Near critical preferential attachment networks have small giant components

Maren Eckhoff[*]   Peter Mörters[†]   Marcel Ortgiese[‡]

December 1, 2017

Abstract

Preferential attachment networks with power law exponent $\tau > 3$ are known to exhibit a phase transition. There is a value $\rho_c > 0$ such that, for small edge densities $\rho \leq \rho_c$ every component of the graph comprises an asymptotically vanishing proportion of vertices, while for large edge densities $\rho > \rho_c$ there is a unique giant component comprising an asymptotically positive proportion of vertices. In this paper we study the decay in the size of the giant component as the critical edge density is approached from above. We show that the size decays very rapidly, like $\exp(-c/\sqrt{\rho - \rho_c})$ for an explicit constant $c > 0$ depending on the model implementation. This result is in contrast to the behaviour of the class of rank-one models of scale-free networks, including the configuration model, where the decay is polynomial. Our proofs rely on the local neighbourhood approximations of [6] and recent progress in the theory of branching random walks [8].

Key words: Power law, small world, scale-free network, preferential attachment, Barabási-Albert model, graph components, percolation, killed branching random walk, survival probability.

Mathematics Subject Classification (2010): Primary 05C80; Secondary 60J85, 60J80, 90B15

1   Introduction and main results

1.1   Introduction

Sparse random graph models typically undergo a phase transition in their connectivity behaviour depending on the mean number of edges per vertex. A typical case are the Erdős-Renyi graphs with n vertices and $m = m(n)$ edges and asymptotic edge density $\rho = \lim \frac{m(n)}{n}$. There exists a critical density $\rho_c = 1$ such that if the edge density satisfies $\rho \leq \rho_c$ the largest component in the graph comprises a vanishing proportion of vertices, whereas for $\rho > \rho_c$ this proportion $\theta(\rho)$ is strictly positive. The behaviour of $\theta(\rho)$ as $\rho \downarrow \rho_c$ is characterised by an exponent β defined by

$$\theta(\rho) \sim (\rho - \rho_c)^\beta, \qquad \text{as } \rho \downarrow \rho_c,$$

---

[*]QuantumBlack, Kinnaird House, 1 Pall Mall East, London SW1Y 5AU, United Kingdom.
Email: `eckhoff.maren@gmail.com`

[†]Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany.
Email: `moerters@math.uni-koeln.de`

[‡]Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, United Kingdom.
Email: `m.ortgiese@bath.ac.uk`

which is $\beta = 1$ in the Erdős-Renyi case. A natural extension of the Erdős-Renyi model is the configuration model, which allows to construct random graphs with n vertices and a given degree sequence $m_1, \ldots, m_n$. Of particular interest is the case when the degree sequence is heavy-tailed,

$$\frac{1}{n}\#\{1 \leq j \leq n \colon m_j = k\} \stackrel{n \to \infty}{\longrightarrow} \mu(k) = k^{-\tau + o(1)}, \quad \text{as } k \to \infty,$$

where the parameter $\tau > 2$ is the power-law exponent. The connectivity behaviour of the Erdős-Renyi model persists with $\beta = 1$ for the configuration model with $\tau > 4$, and with $\beta = 1/(\tau - 3)$ if $3 < \tau < 4$, see Cohen et al. [4]. The paper by Cohen is based on an informal approximation of local neighbourhoods in the graph by Galton-Watson trees and thereby extends to a wide range of scale-free network models, where similar approximations hold.

Cohen et al. [4] claim their result for scale-free networks in general without specifying a model. This reflects the belief that the behaviour observed in the configuration model extends to all natural scale-free network models including the class of preferential attachment networks. In the present paper however we show for the first time that preferential attachment networks have a qualitatively completely different behaviour than predicted in [4]. In fact, for all $\tau > 3$ the size of the giant component is decaying exponentially as one approaches the critical edge density. More precisely, we show that the relative size of the giant component in preferential attachment networks is

$$\theta(\rho) = \exp\left(\frac{-c + o(1)}{\sqrt{\rho - \rho_c}}\right), \quad \text{as } \rho \downarrow \rho_c,$$

where c is an explicit constant depending on the way in which the edge density is controlled. This demonstrates once again that preferential attachment networks belong to a different universality class than the configuration model and other models based on rank-one connection probabilities.

The underlying phenomenon of the 'small giant component' or 'slow emergence of the giant' has first been discovered and discussed by Oliver Riordan in the seminal paper [16], and in collaboration with Bollobás and Janson in [2] and [3]. Riordan [16] finds that for the original Barabasi-Albert model subject to Bernoulli percolation with retention parameter p, one has

$$\theta(p) = \exp\left(-\frac{c_m + o(1)}{p}\right) \quad \text{as } p \downarrow 0, \quad \text{for } c_m = \frac{1}{2}\log\left(\frac{m+1}{m-1}\right), \tag{1.1}$$

where $m \geq 2$ is the outdegree of all vertices. As this model corresponds to the critical case $\tau = 3$ this is not at odds with the results of Cohen et al. [4]. The merit of our work is to extend the phenomenon of slow emergence to a regime where it is most surprising because it defies the predictions in [4].

In his proof, Riordan exploits the fact that there are local approximations of the network by multitype branching processes whose survival probability can be studied analytically by looking at the associated Laplace operators. These branching processes are more complex than those used in [4].

In the present paper local approximation is by a branching random walk with killing. Instead of the analytical techniques used by Riordan, which we could not apply effectively in our case due to the higher complexity of the approximating process, we use geometric properties of killed branching random walks. Studying their survival probabilities requires sophisticated techniques from the theory of branching random walks, which became available only in the past few years, see in particular [8]. Our results are therefore pleasingly probabilistic and highly timely.

1.2 Statement of results

We start by describing the preferential attachment network introduced in [5], which gives scale-free networks with arbitrary power law exponent $\tau > 2$ by variation of a parameter.

A concave function $f \colon \mathbb{N}_0 \to (0, \infty)$ is called an *attachment rule* if $f(0) \leq 1$ and

$$\triangle f(k) := f(k+1) - f(k) < 1 \text{ for all } k \in \mathbb{N}_0.$$

2

The maximal increment is denoted by $\gamma^+ := \sup\{\triangle f(k) \colon k \in \mathbb{N}_0\}$. By concavity, f is non-decreasing, $\gamma^+ = f(1) - f(0)$ and the limit $\gamma := \lim_{k \to \infty} f(k)/k$ exists and equals $\gamma = \inf\{\triangle f(k) \colon k \in \mathbb{N}_0\}$. Given an attachment rule f, we define a growing sequence $(\mathsf{G}_n \colon n \in \mathbb{N})$ of random graphs as follows

- Start with the graph G_1 given by one vertex labelled 1 and no edges;

- Given the graph G_n, we construct G_{n+1} from G_n by adding a new vertex labelled $n + 1$ and, for each $m \leq n$ independently, inserting the directed edge $(n + 1, m)$ with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n}.$$

Formally, we are dealing with a sequence of directed graphs but all edges point from the younger to the older vertex. Hence, directions can be recreated from the undirected, labelled graph. For all structural questions, particularly regarding connectivity and the length of shortest paths, we regard $(\mathsf{G}_n \colon n \in \mathbb{N})$ as an undirected network. It is shown in [5] that when $\gamma > 0$ the networks have a degree distribution which is a power law with exponent $\tau = 1 + 1/\gamma$. We are also interested in the percolated version of the network $(\mathsf{G}_n \colon n \in \mathbb{N})$. For $p \in [0, 1]$, we write $\mathsf{G}_n(p)$ for the graph obtained from G_n by deleting each edge with probability $1 - p$ independent of all other edges.

Let $(\mathsf{G}_n \colon n \in \mathbb{N})$ be a sequence of (random or deterministic) graphs, where G_n has n vertices. For $n \in \mathbb{N}$, we denote by $|\mathsf{C}_n|$ the size of the largest component in G_n. The network $(\mathsf{G}_n \colon n \in \mathbb{N})$ has a *giant component* if there exists a constant $\theta > 0$ such that

$$\frac{|\mathsf{C}_n|}{n} \to \theta \qquad \text{as } n \to \infty,$$

where the convergence holds in probability. The limit θ is called the *size of the giant component.*

Dereich and Mörters showed in [6, Theorem 1.6] that when $\gamma \geq \frac{1}{2}$, or equivalently $2 < \tau \leq 3$, then $(\mathsf{G}_n(p) \colon n \in \mathbb{N})$ has a giant component for all $p \in (0, 1]$. When $\gamma < \frac{1}{2}$, or equivalently $\tau > 3$, then there exists a critical percolation parameter $p_\mathrm{c} > 0$ such that $(\mathsf{G}_n(p) \colon n \in \mathbb{N})$ has giant component if and only if $p > p_\mathrm{c}$. We denote the relative size of the giant component in $(\mathsf{G}_n(p) \colon n \in \mathbb{N})$ by $\theta(p, f)$ and omit p or f from the notation when the percolation parameter or the attachment rule are fixed.

We are interested in the decay in the size of the giant component as we approach p_c from above in the case $p_\mathrm{c} > 0$, or equivalently $\gamma < \frac{1}{2}$. It was shown in Lemma 3.3 of [6] that the critical retention probability for the network $(\mathsf{G}_n \colon n \in \mathbb{N})$ is given by

$$p_\mathrm{c} = \rho(\alpha^*)^{-1},$$

where $\rho(\cdot)$ is the spectral radius of the score operator, a function on a nonempty open interval, which we describe explicitly in Section 2, and α^* is the minimizer of ρ.

Our first result shows the exponential decay of the size of the giant component of the percolated network, when the retention parameter approaches the critical value.

Theorem 1.1. *Let f be an attachment rule with $\gamma < \frac{1}{2}$ and $\theta(1, f) > 0$. Then*

$$\lim_{p \downarrow p_\mathrm{c}} \sqrt{p - p_\mathrm{c}} \log \theta(p, f) = -\sqrt{\frac{\pi^2 \rho''(\alpha^*)}{2}} \alpha^* p_\mathrm{c}.$$

An alternative way of reducing the edge density and thereby destroying the giant component is to alter the attachment rule instead of percolating the network. For a linear attachment rule $f(k) = \gamma k + \beta$ with $\gamma < \frac{1}{2}$, Dereich and Mörters [6] show that there exists a giant component if and only if

$$\beta > \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma} =: \beta_\mathrm{c}(\gamma) = \beta_\mathrm{c}.$$

Therefore, one could fix γ and decrease β to β_c. Another idea would be, for a given β, to decrease γ until $\beta = \beta_\mathrm{c}(\gamma)$. To analyse the behaviour of the size of the giant component under this procedure, let $(f_t)_{t \geq 0}$

3

be a sequence of attachment rules with $\gamma_t := \inf\{\triangle f_t(k) : k \in \mathbb{N}_0\} < \frac{1}{2}$ for all $t \geq 0$. We denote by ρ_t and α_t^* the spectral radius and its minimizer corresponding to the score operator for the unpercolated branching random walk derived from attachment rule f_t.

Theorem 1.2. *Let $(f_t)_{t \geq 0}$ be a pointwise decreasing sequence of attachment rules with $\gamma_t < \frac{1}{2}$ for all $t \geq 0$ and pointwise limit f. Suppose that $\theta(1, f_t) > 0$ for all t and $\theta(1, f) = 0$. Then*

$$\lim_{t \to \infty} \sqrt{\log \rho_t(\alpha_t^*)} \log \theta(1, f_t) = -\sqrt{\frac{\pi^2 \rho''(\alpha^*)}{2}} \alpha^*,$$

where α^* and $\rho''(\alpha^*)$ are derived from f.

