# General random walk in a random environment defined on Galton–Watson trees

A. D. Barbour[*] and Andrea Collevecchio[†]

Universität Zürich and Monash University

**Abstract.** We consider a particle performing a random walk on a Galton–Watson tree, when the probabilities of jumping from a vertex to any one of its neighbours is determined by a random process. We introduce a method for deriving conditions under which the walk is either transient or recurrent. We first suppose that the weights are i.i.d., and re-prove a result of Lyons & Pemantle [10]. We then assume a Markovian environment along each line of descent, and finally consider a random walk in a Markovian environment that itself changes the environment. Our approach involves studying the typical behaviour of the walk on fixed lines of descent, which we then show determines the behaviour of the process on the whole tree.

**Keywords:** Random walk in random environment, Galton–Watson, reinforcement.

## 1 Introduction

We consider the behaviour of a random walk in a random environment, which consists of a randomly sampled Galton–Watson tree, with the jump probabilities at each vertex

---

being prescribed by a further random mechanism. We derive conditions on the environment under which the walk is transient — that is, the event that the walk never returns to the root has positive probability — and under which it is recurrent, when the probability of returning to the root is 1. Our approach, which has its roots in that of [1], involves studying the typical behaviour of the process on fixed lines of descent, which we then show determines the behaviour of the process on the whole tree. We combine these ideas with suitable large deviation principles, and an analysis of the resulting variational formula, enabling rather satisfactory results to be obtained under relatively weak conditions.

The Galton–Watson tree is sampled first, starting from a root $\varrho$. Given the tree, positive weights are then assigned at random to its edges, and the jump probabilities are determined from the weights. In Section 2, we suppose that the weights are i.i.d., and recover a result of Lyons & Pemantle [10], in Theorem 2.1. A different proof of the Lyons–Pemantle theorem is contained in [12]; see also [7] for the multitype Galton-Watson case. In Section 3, we extend the argument to an environment in which the values of the weights evolve as a Markov chain along rays, giving sufficient conditions for both transience and recurrence in Theorem 3.6. Finally, in Section 4, we illustrate the power of our method by considering a random walk in a Markovian environment that itself changes the environment; see Theorems 4.1 and 4.4. The results of this section should be compared to the behaviour of once-reinforced random walk; see, for example, [1], [2] or [6], and the general survey of reinforcement in [15]. The model that we discuss is a strong generalization of the once-reinforced walk, and it exhibits multiple phases (see Theorem 4.4), whereas the once-reinforced walk on the supercritical Galton–Watson tree is always transient (see [1] or [2]). This should be compared to results obtained in a recent preprint by Kious & Sidoravicius [9].

Let $\mathcal{G}$ be an infinite tree with root $\varrho$. We augment $\mathcal{G}$ by adjoining a parent $\varrho^{-1}$ to the root $\varrho$. If two vertices $\nu$ and $\mu$ are the endpoints of the same edge, they are said to be neighbours, and this property is denoted by $\nu \sim \mu$. The distance $|\nu - \mu|$ between any pair of vertices $\nu, \mu$, not necessarily adjacent, is the number of edges in the unique self-avoiding path connecting $\nu$ to $\mu$. We set $|\varrho^{-1}| = -1$. For any other vertex $\nu$, we let $|\nu|$ be the distance of $\nu$ from the root $\varrho$. We denote by $b(\nu)$ the number of neighbors of $\nu$ at level $|\nu| + 1$, its offspring number, and we use $\nu^{-1}$ to denote the parent of $\nu$. We write $\nu < \mu$ if $\nu$ is an ancestor of $\mu$.

For $\nu$ a vertex of $\mathcal{G}$, we write

$$\mathbf{A}_\nu \;=\; (A_{\nu 1}, A_{\nu 2}, \ldots)$$

to denote the (finite, positive) weights on the edges between $\nu$ and its offspring. For simplicity, we index the weight associated to edge $e$ by the endpoint of $e$ with larger distance from $\varrho$. The environment $\omega$ for the random walk on the tree is then defined, for any vertex $\nu$ with offspring $\nu i$, $1 \le i \le b(\nu)$, by the probabilities

$$\omega(\nu, \nu i) \;:=\; \frac{A_{\nu i}}{1 + \sum_{1 \le j \le b(\nu)} A_{\nu j}}; \qquad \omega(\nu, \nu^{-1}) \;:=\; \frac{1}{1 + \sum_{1 \le j \le b(\nu)} A_{\nu j}}. \qquad (1.1)$$

We set $\omega(\nu, \mu) = 0$ if $\mu$ and $\nu$ are not neighbours. Given the environment $\omega$, we define the random walk $\mathbf{X} = \{X_n,\, n \ge 0\}$ that starts at $\varrho$ to be the Markov chain with $\mathbf{P}_\omega(X_0 = \varrho) = 1$, having transition probabilities

$$\mathbf{P}_\omega(X_{n+1} = \mu_1 \mid X_n = \mu_0) \;=\; \omega(\mu_0, \mu_1).$$

Moreover, we assume that $\varrho^{-1}$ is an absorbing state for the walk. The environment is random in two respects. First, the Galton–Watson tree $\mathcal{G}$ is realized; then, for each vertex $\nu \in \mathcal{G}$, the weights $\mathbf{A}_\nu$ are realized. The combined probability measure from which the environment is realized is denoted by $\mathbb{P}$ and its expectation by $\mathbb{E}$, and the semi-direct product $\mathbf{P} := \mathbb{P} \times \mathbf{P}_\omega$ represents the annealed measure. The details of the probability measures used to construct the environment are given in the subsequent sections.

We use $[\nu, +\infty)$ to denote a generic infinite line of descent from $\nu$.

## 2 Random walks in i.i.d. environment.

In this section, we assume that $\mathcal{G}$ is a Galton–Watson tree with offspring mean $b > 1$. Given the realization of the tree, we assume that the sets of weights $(\mathbf{A}_\nu,\, \nu \in \mathcal{G})$ are independent, and that, for each $\nu$, the weights $(A_{\nu i},\, 1 \le i \le b(\nu))$ are exchangeable, with the distributions of the $A_{\nu 1}$, $\nu \in \mathcal{G}$, all identical. Under these assumptions, we prove the following theorem, first given by Lyons & Pemantle [10], as part of a sharp result.

**Theorem 2.1** *Assume that $\mathcal{G}$ and the environment are distributed as above. If $\inf_{\lambda \in [0,1]} \mathbb{E}[A_{\varrho 1}^{\lambda}] > b^{-1}$, then $\mathbf{X}$ is transient; that is, with positive probability, $\mathbf{X}$ does not hit $\varrho^{-1}$.*

Our proof relies on the Mogulskii large deviations principle.

We assume that $\mathbf{X}$ is recurrent and find a contradiction. We consider the behaviour of the random walk $\mathbf{X}$ observed along any infinite line of descent $\sigma = [\varrho^{-1}, \infty)$, if one exists. Such lines exist with positive probability, since $b > 1$. We call this restricted process $\mathbf{X}^{(\sigma)}$. Note that, by our assumption of recurrence, the process $\mathbf{X}^{(\sigma)}$ has the following transition probabilities:

$$\mathbf{P}_{\omega}[X_{n+1}^{(\sigma)} = \sigma_{r+1} \,|\, X_n^{(\sigma)} = \sigma_r] \;=\; \frac{A_{\sigma_{r+1}}}{1 + A_{\sigma_{r+1}}}; \quad \mathbf{P}_{\omega}[X_{n+1}^{(\sigma)} = \sigma_{r-1} \,|\, X_n^{(\sigma)} = \sigma_r] \;=\; \frac{1}{1 + A_{\sigma_{r+1}}},$$

where we denote the successive vertices in $\sigma$ by $\sigma_j$, $j \geq -1$, with $\sigma_0 := \varrho$ and $\sigma_{-1} := \varrho^{-1}$. We define $T_{-1}$ to be the first time $\mathbf{X}^{(\sigma)}$ hits $\varrho^{-1}$, and $T_n$ the first time the process hits $\sigma_n$. Note that the $\mathbf{P}$-distributions of $T_{-1}$ and $T_n$ are not affected by the choice of $\sigma$.

**Proposition 2.2** *If*

$$\limsup_{n \to \infty} \frac{1}{n} \ln \mathbf{P}(T_{-1} > T_n) \;>\; -\ln b, \tag{2.1}$$

*then $\mathbf{X}$ is transient.*

**Proof.** We mimic the proof in [1]. Assume that $\mathbf{X}$ is recurrent. By assumption, there exists an $n^*$ such that $b^{n^*} \mathbf{P}(T_{-1} > T_{n^*}) > 1$. We now construct a branching process as follows. Set $\tau := \inf\{i > 0 \colon X_i = \varrho^{-1}\}$. We color green the vertices $\nu$ at level $n^*$ which are visited before time $\tau$. Define

$$S_{\nu} = \inf\{n \geq 0 \colon X_n = \nu\}.$$

Under our assumptions, $S_{\nu} < \infty$ a.s. for each $\nu$. A vertex $\nu$ at level $jn^*$, for some integer $j \geq 2$, is colored green, if its ancestor $\mu$ at level $(j-1)n^*$ is green, and $(\mathbf{X}_j, j \geq S_{\mu})$ hits $\nu$ before it returns to $\mu^{-1}$. The green vertices evolve as a Galton–Watson tree, with offspring mean $b^{n^*} \mathbf{P}(T_{-1} > T_{n^*}) > 1$. Hence this random tree is supercritical, and thus the probability of there being an infinite number of green vertices is positive. But this contradicts the assumption that $\mathbf{X}$ is recurrent. ∎

