid trivialities.) With its present definition, the foundation F is necessarily empty. We alter the definition as follows. In a rooted tree T, define a vertex v to be standard if the path $P_{\phi,v}$ has the property that $u \in V(P_{\phi,v})$, $u \neq v$, implies that u has κ offspring; that all of the $\kappa - 1$ of these offspring not lying in $V(P_{\phi,v})$ have no offspring; and that, of the weights attached to the κ edges connecting u to its offspring, there is a unique largest,

and it is associated to the edge connecting u to its offspring in $V(P_{\phi,v})$. Note that ϕ is standard.

We now define v_{fbp} to be the final vertex in the initial sequence of standard vertices in the path $P_{\phi,v_{\max}}$. The foundation is then taken to be the connected component of ϕ in the graph with vertex set $V(T)$ and edge-set $E(T) \setminus G_{v_{\text{fbp}}}$, where $G_{v_{\text{fbp}}}$ denotes the set of edges connecting v_{fbp} to its offspring (that is, the foundation is taken to be the part of T "above" v_{fbp}).

Certain changes in the argument are forced. We comment only on the principal ones. The weighted trees L_k and ℓ_k in Definition 24 are now defined to be of depth k, with vertices at distance from the root less than k each having κ offspring. The requirement on edge-biases in the new definition of foundation ensures that an arbitrary choice of optional foundation OF is compatible with a given i-absent verdant foliage, since it implies that vertices v lying in OF but not in $P_{\phi,v_{\max}}$ have $\omega(v) < \omega(v_{\max})$. In deriving the final assertion (66) of Lemma 23, we must show (67). We may no longer consider the event that $R = \{\phi\}$, but instead consider the event $F_{\kappa-1}$ that R consists of a root with $\kappa - 1$ offspring, with these vertices each having no offspring. This event has Θ-probability at least $c > 0$. To show (67), we need to confirm that

$$F_{\kappa-1} \subseteq \Big\{\omega^R(v) < \omega_{\chi_i}(v_{\max}) \text{ for all } v \in V(R)\Big\}. \tag{68}$$

To do so, we note that, if there is no tree R of depth ℓ such that $\omega^R(v) < \omega_{\chi_i}(v_{\max})$ for all $v \in V(R)$, then (66) certainly holds, since its left-hand-side is zero. This being what we seek to show, we may assume the other case, and let T_i' denote some such tree. Provided that $\ell \geq \frac{\log Q}{\log q}$, any choice of R in the support of $F_{\kappa-1}$ has the property that each element of $V(R)$ has weight less than that of some vertex in T'. This confirms (68), so that we have (66) for all ℓ sufficiently high. We obtain the general statement by decreasing the value of $c > 0$.

We also suspend the assumption that ν have no atoms. In defining the FSO-decomposition, we must decide which of the vertices of maximal weight $v_{\max}$ should be. For ordered trees, we might choose $v_{\max}$ to be the lexicographically minimal vertex among those of maximal weight. However, as we discussed after the statement of Proposition 2, we may not work with ordered trees. Our solution is to use the lexicographical ordering of a sample of $\mathbb{P}_{h,\nu}$ merely to identify $v_{\max}$, after which, we forget about this ordering. This is the formal definition:

Definition 28 *For a finite rooted ordered weighted tree T, we write $v_{\max}$ for the lexicographically minimal element of $V(T)$ among those of maximal weight. Two such trees T and T' are said to be $v_{\max}$-isomorphic if there exists*

an isomorphism $\psi : V(T) \to V(T')$ in the sense of Definition 18 with the property that $\psi(v^T_{\max}) = v^{T'}_{\max}$.

A finite rooted unordered weighted tree with declared maximum vertex is an equivalence class of such trees under the equivalence relation of $v_{\max}$-isomorphism.

An unordered weighted Galton-Watson tree with declared maximum vertex is the law on finite rooted unordered weighted trees with declared maximum vertex given by the $v_{\max}$-isomorphism class of a sample of an ordered tree under $\mathbb{P}_{h,\nu}$.

We now suspend the notational abuse by which $\mathbb{P}_{h,\nu}$ denoted the unordered Galton-Watson law. For the ensuing argument, we will abuse notation in a different way, writing $\mathbb{P}_{h,\nu}$ for the unordered weighted Galton-Watson tree with declared maximum vertex. It is for this new choice of $\mathbb{P}_{h,\nu}$ that the lemmas in the preceding proof are to be understood, in the case that ν has no atoms.

Regarding the statement of Lemma 23, let S denote the distribution of O_i under $\mathbb{P}^{(i),+}_{h,\nu}$. We now have that, for R an unordered weighted tree in the support of Θ, the Radon-Nikodym derivative $\frac{dS}{d\Theta}$ satisfies

$$Z^{-1}\frac{dS}{d\Theta}(R) = \begin{cases} 1 & \text{if } \max_{v\in V(R)} \omega^R(v) < \omega_{\chi_i}(v_{\max}), \\ p_R & \text{if } \max_{v\in V(R)} \omega^R(v) = \omega_{\chi_i}(v_{\max}), \\ 0 & \text{otherwise,} \end{cases} \tag{69}$$

where $Z > 0$ is a normalization. The quantity p_R has the following interpretation. Consider the (unordered weighted) tree (with declared maximum vertex) arising from $\mathbb{P}^{(i),+}_{h,\nu}$ with the choice $O_i = R$ being made. Recall that this tree has no ordering, but that it does carry a declared vertex of maximal weight. Now assign a random ordering to the set of offspring of each vertex in this tree, with the uniform law. Then p_R is equal to the probability that the lexicographically minimal vertex of maximal weight, as selected by this random ordering, is equal to the declared vertex. (Note that, in the case that $\max_{v\in V(R)} \omega^R(v) = \omega_{\chi_i}(v_{\max})$, we have that $p_R < 1$, since the copy of R playing the role of O_i contains at least one vertex of maximal weight.)