The existence of ρ and α^* corresponding to f is proved in Proposition 2.3. There we will also see that $\lim_{t \to \infty} \rho_t(\alpha_t^*) = 1$. The following corollary exemplifies Theorem 1.2 for linear attachment rules. We denote

$$\beta_{\mathrm{c}}(\gamma) = \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma} \quad \text{and} \quad \gamma_{\mathrm{c}}(\beta) = \frac{1}{2}\big(1 - \beta - \sqrt{\beta^2 + 2\beta}\big).$$

Corollary 1.3 (Linear attachment rule). *Let $\gamma \in [0, \frac{1}{2})$, $\beta \in (0, 1]$. Then*

$$\lim_{\beta \downarrow \beta_c(\gamma)} \sqrt{\beta - \beta_c(\gamma)} \log \theta(1, \gamma \cdot + \beta) = -\frac{\pi}{2\sqrt{1 - \gamma}}. \tag{1.2}$$

If $\beta \in (0, 1/4]$, then

$$\lim_{\gamma \downarrow \gamma_c(\beta)} \sqrt{\gamma - \gamma_c(\beta)} \log \theta(1, \gamma \cdot + \beta) = -\frac{\pi}{2(\beta^2 + 2\beta)^{1/4}}. \tag{1.3}$$
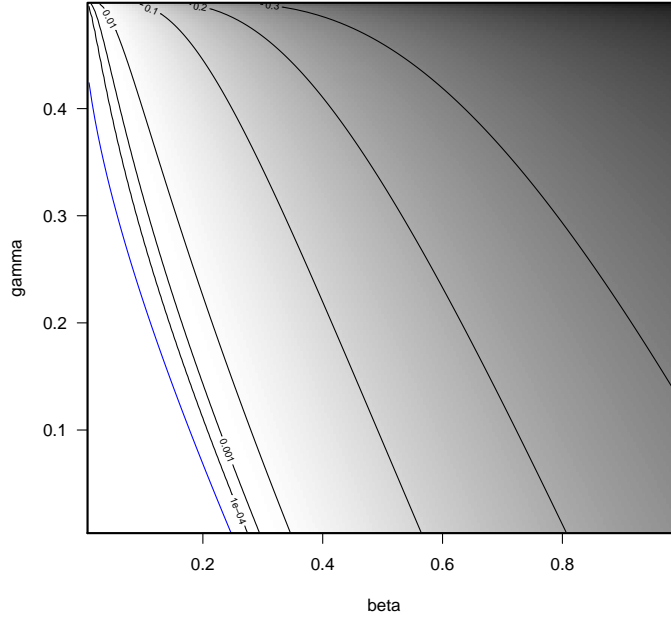

Figure 1: Theorem 1.2 allows us to approach the critical line from above, from the right and from any angle in between.

Remark 1.4. Two cases in our phase diagram are covered in the work of Riordan [16]. The first of these cases corresponds to an approximation from the right of the point $\beta = 0, \gamma = \frac{1}{2}$ which is equivalent to the original Barabasi-Albert, or LCD, model.[1] Note that our results refer to the subcritical case $\gamma < \frac{1}{2}$ and the critical case $\gamma = \frac{1}{2}$ is not included. The second is the case $\beta = \frac{1}{4}, \gamma = 0$, the Dubins model, in which there is no preferential attachment and our results are consistent with those of Riordan [16].

Corollary 1.3 allows a quantitative comparison of the decay of the giant component for different models. The smaller γ (or the larger τ), the slower is the decay. The LCD model, or equivalent models with $\gamma = \frac{1}{2}$, have faster decay of the size of the giant component than preferential attachment networks with attachment rules satisfying $\gamma < \frac{1}{2}$.

Throughout, we will use the following notation: For every integer $k \in \mathbb{N}$, we write $[k] = \{1, \ldots, k\}$, $\mathbb{R}_+ := [0, \infty)$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

The remaining paper is structured as follows: We prove Theorems 1.1 and 1.2 simultaneously. In Section 2 we collect several auxiliary results that we need later on. In particular, in Section 2.1 we recall the relevant results from [6] who relate the size of the giant component to the survival probability of a multitype branching random walk with killing. In Section 2.2 we derive the main tool to analyse these branching random walks, a version of the well-known many-to-one lemma. In Section 2.3 we collect various moment estimates, while in Section 2.4 we state a large deviation result originally due to Mogulskii in a suitable adaptation to our setting. The actual proofs of Theorems 1.1 and 1.2 are split into an upper bound carried out in Section 3 and a lower bound in Section 4. In Section 5 we show how to derive Corollary 1.3 from Theorem 1.2. Finally, in Appendix A we prove the large deviation result, Theorem 2.10.

2 Proofs: General Preparations

2.1 The approximating branching process

We introduce a pure jump Markov process with generator

$$Lg(k) := f(k)\big(g(k+1) - g(k)\big).$$

This defines an increasing, integer-valued process, which jumps from k to $k+1$ after an exponential waiting time with mean $1/f(k)$, independently of the previous jumps. Under the probability P we denote by $(Z_t : t \geq 0)$ the process started in zero, by $(\hat{Z}_t : t \geq 0)$ the process started in one, and by $(Z_t^{(\tau)} : t \geq 0)$ the process started in zero conditioned to have a point at $\tau \geq 0$.

This process is used to define a multitype branching random walk with type space $\mathcal{T} := [0, \infty) \cup \{\ell\}$, where ℓ is a non-numerical symbol for 'left'. A particle in location $x \in \mathbb{R}$ and of type $\tau \in \mathcal{T}$, produces offspring to its left whose displacements have the same distribution as the points of the Poisson point process with intensity measure

$$e^t E[f(Z_{-t})] \mathbb{1}_{(-\infty, 0]}(t) \, dt.$$

The type of an offspring on the left equals the distance to its parent.

The distribution of the offspring to the right depends on the type of the particle. When the particle is of type ℓ, then the relative positions of its right offspring follow the same distribution as the jump times of $(Z_t : t \geq 0)$. When the particle is of type $\tau \geq 0$, then the displacements follow the same distribution as the jump times of $(Z_t^{(\tau)} - \mathbb{1}_{[\tau, \infty)}(t) : t \geq 0)$. All offspring on the right are of type ℓ.

The offspring to the right do *not* form a Poisson point process. The more particles are born, the higher the rate of new particles arriving. Moreover, the total number of particles produced is infinite without accumulation point. The expected distance between a particle and its kth offspring on the right equals

---

[1]There is a difference in the set-up of the LCD model and our model, as the former uses a fixed number of connections for every new vertex. This technical difference does not affect the results beyond the form of the constants involved.

$\sum_{j=0}^{k-1} \frac{1}{f(j)} = \sum_{j=0}^{k-1} \frac{1}{f(j)/j} \frac{1}{j}$. Since $\lim_{j\to\infty} \frac{f(j)}{j} = \gamma$, this distance behaves asymptotically like $\gamma^{-1} \log(k)$ when $\gamma \neq 0$ and like k when $\gamma = 0$.

We call the described process *idealized branching random walk* (IBRW) in accordance with [6]. Dereich and Mörters [6] show that the genealogical tree of the IBRW is related to the local neighbourhood of a vertex in G_n. To obtain a branching process approximation to $(\mathsf{G}_n(p)\colon n \in \mathbb{N})$, we define the *percolated* IBRW by associating to every offspring in the IBRW an independent Bernoulli(p) random variable. If the random variable is zero, we delete the offspring together with its descendants. Otherwise, the offspring is retained in the percolated IBRW. When the percolated IBRW is started with one particle in location x and type τ, then we write $P^p_{(x,\tau)}$ for its distribution and $E^p_{(x,\tau)}$ for the corresponding integral operator; $P_{(x,\tau)} := P^1_{(x,\tau)}$, $E_{(x,\tau)} := E^1_{(x,\tau)}$.

The percolated IBRW can be interpreted as a labelled tree Γ where every node represents a particle and is connected to its children and (apart from the root) to its parent. The vertices are identified as finite sequences of natural numbers $x = j_1 \ldots j_k$, including the empty sequence $\emptyset$ which denotes the root. We concatenate sequences $x = i_1 \ldots i_k$ and $y = j_1 \ldots j_m$ to form the sequence $xy = i_1 \ldots i_k j_1 \ldots j_m$. When, for $j \in \mathbb{N}$, xj is a vertex in Γ, then $x, x1, \ldots, x(j-1)$ are also vertices in the tree and $x =: \mathrm{p}(xj)$ is the parent of xj. The length $|x| = k$ is the generation of x. For $|x| \geq k$, we abbreviate $\mathrm{p}^k(x)$ for the k-fold composition of $\mathrm{p}(\cdot)$. The ancestor of x in generation k is denoted by x_k, i.e. $x_k := \mathrm{p}^{n-k}(x)$ when $|x| = n$. In particular, x_0 always denotes the root. To every vertex we associate two functions, S and τ. Here $S(x)$ is the location of the particle on the real line and $\tau(x)$ denotes its type.

To obtain a branching process approximation to the local neighbourhood of a vertex in $(\mathsf{G}_n(p)\colon n \in \mathbb{N})$, we consider the percolated IBRW with a killing barrier at zero. That is, every particle with location on the nonnegative half-line is deleted together with its descendants. Dereich and Mörters prove the following identification.

Theorem 2.1 (**Dereich and Mörters** [6])**.** *For all $p \in [0,1]$ and attachment rules f, $\theta(p,f)$ equals the survival probability of the percolated* IBRW *with a killing barrier at zero, started with one particle of type ℓ whose location is given by $-E$, where E is an exponential random variable with mean one.*

Next we collect some spectral properties that will be used in the analysis of the IBRW.

Denote by $C(\mathcal{T})$ the Banach space of bounded, continuous functions on $\mathcal{T}$ equipped with the supremum norm $\|\cdot\|$. For $\alpha \in (0,1)$, we consider the score operator

$$A^p_\alpha g(\tau_0) = E^p_{(0,\tau_0)}\Big[\sum_{|x|=1} g(\tau(x)) e^{-\alpha S(x)} \Big], \qquad \tau_0 \in \mathcal{T}, \tag{2.1}$$

on $C(\mathcal{T})$, $A_\alpha := A^1_\alpha$. The spectral radius of A^p_α is denoted by $\rho^p(\alpha)$. The dependence of A^p_α on the attachment rule f is suppressed in notation but it will always be clear from the context which f is considered. Since by definition,

$$A^p_\alpha g(\tau_0) = E^p_{(0,\tau_0)}\Big[\sum_{|x|=1} g(\tau(x)) e^{-\alpha S(x)} \Big] = p E_{(0,\tau_0)}\Big[\sum_{|x|=1} g(\tau(x)) e^{-\alpha S(x)} \Big] = p A_\alpha g(\tau_0), \tag{2.2}$$

it suffices to analyse A_α. We write $\mathbf{1}$ for the constant function with value 1 and let $\mathcal{I} := (\gamma, 1-\gamma)$ for $\gamma < \frac{1}{2}$ and $\mathcal{I} = \emptyset$ for $\gamma \geq \frac{1}{2}$.