4

**Proof of Theorem 2.1.** In view of Proposition 2.2, it is enough to show that (2.1) is satisfied. We use a well-known formula for the hitting probability for random walk in random environment (see, for example, Sznitman [16], Equation 44),

$$\mathbf{P}(T_n < T_{-1}) \;=\; \mathbb{E}\Big[\big(\sum_{r=0}^{n}\prod_{j=1}^{r} A_{\sigma_j}^{-1}\big)^{-1}\Big].$$

Denote by $\lfloor x \rfloor$ the integer part of $x$. Then it follows directly, because

$$\Big(n\max_{r\leq n}\prod_{j=1}^{r} A_{\sigma_j}^{-1}\Big)^{-1} \leq \Big(\sum_{r=0}^{n}\prod_{j=1}^{r} A_{\sigma_j}^{-1}\Big)^{-1} \leq \Big(\max_{r\leq n}\prod_{j=1}^{r} A_{\sigma_j}^{-1}\Big)^{-1},$$

that

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n}\ln \mathbf{P}(T_n < T_{-1}) &= \liminf_{n\to\infty}\frac{1}{n}\ln \mathbb{E}\Big[\big(\sum_{r=0}^{n}\prod_{j=1}^{r} A_{\sigma_j}^{-1}\big)^{-1}\Big] \\
&= \liminf_{n\to\infty}\frac{1}{n}\ln \mathbb{E}\Big[\min_{r\leq n}\prod_{j=1}^{r} A_{\sigma_j}\Big] \\
&= \liminf_{n\to\infty}\frac{1}{n}\ln \mathbb{E}\Big[e^{\min_{r\leq n}\sum_{j=1}^{r}\ln A_{\sigma_j}}\Big] \\
&\geq \liminf_{n\to\infty}\frac{1}{n}\ln \mathbb{E}\Big[e^{\min_{t\in[0,1]}\sum_{j=1}^{\lfloor nt\rfloor}\ln A_{\sigma_j}}\Big].
\end{aligned} \tag{2.2}$$

Denote by $D[0,1]$ the space of functions $f\colon [0,1]\to\mathbb{R}$, which are right-continuous, have limits from the left and have $f(0)=0$. Endow this space with the uniform convergence topology. We write $\mathcal{AC}$ for the subspace of $D[0,1]$ consisting of all absolutely continuous functions. By the Mogulskii theorem (see [4], Theorem 5.1.2), the distribution of $\{(1/n)\sum_{j=1}^{\lfloor nt\rfloor}\ln A_{\sigma_j},\, t\in[0,1]\}$ satisfies a large deviation principle in $D[0,1]$. The rate function for this large deviation principle is

$$I(f) \;:=\; \int_0^1 \sup_\lambda\Big\{f'(t)\lambda - \ln\mathbb{E}[A_{\varrho1}^\lambda]\Big\}\, \mathrm{d}t,$$

if $f\in\mathcal{AC}$, and $I(f)=+\infty$ if $f\notin\mathcal{AC}$. Note that $I(f)$ is lower semicontinuous, but does not necessarily have compact level sets, so that it is not necessarily a 'good' rate function.

The function $g\colon \mathcal{AC}\to(-\infty,0]$ defined by $g(f)=\min_{t\in[0,1]}f(t)$ is continuous in $\mathcal{AC}$. By the lower bound in Varadhan's lemma (see [4], Lemma 4.3.4), we get

$$\liminf_{n\to\infty}\frac{1}{n}\ln \mathbb{E}\Big[e^{\min_{t\in[0,1]}\sum_{j=1}^{\lfloor nt\rfloor}\ln A_{\sigma_j}}\Big] \;\geq\; \sup_{f\in\mathcal{AC}}\Big\{\min_{t\in[0,1]}f(t) - I(f)\Big\}. \tag{2.3}$$

Since the function $\phi(\lambda) := \ln \mathbb{E}[A_{\varrho 1}^{\lambda}]$ is convex, it follows from Proposition 5.1 in the Appendix that the solution to the variational formula on the right hand side of (2.3) is given by

$$\sup_{f \in \mathcal{AC}} \left\{ \min_{t \in [0,1]} f(t) - \int_0^1 \sup_{\lambda} \{f'(u)\lambda - \ln \mathbb{E}[A_{\varrho 1}^{\lambda}]\}\, \mathrm{d}u \right\} = \inf_{\lambda \in [0,1]} \ln \mathbb{E}[A_{\varrho 1}^{\lambda}]. \tag{2.4}$$

Combining (2.2), (2.3) and (2.4), it follows that (2.1) is satisfied, proving Theorem 2.1. ∎

# 3 Markovian environment

We now show that the proof used in the previous section allows us to treat more general dependence between the weights, provided that we have a suitable large deviation principle.

Let $\sigma$ be an infinite line of descent $[\varrho, \infty)$. In this section, we assume that there is a process $\{M_{\sigma_i}, i \geq 1\}$ in a Polish space $\Sigma$, such that the pair $\Gamma_{\sigma_i} := (A_{\sigma_i}, M_{\sigma_i})$, with $i \geq 0$, is a Markov chain on $\Sigma' = (0, \infty) \times \Sigma$, with transition kernel

$$K(x, B) := \mathbb{P}\big(\Gamma_{\sigma_i} \in B \,|\, \mathcal{F}_{i-1} \cap \{\Gamma_{\sigma_{i-1}} = x\}\big), \tag{3.1}$$

for any $B \in \mathcal{B} := \mathcal{B}(\Sigma')$; here, $\mathcal{F}_i$, $i \geq 1$, is the natural filtration of the process $\Gamma_{\sigma_i}$, $i \geq 1$.

**Remark 3.1** *We assume, for each $\nu$, that the random variables $(\Gamma_{\nu i}, i \geq 1)$ are generated from some joint distribution whose marginals, conditionally on $\Gamma_\nu$ and the further past, are equal to $K(\Gamma_\nu, \cdot)$. Note that we do not need to assume independence among the $(\Gamma_{\nu i}, i \geq 1)$. The construction can proceed sequentially along the tree, using any initial condition for $\Gamma_\varrho \in \Sigma'$.*

For any vertex $\nu$, recall that the set of vertices which are descendants of $\nu$ consists of those vertices $\mu$ such that $\nu$ lies on the shortest path connecting $\mu$ to the root $\varrho$. We deem $\nu$ to be its own descendant. We are motivated by examples where the process $\{A_{\sigma_i}, i \geq 0\}$ is determined as a functional of Markov processes defined on rays.

In order to make use of a uniform large deviation principle for Markov chains, we make the following assumption. It is somewhat reminiscent of the requirement for

Harris recurrence, but is much stronger, in that many specific measures must be dominated. We also make use of the assumption to construct regeneration events for the environment.

**Assumption 1.** *There exist integers $0 < \ell \leq N$ and a constant $\kappa \geq 1$ such that, for all $x, y \in \Sigma'$ and $B \in \mathcal{B}$, we have*

$$K^{(\ell)}(x, B) \leq \frac{\kappa}{N} \sum_{m=1}^{N} K^{(m)}(y, B), \tag{3.2}$$

*where $K^{(\ell)}$ stands for the $\ell$-th convolution of the kernel $K$.*

Note that i.i.d. $\{A_{\sigma_i}\}$ satisfy Assumption 1, and so does any finite state space irreducible Markov chain $(A_{\sigma_i}, M_{\sigma_i})$, but there are of course many other possibilities.

Although the classical results on large deviations require the finiteness of all moments (see condition $(\hat{\mathbf{U}})$, page 95 of [5]), we do not assume that the support of the $A_{\sigma_i}$ is either compact or bounded away from zero; nor do we make any assumptions on the moments of $A_{\sigma_i}$. Instead, we use truncation in order to apply the general results. We nonetheless need one further assumption. Setting

$$\eta_{\varepsilon, r} := 1 - \inf_{y \in \Sigma'} \mathbb{P}\big(\varepsilon < A_{\sigma_1} \leq r \,|\, \Gamma_{\sigma_0} = y\big), \tag{3.3}$$

we require:

**Assumption 2.** *For $\eta := \liminf_{\varepsilon \downarrow 0, r \to \infty} \eta_{\varepsilon, r}$, we have $\eta < 1$.*

The following example shows that, even when $A_{\sigma_i}$ itself is a Markov chain, Assumption 1 does not in general imply Assumption 2.