In fact, the actual value that p_R takes is irrelevant: our purpose is served by verifying (66), and, given that (66) is valid for Θ, (as follows immediately from Lemma 8), it remains only to show that $Z \geq c > 0$. However,

$$Z \geq \Theta\Big(\max_{v\in V(R)} \omega^R(v) < \omega_{v_{\text{fbp}}}(v_{\max})\Big),$$

The right-hand-side was shown to be positive by the argument involving (68).Hence, $Z > c > 0$, as required for (66) in the case that neither element of Hypothesis 3 is assumed. This completes the proof of Proposition 2. □

B.2 Deriving Theorem 6 and Lemma 12 from Proposition 2

Proof of Lemma 12. Recalling that $s = |V(S)| - 1$, we label the offshoots $\{O_i : 0 \le i \le s\}$ appearing in the FSO-decomposition of a tree T in reverse order as $\hat{O}_i = O_{s-i}$, $0 \le i \le s$, setting $\hat{O}_i = \emptyset$ if $i > s$.

By averaging Proposition 2 over the distribution of the $(|V(S)| - i)$-th absent foliage under $\mathbb{P}_{h,\nu,u}$, we obtain the following consequence.

Lemma 24 *There exists $c > 0$ such that, for all $i, k \in \mathbb{N}$ and for all $u > 0$,*

$$\mathbb{P}_{h,\nu,u}\Big(|V(\hat{O}_i)| \ge k\Big) \le \exp\{-ck\}. \tag{70}$$

Lemma 24 implies that, for some $c > 0$,

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}\Big(\sum_{i=m}^{|V(S)|} (2Q)^{|V(\hat{O}_i)|} q^{-i} \ge \exp\{-cm\}\Big) \le \exp\{-cm\} \tag{71}$$

for each $m \ge 1$. Let

$$A_\ell = \Big\{d(v_{\text{base}}, v_{\max}) \le \ell/2\Big\}.$$

We observe that, on A_ℓ, the sequence $\{J_{D(T)-\ell-i} : i \ge 0\}$ coincides with $\{\hat{O}_{m+i} : i \ge 0\}$ for some $m \in \{\ell/2, \dots, \ell\}$. We now verify this, with the aid of Figure 3. The outgrowth $J_{D(T)-\ell}$, whose root is the vertex v lying at distance ℓ from v_{base}, is also the offshoot $\hat{O}_m$, where $m = d(v_{\max}, v)$. We have that $m \le \ell$, since $d(\phi, v_{\text{base}}) \ge d(\phi, v_{\max})$, while $d(v_{\text{base}}, v_{\max}) \le \ell/2$ forces $m \ge \ell/2$.

Hence, on A_ℓ,

$$\sum_{i=0}^{D(T)-\ell} (2Q)^{|V(J_{D(T)-\ell-i})|} q^{-\ell-i} \le \sum_{i=0}^{|V(S)|-\ell/2} (2Q)^{|V(\hat{O}_{i+\ell/2})|} q^{-\ell/2-i},$$

again by $d(\phi, v_{\text{base}}) \ge d(\phi, v_{\max})$. By (71), then,

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}\Big(A_\ell, \sum_{i=0}^{D(T)-\ell} (2Q)^{|V(J_{D(T)-\ell-i})|} q^{-i-\ell} \ge \exp\{-c\ell/2\}\Big) \le \exp\{-c\ell/2\}.$$

For the statement of Lemma 12, it suffices then to show that

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}(A_\ell^c) \le \exp\{-c\ell\}. \tag{72}$$

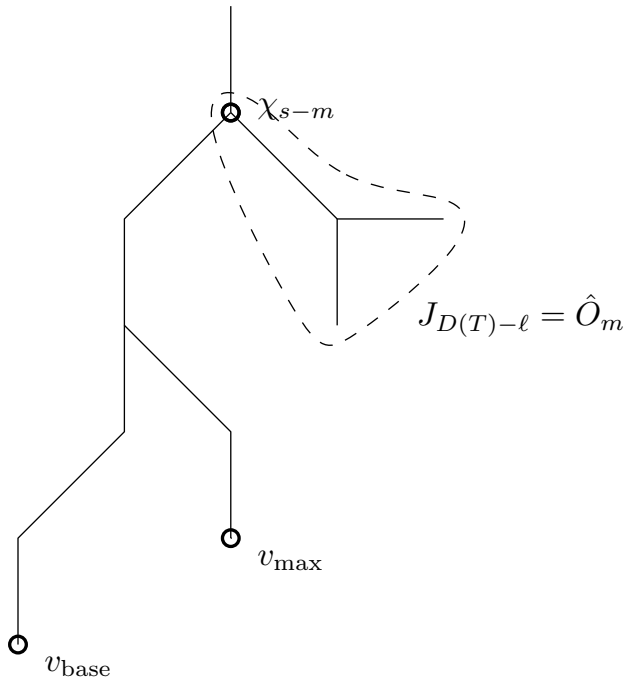

Figure 3: A tree viewed in the neighbourhood of v_{base} and $v_{\max}$.