Lemma 2.2. *Let $p \in (0,1]$. If $\gamma \geq \frac{1}{2}$, then $A^p_\alpha \mathbf{1}(0) = \infty$. If $\gamma < \frac{1}{2}$, then the following holds:*

(i) *$\rho^p(\alpha)$ is finite for all $\alpha \in \mathcal{I}$, $\rho^p(\alpha) = p\rho(\alpha)$ and $\rho(\alpha) \to \infty$ for $\alpha \to \partial\mathcal{I}$.*

(ii) *There exists a unique positive eigenfunction v_α of A^p_α corresponding to $\rho^p(\alpha)$ with $\|v_\alpha\| = 1$. Moreover, v_α does not depend on the retention probability p and $\min_{\tau \in \mathcal{T}} v_\alpha(\tau) > 0$.*

(iii) *The function ρ is twice differentiable on $\mathcal{I}$ with*

$$\rho^{(i)}(\alpha) = E_{(0,\tau_0)}\Big[\sum_{|x|=1} \frac{v_\alpha(\tau(x))}{v_\alpha(\tau_0)} e^{-\alpha S(x)} (-S(x))^i \Big] \qquad \textit{for all } \tau_0 \in \mathcal{T}, i \in \{1,2\}. \tag{2.3}$$

6

(iv) *The function ρ is strictly convex on $\mathcal{I}$ and there exists a unique minimizer $\alpha^* \in \mathcal{I}$.*

(v) *For any $\tau \in [0, \infty)$*
$$v_\alpha(\ell) \le v_\alpha(\tau) \le v_\alpha(0).$$

Proof. By (2.2), it suffices to prove Lemma 2.2 in the case $p = 1$. For that case, it was shown in [6, Lemma 3.1] that $A_\alpha 1(0) < \infty$ is equivalent to A_α being a strongly positive, compact operator with $A_\alpha g \in C(\mathcal{T})$ for all $g \in C(\mathcal{T})$. Moreover, it is proved that for all $g \in C(\mathcal{T})$, $g \ge 0$, we have $\underline{A}_\alpha g \le A_\alpha g \le \overline{A}_\alpha g$, where

$$\underline{A}_\alpha = \begin{pmatrix} \mathsf{a}(\alpha) & \mathsf{a}(1-\alpha) \\ \mathsf{a}(\alpha) & \mathsf{a}(1-\alpha) \end{pmatrix}, \qquad \overline{A}_\alpha = \begin{pmatrix} \mathsf{a}(\alpha) & \mathsf{a}(1-\alpha) \\ \mathsf{c}(\alpha) & \mathsf{a}(1-\alpha) \end{pmatrix},$$

and

$$\mathsf{a}(\alpha) = \int_0^\infty e^{-\alpha t} E[f(Z_t)]\, dt, \qquad \mathsf{c}(\alpha) = \int_0^\infty e^{-\alpha t} E[f(\hat{Z}_t)]\, dt.$$

Here, for example, $\overline{A}_\alpha g$ means that for all $\tau \ge 0$,

$$\overline{A}_\alpha g(\tau) = \int_0^\infty g(\ell) e^{-\alpha t} E[f(\hat{Z}_t)]\, dt + \int_0^\infty g(t) e^{-(1-\alpha)t} E[f(Z_t)]\, dt.$$

We have, $\rho(\alpha) \ge \rho(\underline{A}_\alpha) = \mathsf{a}(\alpha) + \mathsf{a}(1-\alpha)$. The values of a and c are identified as (cf. proof of Proposition 1.10 in [6])

$$\mathsf{a}(\alpha) = \sum_{k=0}^\infty \prod_{j=0}^k \frac{f(j)}{f(j)+\alpha}, \qquad \mathsf{c}(\alpha) = \sum_{k=0}^\infty \prod_{j=0}^k \frac{f(j+1)}{f(j+1)+\alpha}. \tag{2.4}$$

We analyse the convergence properties of $\mathsf{a}(\alpha)$. Using $\log(1+x) \le x$ for $x \ge 0$ and $f(j) = \sum_{i=0}^{j-1} \triangle f(i) + f(0) \ge j\gamma$, we estimate

$$\prod_{j=0}^k \frac{f(j)}{f(j)+\alpha} = \exp\left(-\sum_{j=0}^k \log\left(1+\frac{\alpha}{f(j)}\right)\right) \ge \exp\left(-\sum_{j=0}^k \frac{\alpha}{f(j)}\right)$$

$$\ge \exp\left(-\sum_{j=1}^k \frac{\alpha}{j\gamma} - \frac{\alpha}{f(0)}\right) = \exp\left(-\frac{\alpha}{\gamma}\log k - C - \delta_k\right)$$

for some $C > 0$ and a null sequence $(\delta_k)_{k \in \mathbb{N}}$. In particular, there exists a $C' > 0$ such that for all $k \in \mathbb{N}$

$$\prod_{j=0}^k \frac{f(j)}{f(j)+\alpha} \ge C'\left(\frac{1}{k}\right)^{\alpha/\gamma}.$$

Hence, $\mathsf{a}(\alpha) = \infty$ for all $\alpha \le \gamma$. On the other hand, by Cauchy's condensation test, $\mathsf{a}(\alpha) < \infty$ for all $\alpha > \gamma$. Hence, $A_\alpha 1(0) < \infty$ if and only if $\alpha \in (\gamma, 1-\gamma) =: \mathcal{I}$. From now on we assume that $\mathcal{I} \ne \emptyset$, i.e. $\gamma < \frac{1}{2}$. For $\alpha \in \mathcal{I}$, $\rho(\alpha)$ is finite and since $\rho(\alpha) \ge \mathsf{a}(\alpha) + \mathsf{a}(1-\alpha)$, we see that $\rho(\alpha) \to \infty$ for $\alpha \to \partial\mathcal{I}$. The existence and uniqueness of eigenfunction $v_\alpha \colon \mathcal{T} \to (0, \infty)$ follows from the Krein-Rutman theorem (see Theorem 3.1.3 in [15]). The fact that A_α is a strictly positive operator, implies $\min_{\tau \in \mathcal{T}} v_\alpha(\tau) > 0$. Since $\rho(\alpha)$ is an isolated eigenvalue with one-dimensional eigenspace, one can argue along the lines of Chapter II §3 (in particular Remark 2.4) and Theorem II.§5.5.4 of [11] that ρ is twice differentiable and the derivative can be represented as in (2.3). In particular, $\rho''(\alpha) > 0$ for all $\alpha \in \mathcal{I}$, hence, ρ is strictly convex on $\mathcal{I}$ and there exists a unique minimizer $\alpha^* \in \mathcal{I}$.

To prove (v), consider $\tau \ge 0$. Let $\hat{Z}_t^{(\tau)} = Z^{(\tau)} - \mathbb{1}_{[\tau,\infty)}$ for any $\tau \in [0, \infty)$. Then, by the definition of the

7

eigenfunction,

$$
\begin{aligned}
\rho(\alpha)v_\alpha(\tau) = A_\alpha v_\alpha(\tau) &= E_{(0,\tau)}\Big[\sum_{|x|=1} v_\alpha(\tau(x))e^{-\alpha S(x)} \Big] \\
&= E_{(0,\tau)}\Big[\sum_{|x|=1} v_\alpha(-S(x))e^{-\alpha S(x)}\mathbb{1}_{\{S(x)\le 0\}} \Big] + v_\alpha(\ell)E\Big[\int_0^\infty e^{-\alpha t}\,d\hat{Z}_t^{(\tau)} \Big] \\
&\ge E_{(0,\ell)}\Big[\sum_{|x|=1} v_\alpha(-S(x))e^{-\alpha S(x)}\mathbb{1}_{\{S(x)\le 0\}} \Big] + v_\alpha(\ell)E\Big[\int_0^\infty e^{-\alpha t}\,d\tilde{Z}_t \Big] \\
&= A_\alpha v_\alpha(\ell) = \rho(\alpha)v_\alpha(\ell),
\end{aligned}
$$

where we used in the inequality that the distribution of the positions to the left of the origin do not depend on the initial type and for the second expectation we used the monotonicity in types proved in [6]. The upper bound holds by a similar argument. $\qquad\square$

The next proposition collects some of the consequences for the spectral radius if we consider converging attachment rules.

Proposition 2.3. *Let $(f_t)_{t\ge 0}$ be pointwise decreasing attachment rules with $\gamma_t = \inf_{k\ge 0}\frac{f_t(k)}{k} < 1/2$. Define*

$$
f(k) := \lim_{t\to\infty} f_t(k).
$$

Then, f is concave, $f(0) \le 1$ and $f(k+1) - f(k) \le 1$ for all $k \in \mathbb{N}_0$. Let $\gamma = \inf_{k\ge 0}\frac{f(k)}{k}$ and let $\rho(\alpha)$ be the spectral radius of the operator associated to the branching process with attachment rule f, α^* be its unique minimizers and let v_α be the corresponding eigenfunction with $\|v_\alpha\| = 1$. Define the same quantities with index t when referring to the branching process associated to f_t, where we set $\rho_t(\alpha) = \infty$ if $t \notin (\gamma_t, 1 - \gamma_t)$.

(i)
$$
\rho(\alpha) = \lim_{t\to\infty} \rho_t(\alpha) \ \text{for all } t \in (\gamma, 1 - \gamma).
$$

(ii) Suppose that $\lim_{t\to\infty} \rho_t(\alpha_t^*) = 1$, then $\alpha_t^* \to \alpha^*$ as $t \to \infty$ and $\rho(\alpha^*) = 1$;

(iii) The quotient
$$
\frac{v_{t,\alpha_t^*}(0)}{v_{t,\alpha_t^*}(\ell)}
$$
is uniformly bounded in t away from zero and infinity.

(iv) Moreover, as $t \to \infty$,
$$
\rho_t''(\alpha_t^*) \to \rho''(\alpha^*).
$$

Proof. Since $(f_t)_{t\ge 0}$ is a pointwise decreasing sequence of positive functions

$$
f(k) := \lim_{t\to\infty} f_t(k), \qquad \text{for all } k \in \mathbb{N}_0,
$$

exists. As an infimum of concave functions, $f\colon \mathbb{N}_0 \to [0,\infty)$ is concave. The property $f(0) \le 1$ is inherited from f_t. The increments might in general only satisfy $f(k+1) - f(k) \le 1$, but the strict inequality is not needed for the analysis of the branching process and the corresponding operators.

The assumption $\rho_t(\alpha_t^*) > 1$ implies that there exists a giant component for all $t \ge 0$. Hence,

$$
\mathsf{a}_{f_t}(1/2) + \sqrt{\mathsf{a}_{f_t}(1/2)\mathsf{c}_{f_t}(1/2)} > 1 \quad \forall t \ge 0,
$$

by Proposition 1.10 in [6]. Here a_{f_t} and c_{f_t} are the functions given in (2.4). By monotone convergence, $\mathsf{a}_f(1/2) + \sqrt{\mathsf{a}_f(1/2)\mathsf{c}_f(1/2)} \ge 1$ and, in particular, $f(0) > 0$. Hence, f is an attachment rule. From now

on we add a subscript t to all quantities corresponding to f_t and no subscript for quantities corresponding to f.

The assumption $\gamma_t < \frac{1}{2}$ is needed to make the operator $A_{t,\alpha}$ exist for some α. We have for all $t \leq s$

$$\gamma := \lim_{k \to \infty} \frac{f(k)}{k} \leq \lim_{k \to \infty} \frac{f_s(k)}{k} \leq \lim_{k \to \infty} \frac{f_t(k)}{k} = \gamma_t < \frac{1}{2}.$$

In particular, $(\gamma_t)_{t \geq 0}$ is a non-increasing sequence. Let $\mathcal{I} = (\gamma, 1 - \gamma)$ and $\mathcal{I}_t = (\gamma_t, 1 - \gamma_t)$ for $t \geq 0$. Then $\mathcal{I}_t \subseteq \mathcal{I}$ for all $t \geq 0$ and we write $\rho_t(\alpha) = \infty$ whenever $\alpha \notin \mathcal{I}_t$.