**Example 3.2** *Suppose that $K(x, \cdot)$ is the mixture $(1 - \alpha)\mathrm{Exp}\,(1) + \alpha\mathrm{Exp}\,(\hat{x})$, where $\hat{x} := x \vee 1$, $0 \leq \alpha \leq 1$ and $\mathrm{Exp}\,(\lambda)$ denotes the exponential distribution with mean $\lambda^{-1}$. Then it is easy to check that $\eta = \alpha$, and that $K^{(2)}(x, \cdot)$ has a density $k^{(2)}(x, \cdot)$ satisfying*

$$(1 - \alpha e^{-1})e^{-w} \;\leq\; k^{(2)}(x, w) \;\leq\; 3e^{-w},$$

*uniformly in $x$, so that Assumption 1 is satisfied with $\ell = 2$, but Assumption 2 is not satisfied if $\alpha = 1$.*

For all $x \in \Sigma'$ and for all $B \in \mathcal{B}(\Sigma')$, define $\Sigma'_\varepsilon := (\varepsilon, \infty) \times \Sigma$ and

$$\overline{K}_\varepsilon(x, B) \;:=\; \frac{K(x, B \cap \Sigma'_\varepsilon)}{K(x, (\varepsilon, \infty) \times \Sigma)};$$

note that $\overline{K}_\varepsilon$ is a probability kernel on $\Sigma'_\varepsilon$, and that it satisfies Assumption 1 for all $\varepsilon$ such that $\eta_{\varepsilon,\infty} := 1 - \inf_{y \in \Sigma'} \mathbb{P}\big(A_{\sigma_1} > \varepsilon \,|\, \Gamma_{\sigma_0} = y\big) < 1$. To prove the latter fact, observe that, for all Borel sets $B \in \mathcal{B}(\Sigma'_\varepsilon)$, we have

$$
\begin{aligned}
\overline{K}_\varepsilon^{(\ell)}(x, B) \;\le\; (1 - \eta_{\varepsilon,\infty})^{-\ell} K^{(\ell)}(x, B) \;&\le\; \frac{\kappa}{(1 - \eta_{\varepsilon,\infty})^\ell N} \sum_{j=1}^{N} K^{(j)}(y, B) \\
&\le\; \frac{\kappa}{(1 - \eta_{\varepsilon,\infty})^\ell N} \sum_{j=1}^{N} \overline{K}_\varepsilon^{(j)}(y, B).
\end{aligned}
\tag{3.4}
$$

For any $0 < \varepsilon < 1$, and for some $x^* \in [1, \infty) \times \Sigma$, define the measure $\beta_\varepsilon$ on $\Sigma'$ by

$$
\beta_\varepsilon(\cdot) \;:=\; \overline{K}_\varepsilon^{(\ell)}(x^*, \cdot),
\tag{3.5}
$$

where $\ell$ is the same as in Assumption 1. Set $\beta(\cdot) = \lim_{\varepsilon \to 0} \beta_\varepsilon(\cdot) = K^{(\ell)}(x^*, \cdot)$.

**Proposition 3.3** *Under Assumptions 1 and 2, if*

$$
\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \int_{\Sigma'} \ln \mathbf{P}\big(T_{-1} > T_n \,|\, \Gamma_\varrho = y\big) \beta_\varepsilon(\mathrm{d}y) \;>\; -\ln b,
\tag{3.6}
$$

*then $\mathbf{X}$ is transient.*

**Proof.** Under the assumption that $\mathbf{X}$ is recurrent, we construct a random subtree of $\mathcal{G}$, consisting of green vertices, that contains a number of vertices stochastically larger than the number of vertices in a supercritical Galton–Watson tree. These green vertices are such that the random walk $\mathbf{X}$ visits them before it first reaches $\varrho^{-1}$. The fact that this random subtree is infinite with positive probability implies a contradiction, and hence that $\mathbf{X}$ is transient.

A direct calculation shows that, for any $y \in \Sigma$ and $1 \le j \le N$,

$$
\begin{aligned}
p_j^\varepsilon(y, B) \;:=\; &\mathbb{P}\left( \bigcap_{l=1}^{j-1} \{A_l \ge \varepsilon\},\, \Gamma_j \in B \;\middle|\; \Gamma_0 = y \right) \\
\ge\; &(1 - \eta_{\varepsilon,\infty})^N \overline{K}_\varepsilon^{(j)}(y, B),
\end{aligned}
$$

8

for all $B \in \mathcal{B}(\Sigma'_\varepsilon)$. It thus follows, from (3.4) and (3.5), that if $U$ is uniformly distributed on $\{1, 2, \dots, N\}$, independently of $\Gamma$, then, for all $y \in \Sigma'$ and $B \in \mathcal{B}(\Sigma'_\varepsilon)$,

$$
\begin{aligned}
p_y^\varepsilon(B) \ &:= \ \frac{1}{N} \sum_{j=1}^N p_j^\varepsilon(y, B) \ = \ \mathbb{P}\left( \bigcap_{l=1}^{U-1} \{A_l \geq \varepsilon\}, \Gamma_U \in B \ \middle| \ \Gamma_0 = y \right) \\
&\geq \ \kappa^{-1}(1 - \eta_{\varepsilon,\infty})^{N+\ell} \overline{K}_\varepsilon^{(\ell)}(x^*, B) \ =: \ \delta_\varepsilon \beta_\varepsilon(B),
\end{aligned} \tag{3.7}
$$

with $\delta_\varepsilon > 0$ for all $\varepsilon$ small enough, since $\eta_{\varepsilon,\infty} < 1$ for all $\varepsilon$ small enough, in view of Assumption 2. Because of (3.7), $\delta_\varepsilon \beta_\varepsilon$ is absolutely continuous with respect to $p_y^\varepsilon$, and

$$
\delta_\varepsilon \frac{d\beta_\varepsilon}{dp_y^\varepsilon}(y') \ \leq \ 1 \ = \ \sum_{j=1}^N f_j^\varepsilon(y, y'),
$$

where

$$
f_j^\varepsilon(y, y') \ := \ \frac{1}{N} \frac{dp_j^\varepsilon(y, \cdot)}{dp_y^\varepsilon}(y'), \quad 1 \leq j \leq N.
$$

Hence, if we set $g_j^\varepsilon(y, \cdot) := \delta_\varepsilon \frac{d\beta_\varepsilon}{dp_y^\varepsilon}(y') f_j^\varepsilon(y, y')$, $1 \leq j \leq N$, it follows that $0 \leq g_j^\varepsilon(y, y') \leq f_j^\varepsilon(y, y')$ for all $y' \in \Sigma'_\varepsilon$ and $\sum_{j=1}^N g_j^\varepsilon(y, y') = \delta_\varepsilon \frac{d\beta_\varepsilon}{dp_y^\varepsilon}(y')$.

This justifies the following construction. Starting at a vertex $\nu$ that has an infinite line of descent, let $\Gamma_0$ denote the value $y \in \Sigma'_\varepsilon$ at $\nu$. Realize $U = U_\nu$ uniformly distributed on $\{1, 2, \dots, N\}$ and a random variable $U'$ uniformly distributed on $[0, 1]$, independently of all else. Because there is an infinite line of descent from $\nu$, there is at least one line of descent from $\nu$ of length $U$; if there is more than one, choose one at random. Denote it by $\nu_1, \dots, \nu_U$, and set $\nu_0 := \nu$. Independently, realize the chain $\Gamma$ along this line of descent, starting from $\Gamma_0$ at $\nu$. Say that the event $E_\nu$ occurs if $A_j \geq \varepsilon$, $1 \leq j \leq U - 1$, and if $U' f_U^\varepsilon(y, \Gamma_U) \leq g_U^\varepsilon(y, \Gamma_U)$. In this way, the distribution $\beta_\varepsilon$ is obtained as the distribution of $\Gamma_U$ on an event $E_\nu$ of probability $\delta_\varepsilon$, and with $A_j \geq \varepsilon$, $1 \leq j \leq U$. For any pair of vertices $\nu$, $\mu$, with $\mu$ a descendant of $\nu$, denote by $\mathbf{X}(\nu, \mu)$ the process $\mathbf{X}$ restricted to the finite graph consisting of the vertices in the finite ray $[\nu^{-1}, \mu]$ and the edges connecting them. A vertex $\nu'$ is green if it has an infinite line of descent, and is descended from a green vertex $\nu$ in the following way. $E_\nu$ must occur, and then $\mathbf{X}(\nu, \nu_U)$ has to reach $\nu_U$ before hitting $\nu^{-1}$; the latter event has probability at least $\{\varepsilon/(1 + \varepsilon)\}^N$. Finally, $\nu'$ should be a descendant at distance $n$ from $\nu_U$, and $\mathbf{X}(\nu_U, \nu')$ should reach $\nu'$ before it hits $\nu_U^{-1}$, an event of probability

$$
\int_{\Sigma'_\varepsilon} \mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho = y) \beta_\varepsilon(dy).
$$

9

Thus the expected number of green 'offspring' of a green vertex $\nu$ is at least

$$b^n \delta_\varepsilon \{\varepsilon/(1+\varepsilon)\}^N \left\{ \int_{\Sigma'_\varepsilon} \mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho = y) \beta_\varepsilon(dy) \right\} (1-q), \qquad (3.8)$$

where $q$ denotes the probability of the extinction of the underlying Galton–Watson tree. Next, we show that (3.6) implies that we can choose $\varepsilon$ small enough and $n$ large enough that the quantity in (3.8) becomes larger than 1. By taking the natural logarithm of (3.8) and dividing by $n$, we have

$$\frac{1}{n} \ln \left( \delta_\varepsilon \{\varepsilon/(1+\varepsilon)\}^N (1-q) \right) + \ln b + \frac{1}{n} \ln \int_{\Sigma'_\varepsilon} \mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho = y) \beta_\varepsilon(dy). \qquad (3.9)$$

Fix $\varepsilon > 0$ such that

$$\liminf_{n \to \infty} \frac{1}{n} \int_{\Sigma'_\varepsilon} \ln \mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho = y) \beta_\varepsilon(dy) \; > \; -\ln b,$$

as we may, in view of (3.6). Then, for this choice of $\varepsilon$, the liminf of (3.9) as $n \to \infty$ is larger than 0, proving that the quantity in (3.8) is larger than 1 for our choice of $\varepsilon$ and for $n$ large enough.

By construction, the distribution of the number of green offspring is the same for all green vertices. Hence, choosing an appropriate $\varepsilon$, and then $n$ large enough that the quantity in (3.8) is larger than one, the Galton–Watson tree of green vertices is supercritical. ∎

The proofs that follow rely on large deviations results. These cannot be directly applied to $A$, so we need to consider truncations. For this reason, it is convenient to introduce the large deviations results that we shall use applied to a generic process $W := (W_i, \, i \geq 0)$, which, together with a process $\widetilde{M}$ on $\Sigma$, makes $\widetilde{\Gamma}$ defined by $\widetilde{\Gamma}_i := (W_i, \widetilde{M}_i)$ a Markov chain on $\mathbb{R} \times \Sigma$. Let $\widetilde{K}$ denote the kernel of this process.