Let v_{anc} denote the last common ancestor of v_{base} and $v_{\max}$. We argue that

$$A_\ell^c \subseteq \Big\{ \Big| V(\hat{O}_{d(v_{\text{base}}, v_{\max})}) \Big| \geq \max\Big\{ d(v_{\text{anc}}, v_{\max}), \ell/4 \Big\} \Big\}. \tag{73}$$

To see this, note that $v_{\text{base}} \in V(\hat{O}_{d(v_{\text{anc}}, v_{\max})})$, whence $|V(\hat{O}_{d(v_{\text{anc}}, v_{\max})}| \geq d(v_{\text{anc}}, v_{\text{base}})$. Now, $d(v_{\text{anc}}, v_{\text{base}})$ is at least $d(v_{\text{anc}}, v_{\max})$, since $d(\phi, v_{\text{base}}) \geq d(\phi, v_{\max})$, while it exceeds $\ell/4$ on A_ℓ^c, because then $d(v_{\max}, v_{\text{anc}}) + d(v_{\text{anc}}, v_{\text{base}}) = d(v_{\max}, v_{\text{base}}) > \ell/2$.

From (73), we learn that

$$\mathbb{P}_{h,\nu,u}(A_\ell^c) \leq \sum_{i=1}^{\ell/4} \mathbb{P}_{h,\nu,u}\Big(|V(\hat{O}_i)| \geq \ell/4 \Big) + \sum_{i=\ell/4+1}^{\infty} \mathbb{P}_{h,\nu,u}\Big(|V(\hat{O}_i)| \geq i \Big).$$

Applying Lemma 24 confirms (72) and completes the proof of Lemma 12. □
We now provide two lemmas in preparation for the proof of Theorem 6. As we have noted, Theorem 6 is trivial if the hypothesis of the next result is violated.

Lemma 25 *Assume that $h_i > 0$ for some $i \geq 2$. The foundation F in the FSO-decomposition satisfies the following bound. There exists $c > 0$ such that, for all $k \in \mathbb{N}$,*

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}\Big(|V(F)| \geq k\Big) \leq \exp\{-ck\}.$$

Proof. We consider the case that Hypothesis 3 holds, the other case requiring only minor modifications.

Consider a procedure by which the law $\mathbb{P}_{h,\nu}$ is constructed by firstly realizing the offspring of the root, and the associated biases on edges, and then iteratively selecting an as-yet-selected vertex and similarly realizing its offspring and the intervening biases (so that the new offspring are added to the list of vertices awaiting selection). The procedure stops when all vertices in the presently constructed tree have been selected; naturally, this stopping occurs in a finite number of steps for a subcritical law h. The means of choosing the next vertex to be selected may be any fixed previsible procedure. Let $Q \subseteq V(T)$ be such that $q \in Q$ if and only if, at the moment the procedure selects q, $d(\phi, q)$ is strictly greater than $d(\phi, r)$ for any already selected vertex r. Clearly, $|Q|$ is one plus the depth $D(T)$ of the sampled tree T. Note that, if the event $|V(F)| \geq k$ is to occur, it must be the case that, for each of the first $k-1$ elements of Q selected by the procedure, the number of offspring realized is equal to one. It is easy to see, however, that, for any $u > 0$ and $i \in \mathbb{N}$, under the conditional law $\mathbb{P}_{h,nu,u}(\,\cdot\,||Q| \geq i)$, the number of offspring realized for the i-th selected element of Q stochastically dominates the unconditioned offspring distribution h. This means that, under $\mathbb{P}_{h,\nu,u}$, each successive element of Q may have at least two offspring, with the probability of this being bounded below, conditionally on the construction made thus far; the probability that this fails to happen on $k-1$ consecutive occasions decays exponentially in k, whence the result. $\square$

Lemma 26 *Write $P_{\phi,v_{\max}} = [\phi = \chi'_0, \chi'_1, \ldots, \chi'_{d(\phi,v_{\max})} = v_{\max}]$. Let O'_i denote the connected component containing χ'_i in the graph with vertex set $V(T)$ and edge-set $E(T) \setminus E(P_{\phi,v_{\max}})$. Then there exists $c > 0$ such that, for all $i, k \in \mathbb{N}$,*

$$\sup_{u>0} \mathbb{P}_{h,\nu,u}\Big(|V(O'_i)| \geq k\Big) \leq \exp\{-ck\}.$$

Proof. Note that, in the event that $|V(F)| = i+1$, then $O'_j = O_{j-i}$ for $j \geq i$. Hence,

$$\Big\{|V(O'_i)| \geq k\Big\} \subseteq \bigcup_{j=i-k/2}^{i} \Big\{|V(O_j)| \geq k\Big\} \cup \Big\{|V(F)| \geq k/2\Big\}.$$

Thus,

$$\mathbb{P}_{h,\nu,u}\Big(|V(O_i')| \geq k\Big) \leq \frac{k}{2}\sup_j \mathbb{P}_{h,\nu,u}\Big(|V(O_j)| \geq k\Big) + \mathbb{P}_{h,\nu,u}\Big(|V(F)| \geq k/2\Big).$$

By Proposition 2 and Lemma 25, we obtain the statement of the lemma.

Proof of Theorem 6. Recall that v_{anc} denotes the last common ancestor of $v_{\max}$ and v_{base}. Note that $v_{\text{anc}} \in V(P_{\phi,v_{\max}})$, so that we may write $v_{\text{anc}} = \chi'_M$, with $0 \leq M \leq d(\phi, v_{\max})$.

Recall that $i \in \mathbb{N}$ denotes the index fixed in the statement of the theorem. Set $\mathcal{M}_i = \{M > i\}$.

Note that $J_j = O'_j$ for $j \leq M-1$, so that $\mathcal{M}_i \subseteq \{J_i = O'_i\}$. Note further that $\mathcal{M}_i^c \subseteq \{V(J_i) \subseteq V(T_{v_{\text{anc}}})\}$.