Let $\alpha \in \mathcal{I}$. We use the monotonicity of the branching process in the attachment rule to derive

$$\rho(\alpha) = \inf_{n \in \mathbb{N}} \|A_\alpha^n 1\|^{\frac{1}{n}} = \inf_{n \in \mathbb{N}} \inf_{t \geq 0} \|A_{t,\alpha}^n 1\|^{\frac{1}{n}} = \inf_{t \geq 0} \inf_{n \in \mathbb{N}} \|A_{t,\alpha}^n 1\|^{\frac{1}{n}}$$
$$= \inf_{t \geq 0} \rho_t(\alpha) = \lim_{t \to \infty} \rho_t(\alpha).$$

In particular, for every $\alpha \in \mathcal{I}$ there exists $t \geq 0$ such that $\rho_t(\alpha) < \infty$. Hence, $\bigcup_{t \geq 0} \mathcal{I}_t = \mathcal{I}$. Since $\rho(\alpha) \to \infty$ for $\alpha \to \partial \mathcal{I}$, there exists an $\epsilon > 0$ such that $\alpha^*, \alpha_t^* \in [\gamma + \epsilon, 1 - \gamma - \epsilon] =: \hat{\mathcal{I}}$ for all $t \geq 0$.

Convergence $\gamma_t \to \gamma$ implies the existence of a $t_0 > 0$ such that $\hat{\mathcal{I}} \subseteq \mathcal{I}_t$ for all $t \geq t_0$. In particular, we can consider the family $\rho, (\rho_t)_{t \geq t_0}$ of uniformly continuous functions on $\hat{\mathcal{I}}$.

In the next step we argue that $\alpha_t^* \overset{t \to \infty}{\longrightarrow} \alpha^*$. Notice that by assumption $\rho(\alpha) = \lim_{t \to \infty} \rho_t(\alpha) \geq \lim_{t \to \infty} \rho_t(\alpha_t^*) \geq 1$ and

$$\rho(\alpha^*) \leq \rho(\alpha_t^*) \leq \rho_t(\alpha_t^*) \to 1.$$

Hence, $\rho(\alpha^*) = 1$. Suppose that $\alpha_t^* \overset{t \to \infty}{\longrightarrow} \alpha^*$ does not hold. Then there exists a $\delta > 0$ and a subsequence $t_n \uparrow \infty$ such that $|\alpha_{t_n}^* - \alpha^*| \geq \delta$ for all $n \in \mathbb{N}$. Since ρ is strictly convex with unique minimizer α^*, we have

$$\delta' := \min\{\rho(\alpha^* - \delta), \rho(\alpha^* + \delta)\} - 1 > 0$$

and $\rho(\alpha) \geq 1 + \delta'$ for all $\alpha \notin [\alpha^* - \delta, \alpha^* + \delta]$. In particular,

$$\rho_{t_n}(\alpha^*) \geq \rho_{t_n}(\alpha_{t_n}^*) \geq \rho(\alpha_{t_n}^*) \geq 1 + \delta'.$$

Since the term on the left-hand side converges to 1, this is a contradiction and the convergence $\alpha_t^* \overset{t \to \infty}{\longrightarrow} \alpha^*$ is established.

The fact that α_t^* converges to α^* and f_t converges to f implies that A_{t,α_t^*} converges to A_{α^*} because of the uniform continuity of the operator in $\alpha \in \hat{\mathcal{I}}$. Since the eigenspaces of $\rho(\alpha^*)$ and $\rho_t(\alpha_t^*)$ are one dimensional, one can argue along the lines of Note 3 on Chapter II in [11, pages 568-569] to see that v_{t,α_t^*} converges to v_{α^*}. Since the functions are bounded,

$$\frac{v_{t,\alpha_t^*}(0)}{v_{t,\alpha_t^*}(\ell)}$$

is uniformly bounded in t from zero and infinity. With the observed convergences and this uniform bound (2.3) now implies that also

$$\rho_t''(\alpha_t^*) \to \rho''(\alpha^*),$$

as required. □

2.2   The many-to-one Lemma

We first continue the analysis of the IBRW. The following lemma is based on a spine construction which is known as Lyons' change of measure [13]. Recall the Ulam-Harris notation from Section 2.1.

9

Lemma 2.4 (Many-to-one). *For all $\alpha \in \mathcal{I}$ there exists a probability measure $\mathbb{P}^\alpha$ on some measurable space, and a Markov process $((S_n, \tau_n): n \in \mathbb{N}_0; \mathbb{P}^\alpha)$ with state space $\mathbb{R} \times \mathcal{T}$, such that for all $n \in \mathbb{N}$, $(s_0, \tau_0) \in \mathbb{R} \times \mathcal{T}$ and $F: (\mathbb{R} \times \mathcal{T})^n \to \mathbb{R}_+$ measurable*

$$
E^p_{(s_0,\tau_0)} \Big[\sum_{|x|=n} e^{-\alpha(S(x_n) - S(x_0))} \rho^p(\alpha)^{-n} \frac{v_\alpha(\tau(x_n))}{v_\alpha(\tau(x_0))} F(S(x_1), \tau(x_1), \ldots, S(x_n), \tau(x_n)) \Big]
$$
$$
= \mathbb{E}^\alpha_{(s_0,\tau_0)}[F(S_1, \tau_1, \ldots, S_n, \tau_n)]. \tag{2.5}
$$

Moreover, $\sigma((S_i, \tau_i): i \in [n]) = \sigma(S_i: i \in [n])$.

Note that it is easy to check that the distribution of $((S_n, \tau_n) \; : \; n \in \mathbb{N}_0; \mathbb{P}^\alpha)$ does not depend on the percolation parameter p.

Proof. Given a labelled tree (Γ, L), with $L(x) = (S(x), \tau(x))$, we can distinguish an ancestral line $\xi = (\xi_1, \xi_2, \ldots)$ which we call spine. In the space of labelled trees, we denote by $\mathcal{F}_n$ the σ-field generated by the first n generations, $\mathcal{F}_n = \sigma((x, L(x)): |x| \leq n)$. The analogue in the space of trees with spines is denoted by $\mathcal{F}_n^*$.

For every $(s_0, \tau_0) \in \mathbb{R} \times \mathcal{T}$, the distribution of the IBRW started in (s_0, τ_0) can be interpreted as a distribution $P^p_{(s_0,\tau_0)}$ on the set of labelled trees. We extend this measure to the space of labelled trees with spines. Since (s_0, τ_0) and p will remain fixed throughout the proof, we omit it from the notation and write $P = P^p_{(s_0,\tau_0)}$ for brevity.

Note that every $\mathcal{F}_n^*$-measurable function g can be written as

$$
g(\Gamma, L, \xi) = \sum_{|x|=n} g_x(\Gamma, L) \mathbb{1}_{\xi_n = x},
$$

for $\mathcal{F}_n$-measurable functions g_x (see page 24 in [17]). We define P_n^* to be the (non-probability) measure on $\mathcal{F}_n^*$ such that for all nonnegative $\mathcal{F}_n^*$-measurable functions g,

$$
\int g(\Gamma, L, \xi) \, dP_n^* = \int \sum_{|x|=n} g_x(\Gamma, L) \, dP|_{\mathcal{F}_n}.
$$

We now construct a new branching random walk under a new probability measure $\mathbb{P}^\alpha$. The root has again label $L(\emptyset) = (S(\emptyset), \tau(\emptyset)) = (s_0, \tau_0)$. A particle ξ_n on the spine in generation n with label $(S(\xi_n), \tau(\xi_n))$ produces new offspring with distribution

$$
\frac{d\hat{\mathcal{L}}_{\tau(\xi_n)}}{d\mathcal{L}_{\tau(\xi_n)}}(\mu) = \frac{1}{\rho^p(\alpha)} \int e^{-\alpha(S(x) - S(\xi_n))} \frac{v_\alpha(\tau(x))}{v_\alpha(\tau(\xi_n))} \mu(dx)
$$

for all atomic measures μ on $\mathbb{R} \times \mathcal{T}$. Here $\mathcal{L}_\sigma$ denotes the offspring distribution for a particle of type $\sigma \in \mathcal{T}$ in the original process. The new spine particle ξ_{n+1} in generation $n+1$ is chosen from the offspring of ξ_n by choosing an offspring x with probability proportional to

$$
e^{-\alpha(S(x) - S(\xi_n))} v_\alpha(\tau(x)).
$$

Off the spine the new branching random walk behaves exactly as the original one. Then

$$\frac{d\mathbb{P}^\alpha|_{\mathcal{F}_n^*}}{dP_n^*}(\Gamma, L, \xi) = \frac{d\mathbb{P}^\alpha|_{\mathcal{F}_{n-1}^*}}{dP_{n-1}^*}(\Gamma, L, \xi)\, \rho^p(\alpha)^{-1} \sum_{x\colon \mathrm{p}(x)=\xi_{n-1}} e^{-\alpha(S(x)-S(\xi_{n-1}))} \frac{v_\alpha(\tau(x))}{v_\alpha(\tau(\xi_{n-1}))}$$

$$\frac{e^{-\alpha(S(\xi_n)-S(\xi_{n-1}))} v_\alpha(\tau(\xi_n))}{\sum_{x\colon \mathrm{p}(x)=\xi_{n-1}} e^{-\alpha(S(x)-S(\xi_{n-1}))} v_\alpha(\tau(x))}$$

$$= \frac{d\mathbb{P}^\alpha|_{\mathcal{F}_{n-1}^*}}{dP_{n-1}^*}(\Gamma, L, \xi)\rho^p(\alpha)^{-1} e^{-\alpha(S(\xi_n)-S(\xi_{n-1}))} \frac{v_\alpha(\tau(\xi_n))}{v_\alpha(\tau(\xi_{n-1}))}$$

$$= \prod_{k=1}^n \left(\rho^p(\alpha)^{-1} e^{-\alpha(S(\xi_k)-S(\xi_{k-1}))} \frac{v_\alpha(\tau(\xi_k))}{v_\alpha(\tau(\xi_{k-1}))}\right)$$

$$= \rho^p(\alpha)^{-n} e^{-\alpha(S(\xi_n)-S(\xi_0))} \frac{v_\alpha(\tau(\xi_n))}{v_\alpha(\tau(\xi_0))}.$$

In particular, for all $F\colon (\mathbb{R}\times\mathcal{T})^n \to [0,\infty)$ measurable

$$\mathbb{E}^\alpha\big[F(S(\xi_1),\tau(\xi_1),\ldots,S(\xi_n),\tau(\xi_n))\big]$$

$$= E_n^*\Big[F(S(\xi_1),\tau(\xi_1),\ldots,S(\xi_n),\tau(\xi_n))\rho(\alpha)^{-n} e^{-\alpha(S(\xi_n)-S(\xi_0))} \frac{v_\alpha(\tau(\xi_n))}{v_\alpha(\tau(\xi_0))}\Big]$$

$$= E\Big[\sum_{|x|=n} F(S(x_1),\tau(x_1),\ldots,S(x_n),\tau(x_n))\rho(\alpha)^{-n} e^{-\alpha(S(x_n)-S(x_0))} \frac{v_\alpha(\tau(x_n))}{v_\alpha(\tau(x_0))}\Big].$$

We define $(S_n,\tau_n) := (S(\xi_n),\tau(\xi_n))$ and $\mathbb{P}^\alpha_{(s_0,\tau_0)} = \mathbb{P}^\alpha$. The Markov property follows from the definition of the process. Since the offspring distribution of the spine is absolutely continuous with respect to the offspring distribution of the original process and the type of the original process is a function of the locations of its ancestors and the particle itself, the proof is complete. $\square$

Next we also need a higher-dimensional version of the many-to-one lemma that includes the number of offspring of the particles on the spine. For any $x=(x_0,\ldots,x_n)$ in the branching process, define the point measure on $\mathbb{R}$

$$\nu_x = \sum_{y\colon |y|=n+1, y_n=x} \delta_{S(y)-S(x)},$$

with δ_s a Dirac mass in s, which describes the positions of offspring of x relative to the position of x. Denote by $M_p(\mathbb{R})$ the point measures on $\mathbb{R}$.