Define

$$\Lambda^{(\widetilde{K})}(\lambda) \; := \; \limsup_{n \to \infty} \sup_{\tilde{y} \in \mathbb{R} \times \Sigma} \frac{1}{n} \ln \mathbb{E}\left[ e^{\lambda \sum_{i=1}^n W_i} \,|\, \widetilde{\Gamma}_\varrho = \tilde{y} \right], \quad \Lambda^*_{\widetilde{K}}(x) \; := \; \sup_\lambda \{\lambda x - \Lambda^{(\widetilde{K})}(\lambda)\}, \qquad (3.10)$$

and let

$$S_n^{(\widetilde{K})}(t) \; := \; \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} W_i, \qquad t \in [0,1].$$

10

**Theorem 3.4** *Fix $0 < C < R < \infty$, and assume that $W_i \in (C, R)$ a.s., for each $i$. If $\widetilde{K}$ satisfies Assumption 1, then, for any $\Theta \in \mathcal{B}^+$, we have*

$$- \inf_{x \in \Theta^{\circ}} \Lambda_{\widetilde{K}}^*(x) \; \leq \; \liminf_{n \to \infty} \frac{1}{n} \ln \inf_{y \in \Sigma'} \mathbb{P}(S_n^{(\widetilde{K})}(1) \in \Theta \mid \widetilde{\Gamma}_\varrho = y)$$

$$\leq \; \limsup_{n \to \infty} \frac{1}{n} \ln \sup_{y \in \Sigma'} \mathbb{P}(S_n^{(\widetilde{K})}(1) \in \Theta \mid \widetilde{\Gamma}_\varrho = y) \; \leq \; - \inf_{x \in \overline{\Theta}} \Lambda_{\widetilde{K}}^*(x).$$

**Proof of Theorem 3.4.** The kernel $\widetilde{K}$ satisfies condition $(\hat{\mathbf{U}})$, page 95 of [5]. Hence, the theorem is a consequence of the more general Theorem 4.1.14, page 97 of [5], combined with (4.1.24) page 100 of [5], to identify the rate function. ∎

Recall that $\mathcal{AC}$ denotes the space of absolutely continuous functions $f$ defined on $[0, 1]$, with $f(0) = 0$ and $D[0, 1]$ the space of functions $f$ which are right continuous and have limits from the left, and have $f(0) = 0$. Both spaces are endowed with the uniform convergence topology. The following result is due to Dembo & Zajic [3].

**Theorem 3.5** *Under the hypotheses of Theorem 3.4, the sequence $\{S_n^{(\widetilde{K})}(t), t \in [0, 1]\}$ in $D[0, 1]$ satisfies a large deviations principle with the good, convex, rate function*

$$I_{\widetilde{K}}^*(f) \; := \; \begin{cases} \int_0^1 \Lambda_{\widetilde{K}}^*(\dot{f}(u)) \, \mathrm{d}u, & \text{if } f \in \mathcal{AC} \\ +\infty, & \text{otherwise.} \end{cases}$$

**Proof.** In virtue of Theorem 3.4, $S_n^{(\widetilde{K})}(1)$ satisfies a uniform large deviations principle. We can then use Dembo & Zajic ([3], Theorem 3a) to conclude that $\{S_n^{(\widetilde{K})}(t), t \in [0, 1]\}$ satisfies an LDP with rate function $I_{\widetilde{K}}^*(\cdot)$. ∎

Note that, since $(A_{\sigma_i}, M_{\sigma_i})$ is a Markov chain in $(0, \infty) \times \Sigma$, then $(\ln A_{\sigma_i}, M_{\sigma_i})$ is a Markov chain on $\mathbb{R} \times \Sigma$. Define the kernel

$$K_{\ln}((\tilde{u}, z), B) \; := \; K((e^{\tilde{u}}, z), E(B)), \qquad \tilde{u} \in \mathbb{R}, \; z \in \Sigma, \; B \in \mathcal{B},$$

where $E(B) := \{(e^{\tilde{u}}, z) \colon (\tilde{u}, z) \in B\}$. Note that, if $K$ satisfies Assumption 1, then so does the kernel $K_{\ln}$.

For $R \in (0, \infty)$ and $C \in [-\infty, 0)$, define the probability kernel $Q_{C,R}$ on $(C, R] \times \Sigma$ by

$$Q_{C,R}(\tilde{y}, (\mathrm{d}\tilde{u}, \mathrm{d}z)) \; := \; \frac{K_{\ln}(\tilde{y}, (\mathrm{d}\tilde{u}, \mathrm{d}z))}{K_{\ln}(\tilde{y}, (C, R] \times \Sigma)}, \tag{3.11}$$

and set $Q_R := Q_{-\infty, R}$.

**Theorem 3.6** *If $K$ satisfies Assumption 1, then*

(i) *If Assumption 2 holds, the condition*

$$\limsup_{\min\{-C,R\}\to\infty} \inf_{\lambda\in[0,1]} \Lambda^{(Q_{C,R})}(\lambda) \;>\; -\ln b - \ln(1-\eta) \tag{3.12}$$

*implies transience of $\mathbf{X}$ on $\mathcal{G}$. The constant $\eta$ is the one introduced in Assumption 2.*

(ii) *The condition*

$$\inf_{\lambda\in[0,1]} \Lambda^{(K_{\ln})}(\lambda) \;<\; -\ln b \tag{3.13}$$

*implies positive recurrence of $\mathbf{X}$ on $\mathcal{G}$.*

**Remark 3.7** *In the case of an i.i.d. environment, (3.12) coincides with the condition $\inf_{\lambda\in[0,1]} \Lambda^{(K_{\ln})}(\lambda) \;>\; -\ln b$, which is then also the same as that of Theorem 2.1.*

**Corollary 3.8** *In the uniformly elliptic case, i.e. if there exists $\varepsilon > 0$ such that*

$$\inf_x K\big(x, (\varepsilon, \varepsilon^{-1}) \times \Sigma\big) = 1,$$

*so that then $\eta = 0$, we have the following sharp transition:*

(i) *The condition*

$$\inf_{\lambda\in[0,1]} \Lambda^{(K_{\ln})}(\lambda) \;>\; -\ln b \tag{3.14}$$

*implies transience of $\mathbf{X}$ on $\mathcal{G}$.*

(ii) *The condition*

$$\inf_{\lambda\in[0,1]} \Lambda^{(K_{\ln})}(\lambda) \;<\; -\ln b \tag{3.15}$$

*implies positive recurrence of $\mathbf{X}$ on $\mathcal{G}$.*

The following examples show particular ways to compute bounds for $\Lambda^{(K_{\ln})}$.

**Example 3.9** *Suppose that $A_{\sigma_i}$ evolves as a discrete irreducible aperiodic Markov chain, with state space $\Xi = (a_1, a_2, \ldots, a_\ell)$, where $a_i \in (0, \infty)$ for all $i$, and with transition matrix $K = (k_{i,j}, 1 \le i, j \le \ell)$. Note that, in the finite case, $\Lambda^{(K_{\ln})}(\lambda)$ coincides with $\ln \rho(\lambda)$, where $\rho(\lambda)$ is the Perron-Frobenius eigenvalue of the matrix whose $(i, j)$th entry is $a_j^\lambda k_{i,j}$, (see [4], page 74). Using the Gershgorin circle theorem, the Perron–Frobenius eigenvalue is bounded above by the largest row sum. Hence*

$$\rho(\lambda) \;\le\; \max_i \sum_{j=1}^{\ell} k_{i,j} a_j^\lambda.$$

*Hence, Corollary 3.8 implies that if*

$$\inf_{\lambda \in [0,1]} \max_i \sum_{j=1}^{\ell} k_{i,j} a_j^\lambda \;<\; 1/b,$$

*then the process is recurrent. The Gershgorin circle theorem can also be used to get the lower bound*

$$\rho(\lambda) \;\ge\; \min_i \left( k_{i,i} a_i^\lambda - \sum_{j:\, j \ne i} k_{i,j} a_j^\lambda \right),$$

*useful if $K$ is close to being diagonal. Thus Corollary 3.8 implies that if*

$$\inf_{\lambda \in [0,1]} \min_i \left( k_{i,i} a_i^\lambda - \sum_{j:\, j \ne i} k_{i,j} a_j^\lambda \right) \;>\; 1/b,$$

*then the process is transient.*

In the case $\ell = 2$, $\Lambda^{(K_{\ln})}(\lambda)$ can of course be computed explicitly. For

$$K \;:=\; \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \beta & \beta \end{pmatrix},$$

with $\alpha, \beta \in (0, 1)$, we have

$$\Lambda^{(K_{\ln})}(\lambda) = \ln \frac{1}{2} \left( \alpha a_1^\lambda + \beta a_2^\lambda + \sqrt{(\alpha a_1^\lambda + \beta a_2^\lambda)^2 + 4(1 - \alpha - \beta) a_1^\lambda a_2^\lambda} \right). \qquad (3.16)$$

The procedure used in Example 3.9 can be carried out in continuous space through discretization, as the following simple example shows.