We have that

$$\begin{aligned} & \mathbb{P}_{h,\nu,u}\Big(|V(J_i)| \geq k\Big) \\ \leq\ & \mathbb{P}_{h,\nu,u}\Big(|V(J_i)| \geq k, \mathcal{M}_i\Big) + \mathbb{P}_{h,\nu,u}\Big(|V(J_i)| \geq k, \mathcal{M}_i^c\Big) \\ \leq\ & \mathbb{P}_{h,\nu,u}\Big(|V(O_i')| \geq k\Big) + \mathbb{P}_{h,\nu,u}\Big(|V(T_{v_{\text{anc}}})| \geq k\Big). \end{aligned} \tag{74}$$

Hence, the statement of the theorem follows from Lemma 26 and the next assertion.

There exists $c > 0$ such that, for all $u > 0$, and $k \in \mathbb{N}$,

$$\mathbb{P}_{h,\nu,u}\Big(|V(T_{v_{\text{anc}}})| \geq k\Big) \leq \exp\{-ck\}. \tag{75}$$

To prove (75), note that

$$V(T_{v_{\text{anc}}}) \subseteq \bigcup_{j=0}^{d(v_{\max}, v_{\text{base}})} V(\hat{O}_j).$$

Hence, (75) is a consequence of Lemma 24 and (72). □

C The renewal decomposition of a rooted tree

The tree decompositions that we have inroduced have been sufficient for the purposes of proving the results in this paper. In this appendix, we introduce a further decomposition of a rooted tree, whose component enjoys an attractive independence property. This "renewal decomposition" splits the tree at a sequence of points that have a natural interpretation as regeneration points.

This structure permits us to describe explicitly the conditional distribution of $\mathbb{P}_{h,\nu}(\cdot|\omega(T) > u)$, given the form of a finite initial sequence of components from the renewal decomposition. In Proposition 3, we will prove the analogue of Theorem 6 for the new decomposition. In its derivation of a stable limiting law for randomly biased walk on a supercritical Galton-Watson tree, the paper [18] will make an essential use of the renewal decomposition, and of Proposition 3.

C.1 Definition of the renewal decomposition

Definition 29 *By a root-base tree T, we refer to a finite rooted tree, one of whose vertices* base *at maximal distance from ϕ is declared to be the base.*

Given a rooted tree T, a vertex $v \in V(T)$, $v \neq \phi$, is called a cutpoint if it is not a leaf, and any other vertex in T at the same distance from ϕ as v is a leaf. The set of cutpoints naturally decompose a rooted tree into components in the following manner. We write $r(T)$ for the number of cutpoints of T plus one. We may then record these cutpoints in the form c_i, $1 \leq i \leq r(T) - 1$, in increasing order of distance from the root ϕ. We further set $c_0 = \phi$. We set $d_i = d(\phi, c_i)$ for $0 \leq i \leq r(T) - 1$. We also set $d_{r(T)} = D(T)$, where recall that $D(T) = \max\{d(\phi, v) : v \in V(T)\}$. For $1 \leq i \leq r(T)$, we define the i-th component C_i of the tree T to be the subgraph of T induced by the set of vertices in T at a distance from the root of at least d_{i-1} and at most d_i. Then, for $1 \leq i \leq r(T) - 1$, C_i may be regarded as a root-base tree, with $\phi(C_i) = c_{i-1}$ and $\mathrm{base}(C_i) = c_i$. *The final component $C_{r(T)}$, however, is a tree rooted at $c_{r(T)-1}$ that has no natural choice of base.*

In this way, we divide a rooted tree into components. We may also perform the operation in reverse, assembling a collection of such trees into a single one.

Definition 30 *Let $T_1, \ldots, T_r$ be a finite sequence of finite rooted trees, all but the last of which is root-base. We define the concatenation*

$$T_1 \circ \cdots \circ T_r$$

to be the rooted tree with vertex set $\bigcup_{i=1}^{r} V(T_i)$, in which the identifications $\mathrm{base}(T_i) = \phi(T_{i+1})$, $1 \leq i \leq r - 1$, *are made, and whose edges are inherited from the constituent trees. The root of $T_1 \circ \cdots \circ T_r$ is taken to be equal to $\phi(T_1)$.*

Remark. Note that the components $\{C_1, \ldots, C_{r(T)}\}$ of a finite rooted tree T appear as the constituents in the decomposition $T = C_1 \circ \ldots \circ C_{r(T)}$, in which identifications between successive elements are made at the cutpoints of T.

C.2 Sampling the components of a subcritical tree

To gain an understanding of the renewal decomposition, we will form a sample of $\mathbb{P}_h$ iteratively, adding one component at a time. The following definition is convenient.

Definition 31 *Let $\mathbb{P}_h^*$ denote the law of $\mathbb{P}_h$ given that the sample T is non-trivial, that is, given that $E(T)$ contains at least one element. We further write $\mathbb{P}_{h,\nu}^*$ for the law on rooted weighted trees in which the edges of a sample of $\mathbb{P}_h^*$ are independently assigned biases, each having law ν.*

Having constructed the first k components of a sample of $\mathbb{P}_h^*$, we will add the $(k+1)$-st according to the correct conditional distribution. A subtlety arises, because the final component lacks a base. Note that, for each $k \in \mathbb{N}$, under the law $\mathbb{P}_h$ conditional on $r(T) \geq k+1$ and the values of $C_1, \ldots, C_k$, the tree $T_{\mathrm{base}(C_k)}$ has law $\mathbb{P}_h^*$. The next component C_{k+1} is root-base precisely when it is not the final component, namely, with probability $\mathbb{P}_h^*(r(T) \geq 2)$, which is the conditional probability that a non-trivial tree contains a cutpoint; in this event, the distribution of C_{k+1} is that of C_1 under $\mathbb{P}_h^*(\,\cdot\,|r(T) \geq 2)$, and it has a base to which a further component will be added; otherwise, it has the law of C_1 under $\mathbb{P}_h^*(\,\cdot\,|r(T) = 1)$. We now present two lemmas that summarise these conclusions.