Lemma 2.5. *For all $\alpha \in \mathcal{I}$, $(s_0,\tau_0) \in \mathbb{R}\times\mathcal{T}$ there exists a probability measure $\mathbb{P}^\alpha_{(s_0,\tau_0)} = \mathbb{P}^{\alpha,p}_{(s_0,\tau_0)}$ and a Markov process $((S_i,\tau_i,\nu_{i-1})_{i=0,\ldots,n},\colon n \in \mathbb{N}_0; \mathbb{P}^\alpha_{(s_0,\tau_0)})$ started in $(s_0,\tau_0,\delta_{s_0})$ with state space $\mathbb{R}\times\mathcal{T}\times M_p(\mathbb{R})$ such that for all measurable F*

$$E^p_{(s_0,\tau_0)}\Big[\sum_{|x|=n} e^{-\alpha(S(x_n)-S(x_0))}\rho(\alpha)^{-n}\frac{v_\alpha(\tau(x_n))}{v_\alpha(\tau(x_0))}F(S(x_1),\tau(x_1),\nu_{x_0},\ldots,S(x_n),\tau(x_n),\nu_{x_{n-1}})\Big]$$

$$= \mathbb{E}^\alpha_{(s_0,\tau_0)}[F(S_1,\tau_1,\nu_0,\ldots,S_n,\tau_n,\nu_{n-1})].$$

Moreover, for any measurable $A \subset \mathbb{R}$ and any measurable F,

$$\mathbb{E}[F(S_n-S_{n-1},\nu_{n-1}(A))\,|\,\sigma(S_i,\tau_i,\nu_{i-1}, i\le n-1)]$$

$$= \mathbb{E}_{(0,\tau)}\Big[\sum_{|x|=1} e^{-\alpha S(x)}\rho(\alpha)^{-1}\frac{v_\alpha(\tau(x))}{v_\alpha(\tau)}F(S(x),\{y\,|\,|y|=1, S(y)\in A\})\Big]\Big|_{\tau=\tau_{n-1}}.$$

Note that unlike in Lemma 2.4 the distribution of $(S_k,\tau_k,\nu_{k-1})_{k\in\mathbb{N})}$ does depend on p, since we are considering a non-linear function of the point process describing the position of the offspring.

Proof. Consider the IBRW with spine ξ under the measure $\mathbb{P}^\alpha$ as constructed in the proof of Lemma 2.4. Then, define $(S_n, \tau_n, \nu_{n-1}) := (S(\xi_n), \tau(\xi_n), \nu_{\xi_{n-1}})$ and the first statement follows since we know the explicit Radon-Nikodym density of $\mathbb{P}^\alpha$ with respect to P_n^*. The second statement is a consequence of the Markov property of the IBRW with spine combined with a suitable choice of test function F. $\square$

Lemma 2.6 (Moments). *Let $\alpha \in \mathcal{I}$ and $((S_n, \tau_n) : n \in \mathbb{N}_0)$ be the Markov process from Lemma 2.4. For all $\tau_0 \in \mathcal{T}$*

$$\mathbb{E}^\alpha_{(0,\tau_0)}[S_1] = -\frac{\rho'(\alpha)}{\rho(\alpha)} \qquad \mathbb{E}^\alpha_{(0,\tau_0)}[S_1^2] = \frac{\rho''(\alpha)}{\rho(\alpha)}.$$

Proof. Follows immediately from (2.5) and (2.3). $\square$

2.3 Asymptotic moment estimates

In the proofs of Theorems 1.1 and 1.2 we will need estimates for moments of the Markov chains defined in Lemmas 2.4 and 2.5. Suppose that f_{t_n} is a sequence of decreasing attachment rules such that $f_{t_n} \downarrow f$ pointwise for an attachment rule f. Further, suppose that $\gamma_n := \lim_{k\to\infty} f_{t_n}(k)/k < \frac{1}{2}$ for all $n \in \mathbb{N}$, so that by Proposition 2.3 also $\gamma = \lim_{k\to\infty} \frac{f(k)}{k} < \frac{1}{2}$. Let α_n^*, resp. α^* be the unique minimizer of the spectral radius ρ_n, resp. ρ, of the operator corresponding to the attachment rule f_n, resp. f. In the setting of Theorem 1.1 take $f_{t_n} = f$ and write $\rho_n = \rho^{p_n} = p_n \rho(\cdot)$.

The Markov chain from Lemma 2.4 corresponding to attachment rule f_{t_n} is denoted by $((S_i^{(n)}, \tau_i^{(n)}) : i \in \mathbb{N}; \mathbb{P} = \mathbb{P}^{\alpha_n^*}_{(s_0, \ell)})$,

Lemma 2.7. *There exists $\eta > 0$ such that*

$$\sup_{n \in \mathbb{N}, \tau \in \mathcal{T}} \mathbb{E}^{\alpha_n^*}_{0,\tau} \left[e^{\eta |S_1^{(n)}|} \right] < \infty.$$

Proof. We first consider the case $S_1^{(n)} \geq 0$. Note that

$$\mathbb{E}^{\alpha_n} \left[e^{\eta S_1^{(n)}} \right] \leq 1 + E_{(0,\tau)} \left[\sum_{|x|=1} e^{(\eta - \alpha_n^\star) S^{(n)}(x)} \frac{v_{\alpha_n}(\tau(x))}{v_{\alpha_n}(\tau)} \rho_n(\alpha_n)^{-1} \mathbb{1}_{\{S^{(n)}(x) \geq 0\}} \right]$$

$$\leq 1 + \sup_{k \in \mathbb{N}} \frac{v_{\alpha_k}(0)}{v_{\alpha_k}(\ell)} E_{(0,\tau)} \left[\sum_{|x|=1} e^{(\eta - \alpha_n^\star) S^{(n)}(x)} \mathbb{1}_{\{S^{(n)}(x) \geq 0\}} \right]$$

where we used that $\rho_n(\alpha_n) > 1$, the monotonicity in types for v_{α_n} by Lemma 2.2 and the uniform boundedness of the quotient $\frac{v_{\alpha_n}(0)}{v_{\alpha_n}(\ell)}$ in n by Proposition 2.3. Hence, it suffices to show that for the right choice of η the expectation on the right hand side remains bounded in n.

Now, by Lemma 2.2, we have that $\eta := \frac{1}{4}(\alpha^* - \gamma) > 0$. By Proposition 2.3, we have that $\gamma_n \downarrow \gamma$ and $\alpha_n^* \to \alpha^*$. Then, we can choose n_0 sufficiently large such that for all $n \geq n_0$ we have $\gamma_n < \gamma + \eta$ and $\alpha_n^\star \geq \alpha^* - \eta$.

Furthermore, we denote by $\tilde{Z}^{(n_0)}$ the jump process that jumps from k to $k+1$ with rate $f_{t_{n_0}}(k)$ started in 1. Then by comparison with a Yule process with constant branching rate, we can find a constant $C(n_0) > 0$ such that

$$\sup_{t \geq 0} E[\tilde{Z}_t^{(n_0)}] e^{-(\gamma + \eta)t} \leq C(n_0). \tag{2.6}$$

Finally, we obtain from the construction of IBRW and using (2.6) that for $n \geq n_0$

$$E_{(0,\tau)} \left[\sum_{|x|=1} e^{(\eta - \alpha_n^\star) S^{(n)}(x)} \mathbb{1}_{\{S^{(n)}(x) \geq 0\}} \right] \leq \sum_{k \in \mathbb{N}_0} e^{-(\alpha^* - 2\eta)k} E_{(0,\tau)} \left[\sum_{|x|=1} \mathbb{1}_{\{S(x) \in [k, k+1]\}} \right]$$

$$\leq \sum_{k \in \mathbb{N}_0} e^{-(\alpha^* - 2\eta)k} E[\tilde{Z}_{k+1}^{(n_0)}] \leq C(n_0) \sum_{k \in \mathbb{N}_0} e^{-(\alpha^* - 3\eta - \gamma)k + \gamma + \eta},$$

12

which is finite by choice of η.

The fact that the supremum over all n is finite follows from the same argument if we redefine η as $\frac{1}{4}\min\{\alpha^* - \gamma, \alpha_k^* - \gamma_k, k \leq n_0\}$.

For the case that $S_1^{(n)} \leq 0$, it suffices to prove in a second step, that there exists $\eta > 0$ such that

$$E_{(0,\tau)}\Big[\sum_{|x|=1} e^{-(\alpha_n^*+\eta)S^{(n)}(x)} \mathbb{1}_{\{S^{(n)}(x)\leq 0\}} \Big] < \infty.$$

Since the children to the left of a particle in the IBRW form a Poisson process and their distribution is not depending on the type of their ancestor we have by construction

$$E_{(0,\tau)}\Big[\sum_{|x|=1} e^{-(\alpha_n^*+\eta)S^{(n)}(x)} \mathbb{1}_{\{S^{(n)}(x)\leq 0\}} \Big] = \int_0^\infty e^{(\alpha_n^*+\eta)t} e^{-t} E[f_{t_n}(X_t^{(f_{t_n})})]\, dt,$$

where $X^{(f)}$ is the pure birth process with jump rates given by f.

Let $\epsilon_n = \frac{1}{8}(1 - \gamma_n - \alpha_n^*)$ and let $\epsilon = \frac{1}{4}(1 - \gamma - \alpha^*)$, where $\gamma = \inf_{k\geq 1} \frac{f(k)}{k}$. By Lemma 2.2, $\epsilon_n, \epsilon > 0$.

Using that $f_{t_k}(k) \leq k+1$ and a comparison to a Yule process, we have that there exists $C_n > 0$ such that

$$E[f_{t_n}(X_t^{(f_{t_n})})] \leq E[X_t^{(f_{t_n})}] + 1 \leq C_n\, e^{(\gamma_n+\epsilon_n)t} \quad \text{for all } t \geq 0.$$

By Proposition 2.3, we can find n_0 such that for all $n \geq n_0$,

$$|\alpha_n - \alpha^*| < \epsilon, \quad \text{and} \quad \gamma \leq \gamma_n \leq \gamma + \epsilon.$$

Define $\eta := \min\{\frac{3}{8}\epsilon, \epsilon_n, n \leq n_0\}$. Then, we have for $n \leq n_0$,

$$\int_0^\infty e^{(\alpha_n^*+\eta)t} e^{-t} E[f_{t_n}(X^{(f_{t_n})})]\, dt \leq C_n \int_0^\infty e^{(\alpha_n^*+\eta-1+\gamma_n+\epsilon_n)t}\, dt \leq C_n \int_0^\infty e^{-6\eta t}\, dt < \infty.$$