**Example 3.10** *Let $K$ be the kernel of a Markov process with compact state space $U$. Consider a finite cover of $U$, say $U_1, U_2, \ldots, U_\ell$, with the property that if $i \neq j$ then $U_i \not\subset U_j$, and an $\ell \times \ell$ matrix with strictly positive elements $B = \{b_{i,j}\}$ such that, for all $t > 0$,*

$$K(y, (0, t]) \leq \sum_{j=1}^{\lfloor t \rfloor} b_{i,j} \qquad \forall y \in U_i.$$

*We emphasize that $B$ need not be a transition matrix. Set $\Xi = \{a_i, i \leq \ell\}$, where $a_i = \sup\{b \colon b \in U_i\}$. Then*

$$\Lambda^{(K_{\ln})}(\lambda) \leq \limsup_{n \to \infty} \sup_{j_0 \leq \ell} \frac{1}{n} \ln \sum_{j_1=1}^{\ell} \cdots \sum_{j_n=1}^{\ell} e^{\lambda \ln a_{j_1}} \cdots e^{\lambda \ln a_{j_n}} b_{j_0, j_1} \cdots b_{j_{n-1}, j_n}.$$

*Hence, the logarithm of the Perron-Frobenius eigenvalue of the matrix $\{a_j^\lambda b_{i,j}\}$ is an upper bound for $\Lambda^{(K_{\ln})}$. We can then proceed as in the previous example to determine a sufficient condition for recurrence. An analogous procedure, with lower bounds, can be applied to derive sufficient conditions for transience.*

**Proof of Theorem 3.6.** We first prove that condition (3.12) implies that (3.6) holds, and hence, by Proposition 3.3, that $\mathbf{X}$ is transient. Observe that, for any $r > 0$, we have

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \int_{\Sigma'} \frac{1}{n} \ln \mathbf{P}(T_n < T_{-1} \mid \Gamma_\varrho = y) \beta_\varepsilon(\mathrm{d}y)$$

$$= \liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \int_{\Sigma'} \frac{1}{n} \ln \mathbb{E}\Big[ \big( \sum_{l=0}^{n} \prod_{j=1}^{l-1} A_{\sigma_j}^{-1} \big)^{-1} \mid \Gamma_\varrho = y \Big] \beta_\varepsilon(\mathrm{d}y)$$

$$\geq \liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \int_{\Sigma'} \frac{1}{n} \ln \mathbb{E}\Big[ \big( \sum_{l=0}^{n} \prod_{j=1}^{l-1} (A_{\sigma_j} \wedge r)^{-1} \big)^{-1} \mid \Gamma_\varrho = y \Big] \beta_\varepsilon(\mathrm{d}y)$$

$$\geq \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\Big[ \big( \sum_{l=0}^{n} \prod_{j=1}^{l-1} (A_{\sigma_j} \wedge r)^{-1} \big)^{-1} \mid \Gamma_\varrho = y \Big]$$

$$\geq \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\Big[ e^{\min_{t \in [0,1]} \sum_{i=1}^{\lfloor nt \rfloor} \ln(A_{\sigma_i} \wedge r)} \mid \Gamma_\varrho = y \Big].$$

$$(3.17)$$

14

For $r > 1 > c > 0$, let $A_{\sigma_i}(c, r) = (A_{\sigma_i} \vee c) \wedge r$, and set $C = \ln c$ and $R = \ln r$. Then, writing $\tilde{y}_j := (\tilde{u}_j, z_j) \in \mathbb{R} \times \Sigma$ for $j \geq 1$ and $\tilde{y}_0 := (\ln u, z)$ for $(u, z) = y$, we have

$$
\liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} \ln(A_{\sigma_j} \wedge r)} \,\big|\, \Gamma_\varrho = y \right]
$$

$$
\geq \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} \ln(A_{\sigma_j} \wedge r)} \mathbb{1}_{\bigcap_{i=1}^n \{A_{\sigma_i} > c\}} \,\big|\, \Gamma_\varrho = y \right]
$$

$$
= \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} \ln A_{\sigma_j}(c,r)} \mathbb{1}_{\bigcap_{i=1}^n \{A_{\sigma_i} > c\}} \,\big|\, \Gamma_\varrho = y \right] \tag{3.18}
$$

$$
\geq \liminf_{n \to \infty} \inf_{y \in \Sigma'} \int_{([C,R] \times \Sigma)^n} e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{u}_j} \prod_{j=1}^n K_{\ln}\big(\tilde{y}_{j-1}, \mathrm{d}\tilde{y}_j\big).
$$

Choosing $c$ small enough and $r$ large enough that

$$
\inf_{y \in \Sigma'} K_{\ln}(y, [c, r] \times \Sigma) \geq (1 - \eta_{c,r}) > 0,
$$

as we may, because $\eta < 1$, we have

$$
\liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{E}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} (\ln A_{\sigma_j} \wedge r)} \,\big|\, \Gamma_\varrho = y \right]
$$

$$
\geq \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \left[ (1 - \eta_{c,r})^n \int_{([C,R] \times \Sigma)^n} e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{u}_j} \prod_{j=1}^n \frac{K_{\ln}\big(\tilde{y}_{j-1}, \mathrm{d}\tilde{y}_j)\big)}{K_{\ln}(\tilde{y}_{j-1}, (C, R] \times \Sigma)} \right]
$$

$$
= \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \left[ (1 - \eta_{c,r})^n \widetilde{\mathbb{E}}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} W_j} \,\big|\, \Gamma_\varrho = y \right] \right]
$$

$$
= \ln(1 - \eta_{c,r}) + \liminf_{n \to \infty} \inf_{y \in \Sigma'} \frac{1}{n} \ln \widetilde{\mathbb{E}}\left[ e^{\min_{t \in [0,1]} \sum_{j=1}^{\lfloor nt \rfloor} W_j} \,\big|\, \Gamma_\varrho = y \right],
$$

$$
\tag{3.19}
$$

where $\widetilde{\mathbb{E}}$ is the expectation with respect to the Markov chain $\widetilde{G} = (W, \widetilde{M})$ with probability kernel $Q_{C,R}(\tilde{y}', \mathrm{d}\tilde{y})$ introduced in (3.11).

Next we prove that the kernel $Q_{C,R}$ satisfies Assumption 1. Note that, for Borel sets $F \subset (C, R]$ and $E \in \Sigma$, and for any $\tilde{x}, \tilde{y} \in (C, R] \times \Sigma$, we have

$$
Q_{C,R}^{(\ell)}(\tilde{x}, F \times E) \leq (1 - \eta_{c,r})^{-\ell} K_{\ln}^{(\ell)}(\tilde{x}, F \times E) \leq \frac{M}{(1 - \eta_{c,r})^\ell N} \sum_{j=1}^N K_{\ln}^{(j)}(\tilde{y}, F \times E)
$$

$$
\leq \frac{M}{(1 - \eta_{c,r})^\ell N} \sum_{j=1}^N Q_{C,R}^{(j)}(\tilde{y}, F \times E). \tag{3.20}
$$

15

In the last step, we have used the inequality

$$K_{\ln}^{(n)}(\tilde{y}, F \times E) \leq Q_{C,R}^{(n)}(\tilde{y}, F \times E),$$

valid for $F \subset (C, R]$ and $n \geq 1$, which is easily proved by induction.

Combining Theorem 3.5 with Varadhan's lemma, using the **uniform** large deviations stated in Theorem 3.4, we find that

$$\liminf_{n\to\infty} \inf_{y\in\Sigma'} \frac{1}{n} \ln \mathbb{E}\left[e^{\min_{t\in[0,1]} \sum_{j=1}^{\lfloor nt \rfloor} (\ln A_{\sigma_j} \wedge r)} \,\middle|\, \Gamma_\varrho = y\right]$$
$$\geq \ln(1 - \eta_{c,r}) + \sup_{f\in\mathcal{AC}} \left\{ \min_{t\in[0,1]} f(t) - I_{Q_{C,R}}^*(f)\right\}, \tag{3.21}$$

and, since the function $\Lambda^{(Q_{C,R})}$ is convex for any $C$ and $R$, Proposition 5.1 can be used to solve the variational formula on the right hand side of (3.21), giving

$$\liminf_{n\to\infty} \inf_{y\in\Sigma'} \frac{1}{n} \ln \mathbb{E}\left[e^{\min_{t\in[0,1]} \sum_{j=1}^{\lfloor nt \rfloor} (\ln A_{\sigma_j} \wedge r)} \,\middle|\, \Gamma_\varrho = y\right] \geq \ln(1 - \eta_{c,r}) + \inf_{t\in[0,1]} \Lambda^{(Q_{\ln c, \ln r})}(t). \tag{3.22}$$

Recalling (3.17), we thus have

$$\liminf_{\varepsilon\to 0} \liminf_{n\to\infty} \int_{\Sigma'_\varepsilon} \frac{1}{n} \ln \mathbf{P}(T_n < T_{-1} \mid \Gamma_\varrho = y)\beta_\varepsilon(\mathrm{d}y) \geq \ln(1 - \eta_{c,r}) + \inf_{t\in[0,1]} \Lambda^{(Q_{\ln c, \ln r})}(t),$$

for any $c, r > 0$ such that $\eta_{c,r} < 1$. By letting $c \to 0$ and $r \to \infty$, we get

$$\liminf_{\varepsilon\to 0} \liminf_{n\to\infty} \int_{\Sigma'_\varepsilon} \frac{1}{n} \ln \mathbf{P}(T_n < T_{-1} \mid \Gamma_\varrho = y)\beta_\varepsilon(\mathrm{d}y)$$
$$\geq \ln(1 - \eta) + \limsup_{\min\{1/c,r\}\to\infty} \inf_{t\in[0,1]} \Lambda^{(Q_{\ln c, \ln r})}(t) > -\ln b,$$

using (3.12), and (i) follows from Proposition 3.3.