Lemma 27 *Let $\rho_{\mathrm{r,b}}$ denote the law on root-base trees given by C_1 under the law $\mathbb{P}_h$ conditional on T containing a cutpoint. Let ρ_{r} denote the law on rooted trees given by C_1 under $\mathbb{P}_h$ conditional on T being non-trivial but containing no cutpoint.*

Then the decomposition $T = C_1 \circ \ldots \circ C_{r(T)}$ under $\mathbb{P}_h^*$ has the form

$$T = C_1 \circ \ldots \circ C_{L+1},$$

with L an $\mathbb{N}$-valued random variable satisfying

$$\mathbb{P}(L = k) = \mathbb{P}_h^*(r(T) \geq 2)^k \mathbb{P}_h^*(r(T) = 1), \qquad k \geq 0,$$

and with $C_1, \ldots C_L$ denoting a sequence of independent root-base trees, each having law $\rho_{\mathrm{r,b}}$, while C_{L+1} is an independent rooted tree with law ρ_{r}. $\square$

Lemma 28 *Let $k \in \mathbb{N}$. Consider the law $\mathbb{P}_{h,\nu}$ conditionally on $r(T) \geq k+1$ and on an arbitrary form for the first k weighted components of T. Then the conditional distribution of the descendent tree $T_{\mathrm{base}(C_k)}$ is given by $\mathbb{P}_{h,\nu}^*$.*

Contrast Lemma 28 with the analogous result for the outgrowth splitting of a rooted tree introduced in Definition 12. If we condition on $D(T) \geq k+1$, and on the form of the first $k+1$ outgrowths $\{J_i : 0 \leq i \leq k\}$, then the conditional distribution of $T_{\psi_{k+1}}$ is dependent on the form of these initial outgrowths, because of the requirement that the vertex v_{base} lie in $T_{\psi_{k+1}}$. (The vertex v_{base} being of maximal depth, it is necessary that $T_{\psi_{k+1}}$ be deep enough to reach further from the root of T than all of the initial conditioned outgrowths.) As Lemma 28 shows, this difficulty does not arise for the renewal decomposition. For this reason, it is natural to think of the cutpoints of in the renewal decomposition as "regeneration" points for the tree. The renewal decomposition, in the form of Lemma 28, will be exploited in [18], to understand the geometry of a subcritical trap, conditional on its structure near the trap entrance.

Lemma 27 has the following corollary. Recall that we use the notation $P_{h,\nu,u} = P_{h,\nu}(\,\cdot\,|\omega(T) > u)$.

Corollary 1 *Let $k \in \mathbb{N}$ and $u > 0$. Consider the law $\mathbb{P}_{h,\nu,u}$ conditionally on $r(T) \geq k+1$ and on an arbitrary form for the first k weighted components of T. Set $u' \in \mathbb{R}$ according to*

$$u' = \frac{u - \sum \Big\{\omega(v) : v \in \bigcup_{i=1}^{k} V(C_i) \setminus \{\text{base}(C_k)\}\Big\}}{\omega(\text{base}(C_k))}.$$

If $u' > 0$, then the conditional distribution of $T_{\text{base}(C_k)}$ has the law $\mathbb{P}_{h,\nu}(\,\cdot\,|\omega(T) > u')$; if $u' \leq 0$, it has the law $\mathbb{P}^*_{h,\nu}$.

C.3 Components in a high-weight tree are small

We now present the analogue of Theorem 6 for the renewal decomposition.

Proposition 3 *For the statement, we take $C_i = \emptyset$ if $i > r(T)$ (for any rooted tree T). There exists $c > 0$ such that, for all $u > 0$ and $i \in \mathbb{N}$,*

$$\mathbb{P}_{h,\nu,u}\Big(|V(C_i)| \geq k\Big) \leq \exp\{-ck\},$$

for each $k \in \mathbb{N}$.

Our task is to show that the component sizes of a rooted weighted tree have exponential tails, uniformly under conditioning a sample T of $\mathbb{P}_{h,\nu}$ on $\omega(T) > u$. Firstly, consider Figure 4, in which two problematic trees are depicted. Recalling Definition 20, in each sketch, the vertical segment represents the path $P_{\phi,v_{\max}}$, and the slanting segments the offshoots. Each tree

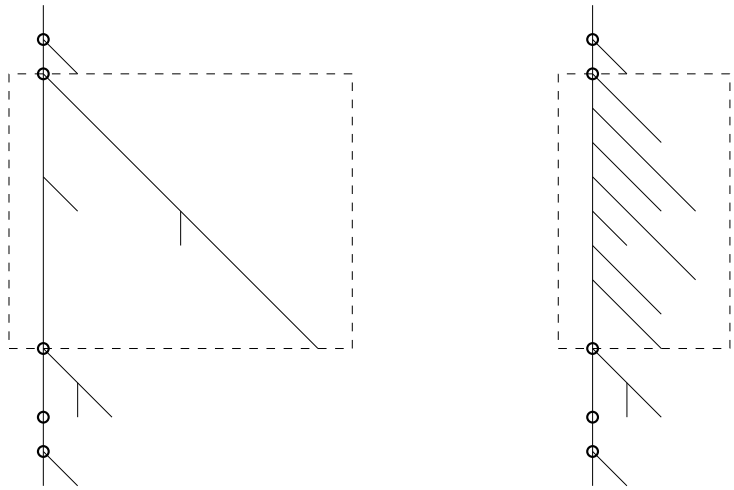

Figure 4: Two trees having a large component. The small circles indicate cutpoints.

has a large component (in its renewal decomposition), indicated in a dashed box. In the left-hand sketch, the cause of the large component is a long outgrowth, so that the scenario is excluded in essence by Proposition 2. In the right-hand sketch, the cause occurs is a succession of overlapping short outgrowths. Not only does Proposition 2 establish that a given offshoot is unlikely to be long, it does so conditionally on the lengths of the other offshoots. Note the dependence in the offshoot lengths in the second sketch in Figure 4. C.3. Proposition 2 will also be valuable in showing that the right-hand scenario is improbable, since it entails much dependence in the lengths of successive offshoots.