Furthermore, for $n \geq n_0$, we can use the monotonicity of f_{t_n} to deduce that

$$\begin{aligned}
\int_0^\infty e^{(\alpha_n^*+\eta)t} e^{-t} E[f_{t_n}(X^{(f_{t_n})})]\, dt &\leq \int_0^\infty e^{(\alpha_n^*+\eta)t} e^{-t} E[f_{t_{n_0}}(X^{(f_{t_{n_0}})})]\, dt \\
&\leq C_{n_0} \int_0^\infty e^{(\alpha_n^*+\eta-1+\gamma_{n_0}+\epsilon_{n_0})t}\, dt \\
&\leq C_{n_0} \int_0^\infty e^{(\alpha^*+\eta-1+\gamma+2\epsilon+\frac{1}{8}(1-\gamma-\alpha^*+\epsilon))t}\, dt \\
&\leq C_{n_0} \int_0^\infty e^{-\epsilon t}\, dt,
\end{aligned}$$

which completes the second step and thus the proof of the lemma. $\qquad\square$

Lemma 2.8. *Let $(S_k^{(n)}, \tau_k^{(n)}, \nu_{k-1}^{(n)})$ be the Markov chain defined in Lemma 2.5 either for attachment rule f_{t_n} and percolation parameter 1 or for fixed attachment rule f (with $\gamma < \frac{1}{2}$) and percolation parameter p_n, where $p_n \downarrow \rho(\alpha^*)^{-1}$. For any sequence $(M_n)_{n\in\mathbb{N}}$ such that $M_n \to \infty$, there exist constants $C > 0$ and $\hat{\gamma} > 0$ such that, for all n,*

$$\sup_{\tau\in\mathcal{T}} \mathbb{E}_{(0,\tau)}^{\alpha_n^*}[\nu_0^{(n)}((-\infty, M_n))\mathbb{1}_{\{S_1^{(n)}\leq M_n\}}] \leq Ce^{\hat{\gamma}M_n}.$$

Proof. To unify notation define $p_n = 1$ in the varying f case and $f_{t_n} = f$ in the percolation case. By the extended many-to-one formula, Lemma 2.5, we have that

$$\begin{aligned}
&\mathbb{E}_{(0,\tau)}^{\alpha_n^*}\big[\nu_0^{(n)}((-\infty, M_n))\mathbb{1}_{\{S_1^{(n)}\leq M_n\}}\big] \\
&\quad = E_{(0,\tau)}^{p_n}\bigg[\bigg(\sum_{|x|=1} e^{-\alpha_n^* S^{(n)}(x)} \rho_n(\alpha_n^*)^{-1} \frac{v_{\alpha_n^*}(\tau(x))}{v_{\alpha_n^*}(\tau)}(\mathbb{1}_{\{S^{(n)}(x)\leq M_n\}})\bigg)\bigg(\sum_{|x|=1}\mathbb{1}_{\{S^{(n)}(x)\leq M_n\}}\bigg)\bigg].
\end{aligned}$$

We can use that $\rho_n(\alpha_n^*) > 1$, the monotonicity in types and in p_n, Lemma 2.2, and that $C := \sup_{n \in \mathbb{N}} \frac{v_{\alpha_n}(0)}{v_{\alpha_n}(\ell)} < \infty$ by Proposition 2.3, in order to bound the above by

$$
E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} e^{-\alpha_n^* S^{(n)}(x)} \frac{v_{\alpha_n}(0)}{v_{\alpha_n}(\ell)}\left(\mathbb{1}_{\{S^{(n)}(x)\leq M_n\}}\right)\right)\left(\sum_{|x|=1} \mathbb{1}_{\{S^{(n)}(x)\leq M_n\}}\right)\right]
$$

$$
\leq C\, E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} \mathbb{1}_{\{0\leq S^{(n)}(x)\leq M_n\}}\right)^2\right]
$$

$$
+ 2C\, E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} e^{-\alpha_n^* S^{(n)}(x)}\mathbb{1}_{\{S^{(n)}(x)\leq 0\}}\right)\right] E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} \mathbb{1}_{\{0\leq S^{(n)}(x)\leq M_n\}}\right)\right]
$$

$$
+ C\, E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} e^{-\alpha_n^* S^{(n)}(x)}\mathbb{1}_{\{S^{(n)}(x)\leq 0\}}\right)\left(\sum_{|x|=1} \mathbb{1}_{\{S^{(n)}(x)\leq 0\}}\right)\right].
$$

$$(2.7)$$

For the first term in (2.7), we note that $f(k) \leq f_{t_n}(k) \leq f_{t_1}(k) \leq \gamma_1^+ k + f_{t_1}(0)$ (for $\gamma_1^+ = \sup\{\Delta f_{t_1}(k), k \in \mathbb{N}_0\}$. Therefore, if we let $(\hat{Z}^{f_{t_1}})_{t\geq 0}$ be the jump process jumping from k to $k+1$ at rate $f_{t_1}(k)$ started in 1, we can conclude that by construction of the IBRW

$$
E_{(0,\tau)}^1\left[\left(\sum_{|x|=1} \mathbb{1}_{\{0\leq S^{(n)}(x)\leq M_n\}}\right)^2\right] \leq E\left[(\hat{Z}_{M_n}^{f_{t_1}})^2\right] \leq \tilde{C}(f_{t_1})e^{\hat{\gamma} M_n},
$$

for some constants $C(f_{t_1})$ and $\hat{\gamma} > 0$, where the latter bound follows by comparison with a Yule process, whose second moments grow at most exponentially.

For the second term on the right hand side in (2.7), the first expectation is bounded uniformly in n by (the second part of) Lemma 2.7 and the second expectation can be bounded by the second moment, so that the first part of the argument applies.

For the final term in (2.7), we use that the particles to the left form a Poisson process, so that we can use a standard identity for Poisson processes, see e.g. [12, Equation (4.26)], to deduce that

$$
E_{(0,\tau)}\left[\left(\sum_{|x|=1} e^{-\alpha_n^* S^{(n)}(x)}\mathbb{1}_{\{S^{(n)}(x)\leq 0\}}\right)\left(\sum_{|x|=1} \mathbb{1}_{\{S^{(n)}(x)\leq 0\}}\right)\right]
$$

$$
= \int_0^\infty e^{\alpha_n^* t} e^{-t} E[f_{t_n}(X^{(f_{t_n})})]\, dt\left(1 + \int_0^\infty e^{-t} E[f_{t_n}(X^{(f_{t_n})})]\, dt\right).
$$

However, as in the proof of Lemma 2.7 the right hand side is bounded uniformly in n. $\qquad\square$

Corollary 2.9. *In the setting of Lemma 2.8, we have for any sequence $N_n \to \infty$ and with $R_n = e^{N^{1/4}}, M_n = N^{1/5}$ that there exist $\tilde{C}, \tilde{\gamma} > 0$ such that*

$$
\inf_{\tau \in \mathcal{T}} \mathbb{P}_{(0,\tau)}^{\alpha_n^*}\left(\bar{\nu}_0^{(n)} \leq R_n, S_1^{(n)} \leq M_n\right) \geq 1 - \tilde{C}e^{-\tilde{\gamma} N^{1/5}}.
$$

Proof. By Lemma 2.7, there exists $\eta > 0$ such that $C := \sup_{n \in \mathbb{N}_0, \tau} \mathbb{E}_{(0,\tau)}^{\alpha_n^*}[e^{\eta S_1^{(n)}}] < \infty$. Moreover, by Lemma 2.8, there exist $C' > 0, \gamma' > 0$ such that $\sup_\tau \mathbb{E}_{(0,\tau)}^{\alpha}[\bar{\nu}_0, S_1 \leq M_n] \leq C'e^{\gamma' M_n}$. Hence, we can estimate by Chebyshev's inequality

$$
\sup_{\tau \in \mathcal{T}} \mathbb{P}_{(0,\tau)}^{\alpha_n^*}\left((\bar{\nu}_0 \leq R_n, S_1 \leq M_n)^c\right)
$$

$$
\leq P_{(0,\tau_{k-1})}^{\alpha}(S_1 \geq M_n) + P_{(0,\tau_{k-1})}^{\alpha}(\bar{\nu}_0 \geq R_n, S_1 \leq M_n)
$$

$$
\leq e^{-\eta M_n} + \frac{1}{R_n}C'e^{\gamma' M_n} = e^{-\eta N^{1/5}} + C'e^{\gamma' N^{1/5} - N^{1/4}},
$$

by our choice of M_n and R_n. Therefore, the statement of the corollary follows by choosing $\tilde{\gamma}$ and $\tilde{C}$ appropriately. $\qquad\square$

2.4 Mogulskii's theorem

The main technical tool in the proof of our main result is the following large deviation result due to Mogulskii in its original form. We state it here in a version adapted to Markov chains as a generalisation of the version for random walks found in [8].

Theorem 2.10 ([14],[8]). *Let $\mathcal{T}$ be a nonempty set. We assume the following:*

(i) *For each $n \in \mathbb{N}$, $((S_i^{(n)}, \tau_i^{(n)}): i \in \mathbb{N})$ is a Markov chain with values in $\mathbb{R} \times \mathcal{T}$.*

(ii) *$(a_n)_{n \in \mathbb{N}}$ and $((k_n)_{n \in \mathbb{N}})$ are positive sequences with $a_n, k_n \to \infty$ and $a_n^2/k_n \to 0$ as $n \to \infty$.*

(iii) *For all $c_1, c_2, c_3, c_4 \in \mathbb{R}$, $A > 0$ and for $r_n := \lfloor Aa_n^2 \rfloor$,*

$$\mathbb{P}_{(0,\tau)}\Big(c_1 \leq \frac{S_i^{(n)}}{a_n} \leq c_2 \,\forall\, i \in [r_n]; c_3 \leq \frac{S_{r_n}^{(n)}}{a_n} \leq c_4\Big) \tag{2.8}$$
$$\xrightarrow{n \to \infty} P(c_1 \leq \sqrt{\sigma^2 A} W_t \leq c_2 \,\forall\, t \in [0,1]; c_3 \leq \sqrt{\sigma^2 A} W_1 \leq c_4)$$

uniformly in $\tau \in \mathcal{T}$, where $\sigma^2 > 0$ is a constant independent of A and $(W_t: t \geq 0; P)$ is a standard Brownian motion.

Let $g_1 < g_2$ be two continuous functions on $[0,1]$ with $g_1(0) \leq 0 \leq g_2(0)$ and denote

$$E_n := \Big\{g_1\Big(\frac{i}{k_n}\Big) \leq \frac{S_i^{(n)}}{a_n} \leq g_2\Big(\frac{i}{k_n}\Big) \,\forall\, i \in [k_n]\Big\}.$$

Then, for all $\tau_0 \in \mathcal{T}$,

$$\lim_{n \to \infty} \frac{a_n^2}{k_n} \log P_{(0,\tau_0)}(E_n) = -\frac{\pi^2 \sigma^2}{2} \int_0^1 \frac{1}{[g_2(t) - g_1(t)]^2} \, dt.$$

Moreover, for any $b > 0$, $\tau_0 \in \mathcal{T}$,

$$\lim_{n \to \infty} \frac{a_n^2}{k_n} \log P_{(0,\tau_0)}(E_n, S_{k_n}^{(n)} \geq (g_2(1) - b)a_n) = -\frac{\pi^2 \sigma^2}{2} \int_0^1 \frac{1}{[g_2(t) - g_1(t)]^2} \, dt.$$

The proof of Theorem 2.10 is postponed to Appendix A.

We will now show that we can apply Mogulskii's results to the Markov chains from Lemmas 2.4 and 2.5. We will treat both the setting of Theorem 1.1 and 1.2 at the same time and so continue using the notation introduced at the beginning of Section 2.3.