Next, we prove that if (3.13) holds, then the process is positive recurrent. In this case we just mimic the proof by Lyons & Pemantle (see [10], proof of Theorem 1.3, page 130). We include the proof for sake of completeness and clarity. We use the well known fact (see [8] Proposition 9-131) that the random walk is positive recurrent if the sum of conductances is a.s. finite, i.e.

$$\sum_{\nu\in\mathcal{G}} \prod_{i=1}^{|\nu|} A_{\nu^{-i}} < \infty \qquad \text{a.s.} \tag{3.23}$$

16

From (3.13) and the definition of $\Lambda$, we can choose $t_0 \in (0,1]$ such that

$$\mathbb{E}\left[\exp\left\{t_0 \sum_{i=1}^{n} \ln A_i\right\} \,\Big|\, \widetilde{G}_0 = \tilde{y}\right] \;<\; (1/b')^n,$$

for some $b' > b$, for all $\tilde{y}$ and all $n$ large enough. Recall that $\nu^{-i}$ denotes the $i$-th ancestor of $\nu$. Because the branching number of the Galton–Watson tree is $b$, this implies that

$$\mathbb{E}\left(\sum_{\nu:\, |\nu|=n} \prod_{i=1}^{n} A_{\nu^{-i}}^{t_0}\right) \;\leq\; (b/b')^n, \tag{3.24}$$

and hence that

$$\sum_{n\geq 1} \sum_{\nu:\, |\nu|=n} \prod_{i=1}^{n} A_{\nu^{-i}}^{t_0} \;<\; \infty \quad \mathbb{P}\text{–a.s.} \tag{3.25}$$

Furthermore, (3.24) also implies that, for all $n$ large enough, $\mathbb{P}(E_n) \leq (b/b')^n$, where

$$E_n \;:=\; \left\{\sum_{\nu:\, |\nu|=n} \prod_{i=1}^{n} A_{\nu^{-i}}^{t_0} \geq 1\right\}.$$

Thus a.s. only finitely many of the events $E_n$ occur, and, on $E_n^c$, since $0 \leq t_0 \leq 1$,

$$\sum_{\nu:\, |\nu|=n} \prod_{i=1}^{n} A_{\nu^{-i}}^{t_0} \;\geq\; \sum_{\nu:\, |\nu|=n} \prod_{i=1}^{n} A_{\nu^{-i}}. \tag{3.26}$$

(3.23) thus follows from (3.25) and (3.26), proving (ii).  ∎

# 4    A walk that changes its environment, once.

In this section, we consider a setting in which the process $\mathbf{X}$ changes the environment. Fix parameters $L, p > 0$, and let $(B_{\sigma_i},\, i \geq 1)$ be a stochastic process, taking values in $[p, +\infty)$, such that the triple $(A_{\sigma_i}, B_{\sigma_i}, M_{\sigma_i})$ is a Markov process along rays. Recalling that

$$S_\nu \;:=\; \inf\{n \geq 0 \colon X_n = \nu\},$$

define

$$G(\nu, n) := \begin{cases} A_\nu & \text{if } \{A_\nu > B_\nu\} \cup \{S_\nu > n\}; \\ L & \text{if } \{A_\nu \leq B_\nu\} \cap \{S_\nu \leq n\}, \end{cases}$$

for each vertex $\nu$ and time $n$. If $X_n = \nu$, given the environment and $\mathcal{F}_n := \sigma\{X_1, X_2, \ldots, X_n\}$, the probability that $X_{n+1} = \nu i$ is given by

$$\frac{G(\nu i, n)}{1 + \sum_{j=1}^{b(\nu)} G(\nu j, n)}, \tag{4.1}$$

so that the probability of a transition from $\nu$ to a state $\nu i$, which has been visited at least once before and for which $A_{\nu i} \leq B_{\nu i}$, is modified by replacing $A_{\nu i}$ by $L$ in its calculation. As before, the process is absorbed at the state $\varrho^{-1}$, and recurrence means that the process is absorbed with probability one at $\varrho^{-1}$. Let

$$D_{\sigma_i} := \begin{cases} L & \text{if } A_{\sigma_i} \leq B_{\sigma_i} \\ A_{\sigma_i} & \text{if } A_{\sigma_i} > B_{\sigma_i}. \end{cases}$$

and denote by $K^*$ the transition kernel of the Markov chain $\Gamma^* := (D_{\sigma_i}, B_{\sigma_i}, A_{\sigma_i}, M_{\sigma_i})$ on $\mathbb{R}_+ \times \Sigma^*$, where $\Sigma^* := \mathbb{R}_+^2 \times \Sigma$ is the state space of $(B_{\sigma_i}, A_{\sigma_i}, M_{\sigma_i})$, and $D_{\sigma_i}$ is singled out. As before, define

$$\eta_{\varepsilon, r} := 1 - \inf_{y \in \mathbb{R}_+ \times \Sigma^*} \mathbb{P}\big(\varepsilon < A_{\sigma_1} \leq r \,|\, \Gamma_{\sigma_0}^* = y\big)$$

and $\eta = \lim_{\varepsilon \to 0, r \to \infty} \eta_{\varepsilon, r}$.

**Theorem 4.1** *Suppose that $K^*$ satisfies Assumption 1 and that Assumption 2 also holds. Suppose that $L, p \geq 1$. Then the condition*

$$\ln(1 - \eta) > -\ln b \tag{4.2}$$

*implies the transience of $\mathbf{X}$ on $\mathcal{G}$.*

**Corollary 4.2** *If $\eta = 0$, then the process $\mathbf{X}$ is transient on $\mathcal{G}$.*

**Proof of Theorem 4.1.** Because the process $\mathbf{X}$ can change $A_\nu$ only at the time $S_\nu$ that $\nu$ is first visited, the proof of Proposition 3.3 can be used to show that, if $\eta < 1$ and

$$\limsup_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \int_{\mathbb{R}_+ \times \Sigma^*} \ln \mathbf{P}\big(T_{-1} > T_n \,|\, \Gamma_\varrho^* = y\big) \beta_\varepsilon(\mathrm{d}y) > -\ln b, \tag{4.3}$$

18

then **X** is transient; here, $\beta_\varepsilon(\cdot)$ is defined as in the previous section, using $K^{*(\ell)}(x^*, \cdot)$ for some $x^* \in \mathbb{R}_+ \times \Sigma^*$, and $\ell$ is chosen in such a way that there exist $N$ and $M$ such that, for all $x, y \in \mathbb{R}_+ \times \Sigma^*$ and Borel sets $B$, we have

$$K^{*(\ell)}(x, B) \leq \frac{M}{N} \sum_{i=1}^{N} K^{*(i)}(y, B).$$

It remains to determine when (4.3) holds.

For a given ray $\sigma = [\varrho, \infty)$, let $Q_i^D := \{\sum_{r=0}^{i} \prod_{j=1}^{r} D_{\sigma_j}^{-1}\}^{-1}$ denote the probability that the random walk starting in $\varrho$ would hit $\sigma_i$ before $\varrho^{-1}$, if the probabilities were determined solely by the $D_{\sigma_i}$, and, for $i \geq 1$, let $q_i^D := Q_i^D / Q_{i-1}^D$ denote the probability that the same random walk starting in $\sigma_{i-1}$ hits $\sigma_i$ before it hits $\varrho^{-1}$. Then, the probability $q_i^A$ that the original walk, after it reached $\sigma_{i-1}$, hits $\sigma_i$ before $\varrho^{-1}$, when started in $\sigma_{i-1}$, is given by $A_{\sigma_i}/\{1 + A_{\sigma_i} - q_{i-1}^D\}$, $i \geq 1$, with $q_0^D$ taken to be zero. This leads us to consider the quantity

$$\Phi_n := \prod_{i=1}^{n} (q_i^A / q_i^D) = \prod_{\substack{i=1 \\ A_{\sigma_i} < b_{\sigma_i}}}^{n} \left( \frac{1 + D_{\sigma_i}^{-1}(1 - q_{i-1}^D)}{1 + A_{\sigma_i}^{-1}(1 - q_{i-1}^D)} \right).$$

Now, since $D_{\sigma_i} \geq \theta := p \wedge L \geq 1$ for all $i$, we have

$$1 - q_i^D \leq i^{-1}, \tag{4.4}$$

so that, on the event $\bigcap_{i=1}^{n}\{A_{\sigma_i} > \varepsilon\}$,

$$\Phi_n \geq \prod_{i=1}^{n}\{1 + i^{-1}\varepsilon^{-1}\}^{-1} \geq kn^{-1/\varepsilon}, \tag{4.5}$$

for a suitable $k$, which depends on $\varepsilon$ only, and

$$\prod_{i=1}^{n} q_i^D = \prod_{i=1}^{n}\{1 - (1 - q_i^D)\} \geq 1/n.$$

Hence, for (4.3), we have

$$
\begin{aligned}
\mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho^* = y) \;&=\; \mathbb{E}\Big[\prod_{i=0}^{n-1} \mathbf{P}_\omega(T_{-1} > T_{i+1} \,|\, T_{-1} > T_i) \,\Big|\, \Gamma_\varrho^* = y\Big] \\
&=\; \mathbb{E}\Big[\prod_{i=1}^{n} q_i^A \,\Big|\, \Gamma_\varrho^* = y\Big] \;=\; \mathbb{E}\Big[\Phi_n \prod_{i=1}^{n} q_i^D \,\Big|\, \Gamma_\varrho^* = y\Big] \\
&\geq\; kn^{-1/\varepsilon}\, \mathbb{E}\Big[\prod_{i=1}^{n} q_i^D \mathbb{1}_{\cap_{i=1}^n \{A_{\sigma_i} > \varepsilon\}} \,\Big|\, \Gamma_\varrho^* = y\Big] \\
&\geq\; kn^{-(1/\varepsilon)-1}\, \mathbb{P}\Big[\bigcap_{i=1}^{n}\{A_{\sigma_i} > \varepsilon\} \,\Big|\, \Gamma_\varrho^* = y\Big].
\end{aligned}
\tag{4.6}
$$