Alongside Proposition 2, The following result will be required.

Lemma 29 *There exists $c > 0$ such that, for any $u > 0$, and for all $i \in \mathbb{N}$,*

$$\mathbb{P}^{(i)}_{h,\nu}\Big(D(O_i) \leq 1 \Big| \omega(T) > u\Big) \geq c.$$

Proof. Consider firstly the case that $h_1 > 0$. The details of this argument are a reworking of those for Proposition 2, and we only sketch them. In Proposition 2, we established, under $\mathbb{P}^{(i)}_{h,\nu}(\cdot|\omega(T) > u)$, that O_i has probability at least c of being of size at most c^{-1}, where a small $c > 0$ may be chosen uniformly in $u > 0$ and the i-absent foliage. Given this, we arrive at the conclusion that $O_i = \emptyset$ has positive probability under the same law, by considering the event that the optional foundation be a finite length longer than it is in the case that $|V(O_i)| \leq c^{-1}$. This requirement on the optional

foundation has a bounded probabilistic cost, but it is enough to ensure that O_i may be empty with the condition $\omega(T) > u$ being satisfied.

In the case that $h_1 = 0$, it is impossible that $O_i = \emptyset$. However, we may have $D(O_i) = 1$, and the preceding argument works to show that the probability of this outcome under $\mathbb{P}^{(i)}_{h,\nu}$ is bounded away from zero, uniformly in the i-absent foliage. $\square$

The following lemma is the tool that we will use to obtain Proposition 3 from Proposition 2 and Lemma 29.

Lemma 30 *Let $\{X_i : i \in \mathbb{N}\}$ and $\{R_i : i \in \mathbb{N}\}$ denote two sequences of $\mathbb{N}$-valued random variables such that $X_i \leq R_i$ almost surely for each $i \in \mathbb{N}$, and such that there exists $c > 0$ for which, for each $n, k \in \mathbb{N}$,*

$$\mathbb{P}\Big(X_n \geq k \Big| \{X_j : 0 \leq j < n\}\Big) \leq \exp\{-ck\}, \tag{76}$$

$\sigma\{X_j : j < n\}$-a.s.,

$$\mathbb{P}\Big(X_n \in \{0,1\} \Big| \{X_j : 0 \leq j < n\}\Big) \geq c, \tag{77}$$

$\sigma\{X_j : j < n\}$-a.s, and

$$\mathbb{P}\Big(R_n \geq k \Big| \{R_j : 0 \leq j < n\}\Big) \leq \exp\{-ck\}, \tag{78}$$

$\sigma\{R_j : j < n\}$-a.s.

Set $D_n = \{n+1, \ldots, n+X_n-1\}$ if $X_n \geq 2$, and $D_n = \emptyset$ otherwise. Set $D = \bigcup_{n=0}^{\infty} D_n$, and $Y = \inf\{i \geq 1 : i \in D^c\}$. (Note that $0 \in D^c$, so that Y denotes the second smallest member of $D^c \cap \mathbb{N}$.) Further, set

$$Z = \sum_{i=0}^{Y-1} R_i.$$

Then there exists $\hat{c} > 0$, which is determined by c, such that, for each $k \in \mathbb{N}$,

$$\mathbb{P}\Big(Z \geq k\Big) \leq \exp\{-\hat{c}k\}. \tag{79}$$

Proof of Proposition 3. Note firstly that Corollary 1 implies that it suffices to prove the statement with the choice $i = 1$.

Let T be a sample of the measure $\mathbb{P}_{h,\nu,u}$. Recall from Definition 20 that, in the FSO-decomposition of T, the offshoot O_i has root χ_i. With a view to applying Lemma 30 with the choice $\mathbb{P} = \mathbb{P}_{h,\nu,u}$, we set $X_i = D(O_i)$ and $R_i = |V(O_i)|$.

Deferring for a moment the verification that the hypotheses of Lemma 30 are satisfied for this choice of sequences $\{X_i : i \geq 0\}$ and $\{R_i : i \geq 0\}$, we note the reason that the lemma is applicable: consider the set D^c appearing in Lemma 30. Then the intersection of the vertex-set $V(S)$ of the spine with the set of cutpoints of T consists precisely of those χ_i, $0 \leq i \leq |S|$, for which $i \in D^c$. As such, the quantity Y in Lemma 30 is the index i of the lowest indexed cutpoint χ_i in $V(S) \setminus \{\chi_0\}$. Let C^* denote the (renewal-decomposition) component of T whose root is χ_0. From the above, we see that

$$V(C^*) \subseteq \bigcup_{i=0}^{Y-1} V(O_i) \cup \{\chi_Y\},$$

whence

$$|V(C^*)| \leq \sum_{i=0}^{Y-1} R_i + 1, \tag{80}$$

We find from Lemma 30 and (80) that

$$\mathbb{P}_{h,\nu,u}\Big(|V(C^*)| \geq k\Big) \leq \exp\{-\hat{c}k\}, \tag{81}$$

for all $u > 0$. Note that, if the foundation F of T has at least one edge, then, necessarily, C_1 consists of a single edge (in the case that $h_1 > 0$). If $F = \emptyset$, then $C_1 = C^*$. Hence, for $k \geq 2$, and for all $u > 0$,

$$\mathbb{P}_{h,\nu,u}\Big(|V(C_1)| \geq k\Big) \leq \mathbb{P}_{h,\nu,u}\Big(|V(C^*)| \geq k, \mathbb{1}_{F=\emptyset}\Big) \leq \exp\{-\hat{c}k\}.$$

(A trivial modification is needed in the case that $h_1 = 0$.) This completes the proof, subject to checking that the hypotheses of Lemma 30 are satisfied for the present choice of sequences and for $\mathbb{P} = \mathbb{P}_{h,\nu,u}$. In checking this, we must ensure that we do so for a choice of the constant $c > 0$ in the hypotheses that is valid for all $u > 0$, since we are claiming that there exists $\hat{c} > 0$ such that (81) holds for all $u > 0$.