To this end, we first recall Donsker's theorem for martingale difference arrays (see for example Theorem 7.7.3 in [7] or Theorem 18.2 in [1]). The theorem is usually stated for $r_n = n$, but it is straightforward to generalize the statement to the following:

Proposition 2.11. *Let $(r_n)_{n \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ be a sequence with $r_n \uparrow \infty$ as $n \to \infty$. For every $n \in \mathbb{N}$, let $(\xi_i^n: 1 \leq i \leq r_n)$ be a family of random variables and denote $\mathcal{F}_i^n = \sigma(\xi_1^n, \ldots, \xi_i^n)$ for all $i \leq r_n$. Assume that*

(i) *$E[\xi_i^n | \mathcal{F}_{i-1}^n] = 0$ for all $i \leq r_n$, $n \in \mathbb{N}$.*

(ii) *For all $t \in [0,1]$, $\sum_{i \leq \lfloor tr_n \rfloor} E[(\xi_i^n)^2 | \mathcal{F}_{i-1}^n] \to t$ in probability as $n \to \infty$.*

(iii) *For all $\epsilon > 0$, $\sum_{i \leq r_n} E[(\xi_i^n)^2 \mathbb{1}_{|\xi_i^n| > \epsilon} | \mathcal{F}_{i-1}^n] \to 0$ in probability as $n \to \infty$.*

Then the linear interpolation of $(\sum_{i \leq m} \xi_i^n : m \leq r_n)$ converges weakly to a standard Brownian motion on $[0,1]$.

We use Donsker's theorem as follows.

Lemma 2.12. *Let $A > 0$, $(a_n)_{n \in \mathbb{N}}$ be a positive sequence with $\lim_{n \to \infty} a_n = \infty$ and write $r_n = \lfloor A a_n^2 \rfloor$ for each $n \in \mathbb{N}$. Moreover, let $(S_i^{(n)}, \tau_i^{(n)})_{i \in \mathbb{N}_0}$ be the Markov chain introduced in Lemma 2.4 for attachment rule f_{t_n}. For all $c_1, c_2, c_3, c_4 \in \mathbb{R}$,*

$$\mathbb{P}_{(0,\tau)}\left(c_1 \leq \frac{S_i^{(n)}}{a_n} \leq c_2 \, \forall i \in [r_n]; c_3 \leq \frac{S_{r_n}^{(n)}}{a_n} \leq c_4\right) \to P(c_1 \leq \sqrt{\sigma^2 A} W_t \leq c_2 \, \forall t \in [0,1]; c_3 \leq \sqrt{\sigma^2 A} W_1 \leq c_4)$$

as $n \to \infty$, uniformly in $\tau \in \mathcal{T}$, where $(W_t : t \geq 0; P)$ is a standard Brownian motion.

Proof. The first step is to show that the conditions of Proposition 2.11 are satisfied by the random variables

$$\xi_i^n := \frac{S_i^{(n)} - S_{i-1}^{(n)}}{\sqrt{r_n \sigma^2}}, \qquad i = 1, \ldots, r_n. \tag{2.9}$$

Let $\mathcal{F}_i^n = \sigma(S_1^{(n)}, \ldots, S_i^{(n)})$. For all $i \in [r_n]$,

$$\mathbb{E}[\xi_i^n | \mathcal{F}_{i-1}^n] = \frac{1}{\sqrt{r_n \sigma^2}} \mathbb{E}_{(S_{i-1}^{(n)}, \tau_{i-1}^{(n)})}^{\alpha_n^*}[S_1^{(n)} - S_0^{(n)}] = \frac{1}{\sqrt{r_n \sigma^2}} \mathbb{E}_{(0, \tau_{i-1}^{(n)})}^{\alpha_n^*}[S_1^{(n)}] = 0$$

by Lemma 2.6 and $\rho_n'(\alpha_n^*) = 0$ as α_n^* minimizes ρ_n. Moreover, as $n \to \infty$,

$$\sum_{i \leq \lfloor tr_n \rfloor} \mathbb{E}[(\xi_i^n)^2 | \mathcal{F}_{i-1}^n] = \sum_{i \leq \lfloor tr_n \rfloor} \frac{1}{r_n \sigma^2} \mathbb{E}_{(S_{i-1}^{(n)}, \tau_{i-1}^{(n)})}^{\alpha_n^*}[(S_1^{(n)} - S_0^{(n)})^2] = \frac{\lfloor tr_n \rfloor}{r_n \sigma^2} \frac{\rho_n''(\alpha_n^*)}{\rho_n(\alpha_n^*)} \to \frac{t}{\sigma^2} \sigma^2 = t,$$

since $r_n \to \infty$ and $\rho_n(\alpha_n^*)/\rho_n''(\alpha_n^*) \to \sigma^2$ by definition. Thus, Condition (ii) of Proposition 2.11 is satisfied. Finally, let η be as in Lemma 2.7, then

$$\sum_{i \leq r_n} \mathbb{E}\left[(\xi_i^n)^2 \mathbb{1}_{|\xi_i^n| > \epsilon} | \mathcal{F}_{i-1}^n\right] = \frac{1}{r_n \sigma^2} \sum_{i \leq r_n} \mathbb{E}_{(0, \tau_{i-1}^{(n)})}^{\alpha_n^*}\left[(S_1^{(n)})^2 \mathbb{1}\{|S_1^{(n)}| > \epsilon \sqrt{r_n \sigma^2}\}\right]$$

$$\leq \sigma^{-2} \sup_{\tau \in \mathcal{T}} \mathbb{E}_{(0, \tau)}^{\alpha_n^*}\left[(S_1^{(n)})^2 \mathbb{1}\{|S_1^{(n)}| > \epsilon \sqrt{r_n \sigma^2}\}\right]$$

$$\leq C \sigma^{-2} \frac{1}{\epsilon \sqrt{r_n \sigma^2}} \sup_{\tau \in \mathcal{T}} \mathbb{E}_{(0, \tau)}^{\alpha_n^*}\left[e^{\eta |(S_1^{(n)})|}\right],$$

where C is a constant such that $x^2 \leq C e^{\eta |x|}$ for all $x \in \mathbb{R}$. Then, by Lemma 2.7 the exponential moment is bounded uniformly in n, so that the right hand side converges to zero as required.

To apply Theorem 2.10 we have to check that the convergence of the distribution functions is uniform in the start type. This is guaranteed by the monotonicity of the IBRW in the start type (which was proven in [6, Remark 2.6]) which entails a monotonicity of $(S_i^{(n)})$ by the many-to-one lemma, and by the fact that the limit is independent of the start type. $\qquad \square$

Lemma 2.13. *Let N_n be a positive sequence with $\lim_{n \to \infty} N_n = \infty$ and set $a_n = N_n^{1/3}$, $M_n = N_n^{1/5}$, $R_n = e^{N^{1/4}}$. Consider the Markov chain $(\tilde{S}_k^{(n)}, \tilde{\tau}_k^{(n)}, \tilde{\nu}_{k-1}^{(n)})_{k \in \mathbb{N}_0}$ with filtration $(\mathcal{F}_k^{(n)})_{k \in \mathbb{N}_0}$ and transitions given for any measurable F by*

$$\mathbb{E}[F(\tilde{S}_k^{(n)} - \tilde{S}_{k-1}^{(n)}, \tilde{\tau}_{\tau_k}^{(n)}, \tilde{\nu}_{k-1}^{(n)}) \mid \mathcal{F}_{k-1}^{(n)}] = \mathbb{E}_{(0, \tau_{k-1})}^{\alpha_n}[F(S_1^{(n)}, \tau_1^{(n)}, \nu_0^{(n)}) \mid \nu_0^{(n)}((-\infty, M_n)) \leq R_n, S_1^{(n)} \leq M_n].$$

where $(S_1^{(n)}, \tau_1^{(n)}, \nu_0^{(n)})$ is the first step of the Markov chain defined in Lemma 2.5 associated either to the IBRW with attachment rule f_{t_n} and $p = 1$ or with attachment rule f and percolation parameter p_n. Let $A > 0$, write $r_n = \lfloor A a_n^2 \rfloor$. For all $c_1, c_2, c_3, c_4 \in \mathbb{R}$,

$$\mathbb{P}_{(0,\tau)}\left(c_1 \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2 \, \forall i \in [r_n]; c_3 \leq \frac{\tilde{S}_{r_n}^{(n)}}{a_n} \leq c_4\right)$$

$$\to P(c_1 \leq \sqrt{\sigma^2 A} W_t \leq c_2 \, \forall t \in [0,1]; c_3 \leq \sqrt{\sigma^2 A} W_1 \leq c_4) \tag{2.10}$$

as $n \to \infty$, uniformly in $\tau \in \mathcal{T}$, where $(W_t : t \geq 0; P)$ is a standard Brownian motion.

Proof. We show that we can replace $\tilde{S}_i^{(n)}$ by $S_i^{(n)}$ in the above probability up to an error that converges to 0 uniformly in τ and then invoke Lemma 2.12. Define the event

$$E_n := \{\bar{\nu}_0^{(n)} \leq R_n, S_1^{(n)} - S_0^{(n)} \leq M_n\}.$$

and note that by Corollary 2.9, there exist constants $\tilde{C}, \tilde{\gamma} > 0$ such that

$$\inf_{\tau \in \mathcal{T}} \mathbb{P}_{(0,\tau)}(E_n) \geq 1 - \tilde{C} e^{-\tilde{\gamma} N^{1/5}}. \tag{2.11}$$

We estimate for any $c_1^i, c_2^i \in \mathbb{R}$

$$\mathbb{P}_{(0,\tau)}^{\alpha_n^*}\Big(c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n]\Big)$$

$$\leq \mathbb{E}_{(0,\tau)}^{\alpha_n^*}\bigg[\mathbb{1}\{c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n-1]\}\frac{\mathbb{P}_{(\tilde{S}_{r_n-1},\tilde{\tau}_{r_n-1})}^{\alpha_n^*}(c_1^{r_n} \leq S_1^{(n)} \leq c_2^{r_n})}{\mathbb{P}_{(0,\tilde{\tau}_{r_n-1})}^{\alpha_n^*}(E_n)}\bigg]$$

$$\leq \mathbb{E}_{(0,\tau)}^{\alpha_n^*}\bigg[\mathbb{1}\{c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n-1]\}\mathbb{P}_{(\tilde{S}_{r_n-1},\tilde{\tau}_{r_n-1})}^{\alpha_n^*}(c_1^{r_n} \leq S_1^{(n)} \leq c_2^{r_n})\bigg](1 + 2\tilde{C} e^{-\tilde{\gamma} N^{1/5}}),$$

where we assume that n is sufficiently large and we used (2.11) together with $\frac{1}{1-x} \leq 1 + 2x$ for $x \in [0, \frac{1}{2}]$. Iterating this estimate yields

$$\mathbb{P}_{(0,\tau)}^{\alpha_n^*}\Big(c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n]\Big) \quad \leq \mathbb{P}_{(0,\tau)}\Big(c_1^i \leq \frac{S_i^n}{a_n} \leq c_2^i \,\forall i \in [r_n]\Big)(1 + 2\tilde{C} e^{-\tilde{\gamma} N^{1/5}})^{r_n}. \tag{2.12}$$

and we note that by the choice of r_n the error converges to 0.