Hence, from the definition of $\eta_{\varepsilon,\infty}$,

$$
\liminf_{n\to\infty}\ \inf_{y\in\mathbb{R}_+\times\Sigma^*}\ \frac{1}{n}\ln \mathbf{P}(T_{-1} > T_n \,|\, \Gamma_\varrho^* = y) \;\geq\; \ln(1 - \eta_{\varepsilon,\infty}),
$$

and the theorem follows by letting $\varepsilon \to 0$ and using (4.3). ∎

**Remark 4.3** *Consider a once-reinforced random walk on a Galton–Watson tree, defined as follows. Each edge is initially assigned weight 1. The walk moves to any one of its nearest neighbours, with probability proportional to the weight of the edge traversed. The first time an edge is traversed, its weight becomes $1+\Delta$, for $\Delta > -1$, and is never changed again. With the choice of $L = 1$ and $A_\nu = 1/(1+\Delta)$ for all $\nu \in \mathcal{G}$, and with $b_\nu = p = \min\{1, 1/(1+\Delta)\}$, for all $\nu$, our walk is exactly a once-reinforced random walk. Theorem 4.1 then implies transience for this class of processes, as already proved in [1] or [2].*

The next result holds for all choices of $L$ and $p$ such that $L < p$. Define the kernel

$$
K_{\ln}^*((\tilde{w}, c, \tilde{u}, z), B) \;:=\; K^*((e^{\tilde{w}}, c, e^{\tilde{u}}, z), E^*(B)),
$$

where $\tilde{w} \in [\ln L, \infty)$, $c \in (p, \infty)$, $\tilde{u} \in \mathbb{R}$, $z \in \Sigma$, $B \in \mathcal{B}$ and $E^*(B) := \{(e^{\tilde{w}}, c, e^{\tilde{u}}, z) \colon (\tilde{w}, c, \tilde{u}, z) \in B\}$. Note that, if $K^*$ satisfies Assumption 1, then so does the kernel $K_{\ln}^*$.

For $R \in (0, \infty)$ and with $C := \ln L$, define the probability kernel $Q^*_{C,R}$ on $[C, R] \times [p, \infty) \times [C, R] \times \Sigma$ by

$$Q^*_{C,R}\big(\tilde{y}, (\mathrm{d}\tilde{w}, \mathrm{d}c, \mathrm{d}\tilde{u}, \mathrm{d}z)\big) \;:=\; \frac{K^*_{\ln}\big(\tilde{y}, (\mathrm{d}\tilde{w}, \mathrm{d}c, \mathrm{d}\tilde{u}, \mathrm{d}z)\big)}{K^*_{\ln}(\tilde{y}, [C, R] \times (p, \infty) \times [C, R] \times \Sigma)}. \tag{4.7}$$

This kernel describes the distribution of the jumps of the process $(\ln D_{\sigma_i}, B_{\sigma_i}, \ln A_{\sigma_i}, M_{\sigma_i})$ when $\ln A_{\sigma_i}$ is conditioned to be in the interval $[C, R]$. This also implies that $\ln D_{\sigma_i}$ takes values in the same interval. If $K^*$ satisfies Assumption 1, then so does the kernel $Q^*_{C,R}$. Define

$$\tilde{\Lambda}^{(Q^*_{C,R})} \;:=\; \limsup_{n \to \infty} \sup_{\tilde{y}} \frac{1}{n} \ln \mathbb{E}\left[ \mathrm{e}^{\lambda \sum_{i=1}^{n} \ln D_{\sigma_i}} \,\big|\, \Gamma^* = \tilde{y} \right],$$

where the expected value is taken with respect to the kernel $Q^*_{C,R}$, and the supremum over the set $[C, R] \times [p, \infty) \times [C, R] \times \Sigma$.

**Theorem 4.4** *Suppose that $K^*$ satisfies Assumption 1 and that $\eta_{L,\infty} < 1$ and $L < p$. Then the condition*

$$\limsup_{R \uparrow \infty} \inf_{\lambda \in [0,1]} \tilde{\Lambda}^{(Q^*_{\ln L, R})}(\lambda) \;>\; -\ln b - \ln(1 - \eta_{L,\infty}) \tag{4.8}$$

*implies the transience of $\mathbf{X}$ on $\mathcal{G}$.*

**Remark 4.5** *Suppose that $\eta = 0$ and $K^*$ satisfies Assumption 1. In this case, if $L, p \geq 1$, then, no matter what is the distribution of the initial environment $(A_\nu, \nu \in \mathcal{G})$, the process $\mathbf{X}$ is transient, by Corollary 4.2. If instead we assume that $L < p$ and $\eta_{L,\infty} < 1$, then the process can also be recurrent. In this case, (4.8) provides a sufficient condition for transience.*

**Proof of Theorem 4.4.** First, note that

$$\Phi_n \geq \prod_{\substack{i=1 \\ A_{\sigma_i} \leq B_{\sigma_i}}}^{n} \left( \frac{1 + D_{\sigma_i}^{-1}(1 - q_{i-1}^D)}{1 + A_{\sigma_i}^{-1}(1 - q_{i-1}^D)} \right) \mathbb{1}_{\{A_{\sigma_i} \geq L\}} \;\geq\; \prod_{\substack{i=1 \\ A_{\sigma_i} \leq B_{\sigma_i}}}^{n} \mathbb{1}_{\{A_{\sigma_i} \geq L\}},$$

21

and hence that

$$\mathbf{P}(T_{-1} < T_n \mid \Gamma^*_\varrho = y) \ \geq \ \mathbb{E}\Big[\prod_{i=1}^{n} q_i^D \mathbb{1}_{\cap_{i=1}^{n-1}\{A_{\sigma_i} \geq L\}} \ \Big| \ \Gamma^*_\varrho = y\Big]$$

$$\geq \ \mathbb{E}\Big[\Big(\sum_{r=0}^{n}\prod_{j=1}^{r-1} D_{\sigma_n}^{-1}\Big)^{-1} \mathbb{1}_{\cap_{i=1}^{n-1}\{A_{\sigma_i} \geq L\}} \mid \Gamma^*_\varrho = y\Big]. \qquad (4.9)$$

Now  the last line of (4.9) is at most

$$\mathbb{E}\Big[\Big((n+1)\max_{r \leq n}\prod_{j=1}^{r-1} D_{\sigma_n}^{-1}\Big)^{-1} \mathbb{1}_{\cap_{i=1}^{n-1}\{A_{\sigma_i} \geq L\}} \mid \Gamma^*_\varrho = y\Big]$$

$$\geq \ \mathbb{E}\Big[\Big((n+1)\max_{r \leq n}\prod_{j=1}^{r-1} (D_{\sigma_n} \wedge R)^{-1}\Big)^{-1} \mathbb{1}_{\cap_{i=1}^{n-1}\{A_{\sigma_i} \geq L\}} \mid \Gamma^*_\varrho = y\Big].$$

This, in turn, implies that

$$\liminf_{n\to\infty}\ \inf_{y\in\Sigma'}\ \frac{1}{n}\ln\mathbb{P}\big(T_{-1} > T_n \mid \Gamma^*_\varrho = y\big)$$

$$\geq \liminf_{n\to\infty}\ \inf_{y\in\Sigma'}\ \frac{1}{n}\ln\mathbb{E}\Big[e^{\min_{t\in(0,1)}\sum_{i=1}^{[nt]}\ln(D_{\sigma_i}\wedge R)}\ \mathbb{1}_{\cap_{i=1}^{n-1}\{A_{\sigma_i} \geq L\}} \mid \Gamma^*_\varrho = y\Big].$$

We now argue much as for (3.19) in the proof of the first part of Theorem 3.6, proving that

$$\liminf_{n\to\infty}\ \inf_{y\in\Sigma'}\ \frac{1}{n}\ln\mathbb{P}\big(T_{-1} > T_n \mid \Gamma^*_\varrho = y\big) \ \geq \ \limsup_{R\to\infty}\ \inf_{\lambda\in[0,1]}\Lambda^{(Q^*_{\ln L, R})}(\lambda)+\ln(1-\eta_L) \ > \ -\ln b.$$

Hence, (4.3) holds, and this ends the proof. ∎

As an example, we consider the case where $B_\nu = p$ is constant for all $\nu \in \mathcal{G}$. Suppose that $L^{-1} = p^{-1} + \varepsilon$ and that $A_{\sigma_i} \in (L, C)$ a.s., for all $\nu \in \mathcal{G}$ and for some constant $C$. Note that then $L < p$ and that $\eta_{L,\infty} = 0$. We prove that $\mathbf{X}$ is transient if $\inf_{\lambda\in[0,1]}\Lambda^{(K^*_{\ln})} > -\ln b$, and recurrent if $\inf_{\lambda\in[0,1]}\Lambda^{(K^*_{\ln})} < -\ln b - \ln(1+\varepsilon)$.