Note that, for each of (76), (77) and (78), the event on which we condition is determined by the data in an n-absent foliage. Note also that the bound $X_i \leq R_i$, for any $i \geq 0$, is trivial. As such, (76) and (78) are implied by Proposition 2, and (77) by Lemma 29. □

Proof of Lemma 30. We firstly show that, for some $c' > 0$,

$$\mathbb{P}\Big(Y \geq k\Big) \leq \exp\{-c'k\}. \tag{82}$$

To derive (82), it suffices to assume that $\{X_i : i \in \mathbb{N}\}$ are independent and identically distributed, with X_1 having the law $\chi = \chi_{c_0}$ for some $c_0 \in (0, 1)$,

that is determined by $c > 0$, where χ on $\{1, \ldots, \infty\}$ has the form

$$\chi(i) = (1 - c_0)c_0^{i-1}$$

for $i \geq 1$. This is because we may choose $c_0 \in (0, 1)$ such that, for each $n \in \mathbb{N}$, the conditional distribution of X_n given $\{X_i : i < n\}$ is stochastically dominated by χ, $\sigma\{X_i : i < n\}$-a.s.

Note that there exists $c_1 > 0$, determined by c, such that, for each $i \geq 1$,

$$\mathbb{P}\Big(i \in D^c\Big) \geq c_1. \tag{83}$$

Indeed, by use of (76), there exists $K \in \mathbb{N}$ such that, for all $i \in \mathbb{N}$,

$$\mathbb{P}\Big(\bigcup_{j=K+1}^{i} \{X_{i-j} > j\}\Big) \leq \frac{1}{2}.$$

Conditional on the occurrence of $\bigcap_{j=K+1}^{i} \{X_{i-j} \leq j\}$, there is, by (77), probability at least c^K of $\bigcap_{j=1}^{K} \{X_{i-j} \leq j\}$. Noting that

$$\Big\{i \in D^c\Big\} = \bigcap_{j=1}^{i} \Big\{X_{i-j} \leq j\Big\},$$

we see that, as required for (83),

$$\mathbb{P}\Big(i \in D^c\Big) \geq \frac{1}{2}c^K$$

for such a value of K.

To show (82), firstly define $y_1 = \sup D_0 + 1$ if $D_0 \neq \emptyset$ and $y_1 = 1$ otherwise. This quantity acts as a candidate for the smallest element of D^c exceeding 0. Note that, for each $n \geq 1$,

$$\begin{aligned} &\mathbb{P}\Big(y_1 \in D^c \Big| y_1 = n + 1\Big) = \mathbb{P}\Big(\bigcap_{j=1}^{n} \{X_{n+1-j} \leq j\} \Big| X_0 = n + 1\Big) \\ = \;& \mathbb{P}\Big(\bigcap_{j=1}^{n} \{X_{n-j} \leq j\}\Big) = \mathbb{P}\Big(n \in D^c\Big) \geq c_1, \end{aligned}$$

the second equality by the assumption that $\{X_i : i \in \mathbb{N}\}$ are independent and identically distributed, the inequality by (83).

If $y_1 \in D$, we seek a second candidate y_2 for the second smallest positive element of D^c. Let $z_1 \in \{1, \ldots, y_1 - 1\}$ be minimal such that $y_1 \in D_{z_1}$, that

is, such that $X_{z_1} \geq y_1 - z_1 + 1$. We set $y_2 = z_1 + X_{z_1}$ (which is $\sup D_{z_1} + 1$). Note that, from the form of the law χ, $X_{z_1} - (y_1 - z_1)$, conditionally on $y_1 \in D$, has the distribution of χ. However, conditionally on $y_1 \in D$, we have that $y_2 - y_1 = X_{z_1} - (y_1 - z_1)$, so that $y_2 - y_1$ has law χ. Note that

$$y_2 \geq \sup D_i + 1$$

for any $i \leq z_1$, while the conditional distribution of $\{X_{z_1+j} : j \geq 1\}$ is the same as the unconditioned one. Thus, for any $n \geq 2$,

$$\mathbb{P}\Big(y_2 \in D^c \Big| y_1 \in D, y_2 - z_1 = n\Big) = \mathbb{P}\Big(n - 1 \in D^c\Big) \geq c_1,$$

the inequality by (83). Note also that $y_2 > y_1 > z_1$, so that $y_2 - z_1 \geq 2$, implying that $\mathbb{P}(y_2 \in D^c | y_1 \in D) \geq c_1$.

We may iterate this procedure, forming a sequence $\{y_i : i \in \mathbb{N}\}$ such that $\{y_{i+1} - y_i : i \in \mathbb{N}\}$ is a sequence of independent terms, each having law χ, with each y_i, $i \in \mathbb{N}$, having probability at least c_1 of belonging to D^c, conditional on all previous terms belonging to D, and on any specific values for these previous terms. As such, we have found that Y is stochastically dominated by a sum of independent geometric random variables, the number of summands being an independent geometric random variable, whence (82). Note further that c' is determined by c_0 and c_1, and, thus, by c.

Turning now to (79), note that, for any $c_2 > 0$,

$$\Big\{Z \geq k\Big\} \subseteq \Big\{\sum_{i=0}^{c_2 k} R_i > k\Big\} \cup \Big\{Y \geq c_2 k\Big\},$$

so that

$$\mathbb{P}(Z \geq k) \leq \mathbb{P}\Big(\sum_{i=0}^{c_2 k} R_i > k\Big) + \exp\{-c' c_2 k\}, \tag{84}$$

by (82). In estimating $\mathbb{P}(\sum_{i=0}^{c_2 k} R_i > k)$, we may assume, due to(78), that $\{R_i : i \in \mathbb{N}\}$ is an independent and identically distributed sequence of random variables whose law has an exponentially decaying tail, with constant in the exponential determined by c. Choosing $c_2 < \frac{1}{\mathbb{E}R_1}$, we obtain

$$\mathbb{P}\Big(\sum_{i=0}^{c_2 k} R_i > k\Big) \leq \exp\{-c_3 c_2 k\}. \tag{85}$$

by means of the elementary large deviations bound

$$\mathbb{P}\Big(\sum_{i=1}^{m} R_i > (\mathbb{E}R_1 + \epsilon)n\Big) \leq \{-\psi(\epsilon)m\},$$

(with $\psi(\epsilon) > 0$ for $\epsilon > 0$). Applying (85) to (84), we obtain (79). This completes the proof of Lemma 30. □

References

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

[2] Martin Barlow. Random walks on graphs: a brief introduction. *http://www.math.ubc.ca/ barlow/cornell/cnotes.pdf.*

[3] Gérard Ben Arous, Alexander Fribergh, Nina Gantert, and Alan Hammond. Biased random walks on Galton-Watson trees with leaves. *Annals of Probability*, to appear.

[4] Gérard Ben Arous and Jiří Černý. Dynamics of trap models. In *Mathematical statistical physics*, pages 331–394. Elsevier B. V., Amsterdam, 2006.

[5] Gérard Ben Arous and Jiří Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.*, 35(6):2356–2384, 2007.

[6] Noam Berger, Nina Gantert, and Yuval Peres. The speed of biased random walk on percolation clusters. *Probab. Theory Related Fields*, 126(2):221–242, 2003.

[7] H. Bottger and V. Bryksin. Hopping conductivity in ordered and disordered systems (iii). *Physica Status Solidi (b)*, 113(1):9–49, 1982.

[8] J.P. Bouchaud. Weak ergodicity breaking and aging in disordered-systems. *J. Phys. (France) I*, (2):1705–1713, 1992.

[9] Ashok K. Chandra, Prabhakar Raghavan, Walter L. Ruzzo, Roman Smolensky, and Prasoon Tiwari. The electrical resistance of a graph captures its commute and cover times. *Comput. Complexity*, 6(4):312–340, 1996/97.

[10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[11] D. Dhar. Diffusion and drift on percolation networks in an external field. *Journal of Physics A*, 17(5):257–259, 1984.

[12] D. Dhar and Stauffer. Drift and trapping in biased diffusion on disordered lattices. *Int. J. Mod. Phys. C*, 9(2):349–355, 1998.

[13] Nathanaël Enriquez, Christophe Sabot, and Olivier Zindy. Limit laws for transient random walks in random environment on $\mathbb{Z}$. *Ann. Inst. Fourier (Grenoble)*, 59(6):2469–2508, 2009.

[14] Nathanaël Enriquez, Christophe Sabot, and Olivier Zindy. A probabilistic representation of constants in Kesten's renewal theorem. *Probab. Theory Related Fields*, 144(3-4):581–613, 2009.

[15] Gabriel Faraud, Yueyun Hu, and Zan Shi. Almost sure convergence for stochastically biased random walks on trees. *arXiv:1003.5505.*

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

[17] Alexander Fribergh and Alan Hammond. Phase transition for the speed of the anisotropic walk on the supercritical percolation cluster. *In preparation.*

[18] Alan Hammond. Stable limit laws for randomly biased walks on supercritical trees. *In preparation.*

[19] Shlomo Havlin and Daniel Ben-Avraham. *Diffusion and reactions in fractals and disordered systems.* Cambridge University Press, Cambridge, 2000.

[20] C. R. Heathcote, E. Seneta, and D. Vere-Jones. A refinement of two theorems in the theory of branching processes. *Teor. Verojatnost. i Primenen.*, 12:341–346, 1967.

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

[22] Takashi Kumagai. Random walks on disordered media and their scaling limits. *to appear in the lectures notes of the 40th probability summer school in St-Flour*, 2010.

[23] Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311 (electronic), 2005.

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

[25] Russell Lyons, Robin Pemantle, and Yuval Peres. Biased random walks on Galton-Watson trees. *Probab. Theory Related Fields*, 106(2):249–264, 1996.

[26] Russell Lyons, Robin Pemantle, and Yuval Peres. Unsolved problems concerning random walks on trees. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 223–237. Springer, New York, 1997.

[27] Dietrich Stauffer and Didier Sornette. Log-periodic oscillations for biased diffusion on random lattice. *Physica A*, 252(3-4):271–277, 1998.

[28] Alain-Sol Sznitman. On the anisotropic walk on the supercritical percolation cluster. *Comm. Math. Phys.*, 240(1-2):123–148, 2003.

[29] Alain-Sol Sznitman. Random motions in random media. In *Mathematical statistical physics*, pages 219–242. Elsevier B. V., Amsterdam, 2006.

[30] Ofer Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.