The transience, when $\inf_{\lambda\in[0,1]}\Lambda^{(K^*_{\ln})} > -\ln b$, is a consequence of Theorem 4.4 with $Q^*_{\ln L, \ln C} = K^*_{\ln}$. For the proof of recurrence under the assumption $\inf_{\lambda\in[0,1]}\Lambda^{(K^*_{\ln})} <$

$- \ln b - \ln(1 + \varepsilon)$, we have that

$$
\begin{aligned}
\Phi_n &= \prod_{\substack{i=1 \\ A_{\sigma_i} \le B_{\sigma_i}}}^{n} \left\{ 1 + \frac{(L^{-1} - A_{\sigma_i}^{-1})(1 - q_{i-1}^D)}{1 + A_{\sigma_i}^{-1}(1 - q_{i-1}^D)} \right\} \\
&\le \prod_{\substack{i=1 \\ A_{\sigma_i} \le p}}^{n} (1 + L^{-1} - A_{\sigma_i}^{-1}) \ \le \ \prod_{\substack{i=1 \\ A_{\sigma_i} \le p}}^{n} (1 + \varepsilon) \ \le \ (1 + \varepsilon)^n.
\end{aligned}
$$

Hence

$$
\mathbf{P}(T_{-1} > T_n \mid \Gamma^* = y) \ = \ \mathbb{E}\left[ \Phi_n \prod_{i=1}^{n} q_i^D \ \Big| \ \Gamma^* = y \right] \ \le \ (1 + \varepsilon)^n \cdot \mathbb{E}\left[ \prod_{i=1}^{n} q_i^D \ \Big| \ \Gamma^* = y \right].
$$

This, by Theorem 3.4, using Varadhan's lemma and Proposition 5.1, implies that

$$
\limsup_{n \to \infty} \sup_{y \in \Sigma'} \frac{1}{n} \ln \mathbb{P}(T_{-1} > T_n \mid \Gamma^* = y) \ \le \ \inf_{\lambda \in [0,1]} \Lambda^{(K_{\ln}^*)}(\lambda) + \ln(1 + \varepsilon) \ < \ -\ln b.
$$

The expected number of vertices at level $n$ which are visited before the first return to the origin is bounded above by $b^n \sup_{y \in \Sigma'} \mathbb{P}(T_{-1} > T_n \mid \Gamma^* = y)$. Hence the expected number of vertices visited before the process returns to the origin is bounded by

$$
1 + \sum_{n=1}^{\infty} b^n \sup_{y \in \Sigma'} \mathbb{P}(T_{-1} > T_n \mid \Gamma^* = y) \ < \ \infty.
$$

The latter proves recurrence.

**Example 4.6** *With the situation as above, suppose that $(D_{\sigma_i}, B_{\sigma_i}, A_{\sigma_i})$ evolves as a two state Markov chain. $A_\sigma$ can take the values 1 and $p < 1$, and $L^{-1} = p^{-1} + \varepsilon$, so that $(D_{\sigma_i}, B_{\sigma_i}, A_{\sigma_i})$ has state space $(L, p, p)$ and $(1, p, 1)$. We assume that the diagonal elements of the transition matrix of this process take the value $3/4$. Using (3.16), we have*

$$
\varrho(\lambda) \ = \ \frac{3}{8} \{ 1 + L^\lambda + \sqrt{(1 + L^\lambda)^2 - 32 L^\lambda / 9} \},
$$

*and $\Lambda^{(K_{\ln}^*)}(\lambda) = \ln \rho(\lambda)$. Since $L < p < 1$, this implies that*

$$
\inf_{\lambda \in [0,1]} \Lambda^{(K_{\ln}^*)}(\lambda) = \ln \left\{ \frac{3}{8} \left( 1 + L + \sqrt{(1 + L)^2 - 32 L / 9} \right) \right\}.
$$

23

*Thus, if $b > 4/3$, the process is always transient. However, for $1 < b < 4/3$, the process is recurrent if*

$$\frac{3}{8}\left(1 + L + \sqrt{(1+L)^2 - 32L/9}\right) \; < \; \frac{1}{b(1+\varepsilon)},$$

*and is transient if*

$$\frac{3}{8}\left(1 + L + \sqrt{(1+L)^2 - 32L/9}\right) \; > \; \frac{1}{b}.$$

# 5    Appendix

**Proposition 5.1** *Suppose that $\phi \colon \mathbb{R} \to [-\infty, +\infty]$ is a convex function, with $\phi(0) = 0$. Then*

$$\sup_{f \in \mathcal{AC}} \left\{ \min_{t \in [0,1]} f(t) - \int_0^1 \sup_\lambda \{f'(u)\lambda - \phi(\lambda)\} \, \mathrm{d}u \right\} \;=\; \inf_{\lambda \in [0,1]} \phi(\lambda). \qquad (5.1)$$

**Proof.**    We first prove that the right-hand side of (5.1) is a lower bound. Let $\phi$ be finite on $F \subset \mathbb{R}$, and let $t^* \in [0,1] \cap \overline{F}$ be such that

$$\lim_{\substack{t \to t^* \\ t \in F}} \phi(t) \;=\; \inf_{0 \le t \le 1} \phi(t).$$

Such a $t^*$ exists, in virtue of the convexity of $\phi$. Then, by convexity, $\phi$ has a (non-empty) sub-derivative $\mathrm{SD}(\phi)\{t^*\}$ at $t^*$. Recall that $c \in \mathrm{SD}(\phi)\{a\}$ means that $\phi(t) \ge \phi(a) + c(t-a)$ for all $t$.

If $t^* \in (0,1)$, then $0 \in SD(\phi)\{t^*\}$, and we choose $f(t) = 0$ for all $t$ to get

$$\inf_{\lambda \in \mathbb{R}} \phi(\lambda) \;=\; \inf_{\lambda \in [0,1]} \phi(\lambda)$$

as a lower bound for the left hand side of (5.1).

If $t^* = 0$, then there is a $c \ge 0$ with $c \in \mathrm{SD}(\phi)\{0\}$, so that $\phi(t) \ge \phi(0) + ct$ for all $t$. Take $f(t) = ct$ for all $t$. Since $c \ge 0$, we have $\min_{0 \le t \le 1} f(t) = 0$, and we get

$$-\sup_t \{ct - \phi(t)\} \;\ge\; -\sup_t \{ct - \phi(0) - ct\} \;=\; \phi(0) \;=\; \inf_{\lambda \in [0,1]} \phi(\lambda)$$

as a lower bound for the left hand side of (5.1).

If $t^* = 1$, then there is a $c \leq 0$ with $c \in \mathrm{SD}(\phi)\{1\}$, so that $\phi(t) \geq \phi(1) + c(t-1)$ for all $t$. Take $f(t) = ct$. As $c \leq 0$, we have $\min_{0 \leq t \leq 1} f(t) = c$, and we get

$$c - \sup_t \{ct - \phi(t)\} \; \geq \; c - \sup_t \{ct - \phi(1) - c(t-1)\} \; = \; c + \phi(1) - c \; = \; \inf_{\lambda \in [0,1]} \phi(\lambda)$$

as a lower bound for the left hand side of (5.1).

Next we turn to the proof of the upper bound. Fix any $t^* \in [0,1]$. Notice that, for any $f \in \mathcal{AC}$, we have $\min_{t \in [0,1]} f(t) \leq 0$ and $\big(f(1) - \min_{t \in [0,1]} f(t)\big) \geq 0$. Hence, taking $\lambda = t^*$ for all $u \in [0,1]$, the left-hand side of (5.1) is bounded above by

$$\sup_{f \in \mathcal{AC}} \Big\{ \min_{t \in [0,1]} f(t) - f(1)t^* + \phi(t^*) \Big\}$$
$$= \sup_{f \in \mathcal{AC}} \Big\{ \min_{t \in [0,1]} f(t)(1 - t^*) - \big(f(1) - \min_{t \in [0,1]} f(t)\big)t^* + \phi(t^*) \Big\} \; \leq \; \phi(t^*).$$

By taking the infimum over $t^* \in [0,1]$ we have the upper bound. ∎

# References

[1] A. Collevecchio (2006). On the transience of processes defined on Galton-Watson trees. *Ann. Probab.* **34**, 870–878.

[2] J. J. Dai (2005). A once edge-reinforced random walk on a GaltonWatson tree is transient. *Statist. Probab. Lett.* **73**, 115–124.

[3] A. Dembo & T. Zajic (1995). Large deviations: from empirical mean and measure to partial sums process, *Stoch. Procs. Applics.* **57**, 191–224.

[4] A. Dembo & O. Zeitouni (1998). *Large deviations techniques and applications*, 2nd edition. Springer, Berlin.

[5] J.-D. Deuschel & D. W. Stroock (1989). *Large deviations*. Academic press, Boston.

[6] R. Durrett, H. Kesten & V. Limic (2002). Once edge-reinforced random walk on a tree. *Probab. Theory Rel. Fields* **122**, 567–592.

[7] G. Faraud (2011). A central limit theorem for random walk in a random environment on marked Galton–Watson trees. *Electronic J. Probab.* **16**, 174–215.

[8] J. G. Kemeny, J. L. Snell & A. W. Knapp (1966). *Denumerable Markov chains.* van Nostrand, Princeton, NJ.

[9] D. Kious & V. Sidoravicius (2016). Phase transition for the once-reinforced random walk on $\mathbb{Z}_d$-like trees. `arXiv:1604.07631`

[10] R. Lyons & R. Pemantle (1992). Random walk in a random environment and first–passage percolation on trees . *Ann. Probab.* **20**, 125–136.

[11] R. Lyons & Y. Peres (2005). *Probability on trees and networks.* Cambridge Series in Statistical and Probabilistic Mathematics (to appear).

[12] M. V. Menshikov & D. Petritis (2002). On random walks in random environment on trees and their relationship with multiplicative chaos. *Mathematics and computer science II (Versailles, 2002)*, 415–422.

[13] P. Ney & E. Nummelin (1987). Markov additive processes I: Eigenvalue properties and limit theorems. *Ann. Probab.* **15**, 561–592.

[14] P. Ney & E. Nummelin (1987). Markov additive processes II: Large deviations. *Ann. Probab.* **15**, 593–609.

[15] R. Pemantle (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4**, 1–79.

[16] A. Sznitman (2002). Topics in random walks in random environment. In *School and Conference in Probability Theory*, ICTP Lect. Notes, XVII 203–266 (electronic). Abdus Salam Int. Cent. Theoret. Phys., Trieste. Available at `http://users.ictp.trieste.it/ pub_off/lectures/lns017/Sznitman/Sznitman.pdf`