For a lower bound, we estimate

$$\mathbb{P}_{(0,\tau)}^{\alpha_n^*}\Big(c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n]\Big)$$

$$\geq \mathbb{E}_{(0,\tau)}^{\alpha_n^*}\bigg[\mathbb{1}\{c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n-1]\}\mathbb{P}_{(\tilde{S}_{r_n-1},\tilde{\tau}_{r_n-1})}^{\alpha_n^*}(c_1^{r_n} \leq S_1^{(n)} \leq c_2^{r_n}, E_n)\bigg]$$

$$\geq \mathbb{E}_{(0,\tau)}^{\alpha_n^*}\bigg[\mathbb{1}\{c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \,\forall i \in [r_n-1]\}\mathbb{P}_{(\tilde{S}_{r_n-1},\tilde{\tau}_{r_n-1})}^{\alpha_n^*}(c_1^{r_n} \leq S_1^{(n)} \leq c_2^{r_n})\bigg] - \sup_\tau \mathbb{P}_{(0,\tau)}^{\alpha_n^*}(E_n)$$

Iterating and using the bound (2.11) gives

$$\mathbb{P}_{(0,\tau)}^{\alpha_n^*}\Big(c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \forall i \in [r_n]\Big) \geq \mathbb{P}_{(0,\tau)}^{\alpha_n^*}\Big(c_1^i \leq \frac{\tilde{S}_i^{(n)}}{a_n} \leq c_2^i \forall i \in [r_n]\Big) - \tilde{C} r_n e^{-\tilde{\gamma} N^{1/5}}, \tag{2.13}$$

where the error converges to 0 by choice of r_n. Then, combining (2.12) and (2.13) together with Lemma 2.12 we can deduce the statement of the lemma. $\qquad\square$

3 Proofs: Upper bound

In this section, we fix the start type ℓ of the IBRW. Recall that in the killed IBRW every particle x with $S(x) > 0$ is deleted together with its descendants. We denote its survival probability by ζ, that is, for $s_0 \leq 0$,

$$\zeta_{s_0}(p, f) = P_{(s_0,\ell)}^p(\text{killed IBRW survives}).$$

Lemma 3.1. *For all $\alpha \in \mathcal{I}$, $s_0 \leq 0$, $n \in \mathbb{N}$ and $b_1, \ldots, b_n \geq 0$,*

$$\zeta_{s_0}(p, f) \leq e^{-\alpha s_0} \rho^p(\alpha)^n I(n) + \sum_{j=0}^{n-1} e^{-\alpha(b_{j+1}+s_0)} \rho^p(\alpha)^{j+1} I(j),$$

where $I(j) = \mathbb{P}_{(s_0,\ell)}^\alpha(-b_i \leq S_i \leq 0 \,\forall i \in [j])$ for $j = 0, \ldots, n$, where $((S_i, \tau_i)_{i \geq 0}, \mathbb{P}_{(s_0,\ell)})$ is the Markov chain from Lemma 2.4 started in (s_0, ℓ).

Proof. By definition,

$$\zeta_{s_0}(p, f) \leq P^p_{(s_0, \ell)}(\exists |x| = n \colon S(x_i) \leq 0 \; \forall \, i \in [n])$$

$$\leq P^p_{(s_0, \ell)}(\exists |x| = n \colon -b_i \leq S(x_i) \leq 0 \; \forall \, i \in [n]) \tag{3.1}$$

$$+ \sum_{j=1}^{n} P^p_{(s_0, \ell)}(\exists |x| = n \colon -b_i \leq S(x_i) \leq 0 \; \forall \, i \in [j-1], S(x_j) < -b_j).$$

For the first summand we use the Markov inequality and Lemma 2.4 to derive

$$P^p_{(s_0, \ell)}(\exists |x| = n \colon -b_i \leq S(x_i) \leq 0 \; \forall \, i \in [n])$$

$$= P^p_{(s_0, \ell)}\Big(\sum_{|x|=n} \mathbb{1}\{-b_i \leq S(x_i) \leq 0 \; \forall \, i \in [n]\} \geq 1 \Big)$$

$$\leq E^p_{(s_0, \ell)}\Big[\sum_{|x|=n} \mathbb{1}\{-b_i \leq S(x_i) \leq 0 \; \forall \, i \in [n]\} \Big]$$

$$= \rho^p(\alpha)^n \mathbb{E}^\alpha_{(s_0, \ell)}\Big[e^{\alpha(S_n - S_0)} \mathbb{1}\{-b_i \leq S_i \leq 0 \; \forall \, i \in [n]\} \frac{v_\alpha(\tau_0)}{v_\alpha(\tau_n)} \Big]$$

$$\leq \frac{v_\alpha(\ell)}{\inf_{\tau \in \mathcal{T}} v_\alpha(\tau)} \rho^p(\alpha)^n e^{-\alpha s_0} I(n) = \rho^p(\alpha)^n e^{-\alpha s_0} I(n), \tag{3.2}$$

where we used in the last step that by Lemma 2.2 $\inf_{\tau \in \mathcal{T}} v^p_\alpha(\tau) = v_\alpha(\ell)$. Similarly,

$$P^p_{(s_0, \ell)}(\exists |x| = n \colon -b_i \leq S(x_i) \leq 0 \; \forall \, i \in [j-1], S(x_j) < -b_j)$$

$$\leq P^p_{(s_0, \ell)}(\exists |x| = j \colon -b_i \leq S(x_i) \leq 0 \; \forall \, i \in [j-1], S(x_j) < -b_j)$$

$$= P^p_{(s_0, \ell)}\Big(\sum_{|x|=j} \mathbb{1}\{-b_i \leq S(x_i) \leq 0 \; \forall \, i \in [j-1], S(x_j) < -b_j\} \geq 1 \Big)$$

$$\leq E^p_{(s_0, \ell)}\Big[\sum_{|x|=j} \mathbb{1}\{-b_i \leq S(x_i) \leq 0 \; \forall \, i \in [j-1], S(x_j) < -b_j\} \Big]$$

$$= \rho^p(\alpha)^j \mathbb{E}^\alpha_{(s_0, \ell)}\Big[e^{\alpha(S_j - S_0)} \mathbb{1}\{-b_i \leq S_i \leq 0 \; \forall \, i \in [j-1], S_j < -b_j\} \frac{v_\alpha(\tau_0)}{v_\alpha(\tau_j)} \Big]$$

$$\leq \rho^p(\alpha)^j e^{\alpha(-b_j - s_0)} I(j-1). \tag{3.3}$$

Combining (3.1)–(3.3), now concludes the proof. $\qquad\square$

Lemma 3.2. *For all $\alpha \in \mathcal{I}$ with $\rho(\alpha) \geq 1$ and for all $s_0 \leq 0, k, N \in \mathbb{N}$ and $(b_j \colon j \in [kN])$ nonnegative and decreasing,*

$$\zeta_{s_0}(p, f) \leq e^{-\alpha s_0} \rho(\alpha)^{kN} I(kN) + k \sum_{l=0}^{N-1} e^{-\alpha(b_{(l+1)k} + s_0)} \rho(\alpha)^{(l+1)k} I(lk).$$

Proof. Using Lemma 3.1, that $j \mapsto I(j)$ is decreasing, $\rho(\alpha) \geq 1$ and that $j \mapsto e^{-\alpha b_j}$ is increasing to obtain

$$\zeta(p, f) - e^{-\alpha s_0} \rho(\alpha)^{kN} I(kN) \leq \sum_{j=0}^{kN-1} e^{-\alpha(b_{j+1} + s_0)} \rho(\alpha)^{j+1} I(j)$$

$$= \sum_{l=0}^{N-1} \sum_{j=lk}^{(l+1)k-1} e^{-\alpha(b_{j+1} + s_0)} \rho(\alpha)^{j+1} I(j)$$

$$\leq k \sum_{l=0}^{N-1} e^{-\alpha(b_{(l+1)k} + s_0)} \rho(\alpha)^{(l+1)k} I(lk). \qquad\square$$

18

For the proof of Theorem 1.1, let $(p_n)_{n\in\mathbb{N}}$ be a sequence of retention probabilities with $p_n \downarrow p_c$. For Theorem 1.2, let $(t_n)_{n\in\mathbb{N}}$ be a sequence of parameters with $t_n \uparrow \infty$. We write,

$$\rho_n(\cdot) = \begin{cases} \rho^{p_n}(\cdot) = p_n\rho(\cdot) & \text{for Theorem 1.1} \\ \rho_{t_n}(\cdot) & \text{for Theorem 1.2} \end{cases} \quad \text{and} \quad \alpha_n^* := \begin{cases} \alpha^* & \text{for Theorem 1.1} \\ \alpha_{t_n}^* & \text{for Theorem 1.2.} \end{cases}$$

Moreover, we denote by v_n the eigenfunction for $\rho_n(\alpha_n^*)$ from Lemma 2.2. The Markov chain from Lemma 2.4 corresponding to $\alpha = \alpha_n^*$ and retention parameter p_n or attachment rule f_{t_n} is denoted by $((S_i^{(n)}, \tau_i^{(n)}): i \in \mathbb{N}; \mathbb{P})$, i.e. $\mathbb{P} = \mathbb{P}_{(s_0,\ell)}^{\alpha_n^*}$. One easily checks that in the setup of Theorem 1.1 the distribution of the Markov chain does not depend on n.

Finally, we introduce for all $n \in \mathbb{N}$,

$$\epsilon_n := \log \rho_n(\alpha_n^*) \quad \text{and} \quad \sigma^2 := \lim_{n\to\infty} \frac{\rho_n''(\alpha_n^*)}{\rho_n(\alpha_n^*)}.$$

Notice, that in the situation of Theorem 1.1, $\sigma^2 = \rho''(\alpha^*)p_c$. The choice $\alpha = \alpha_n^*$ guarantees that for both theorems, $\epsilon_n \downarrow 0$ as $n \to \infty$.

Lemma 3.3. *Let $N \in \mathbb{N}, s_0 \leq 0$ and $a, b > 0$. Let $k_n = \lfloor (a/\epsilon_n)^{\frac{3}{2}} N^{-1} \rfloor$ and $b_i^n(s_0) = a^{-1/2}b(Nk_n - i)^{1/3} - s_0$, $i \in [Nk_n]$. Then, for all $C > 0$, $l \in [N]$,*

$$\limsup_{n\to\infty} \epsilon_n^{\frac{1}{2}} \log \Big(\sup_{-C/\sqrt{\epsilon_n} \leq s_0 \leq 0} I_n^{s_0}(lk_n) \Big) \leq -\frac{\pi^2\sigma^2}{2} \frac{l}{N} \frac{a^{3/2}}{(b+C)^2}.$$

where $I_n^s(j) = \mathbb{P}_{(s,\ell)}(-b_i^n(s) \leq S_i^{(n)} \leq 0 \,\forall\, i \in [j])$.

Notice that the specific choice of parameters implies that

$$I_n^{s_0}(lk_n) = \mathbb{P}_{(0,\ell)}^{\alpha_n^*}\Big(-a^{-1/2}b\Big(\frac{N}{l} - \frac{i}{lk_n}\Big)^{\frac{1}{3}} \leq \frac{S_i^{(n)}}{(lk_n)^{1/3}} \leq -\frac{s_0}{(lk_n)^{1/3}} \,\forall\, i \in [lk_n] \Big). \tag{3.4}$$

Proof. Denote $g_1(t) = -a^{-1/2}b(\frac{N}{l} - t)^{\frac{1}{3}}$ and let $\delta > 0$. For n sufficiently large, (3.4) yields

$$\sup_{-C/\sqrt{\epsilon_n} \leq s_0 \leq 0} I_n^{s_0}(lk_n) \leq \mathbb{P}_{(0,\ell)}^{\alpha_n^*}\Big(g_1\Big(\frac{i}{lk_n}\Big) \leq \frac{S_i^{(n)}}{(lk_n)^{1/3}} \leq \Big(\frac{l}{N}\Big)^{1/3} a^{-1/2}(C+\delta) \,\forall\, i \in [lk_n] \Big).$$

By Lemma 2.12 we can apply Theorem 2.10 with $a_n = (lk_n)^{1/3}$, $n = lk_n$ and $g_2(t) = a^{-1/2}(C+\delta)\frac{l}{N}$, and then take $\delta \searrow 0$ to derive

$$\limsup_{n\to\infty} \frac{1}{(lk_n)^{\frac{1}{3}}} \log \sup_{-C/\sqrt{\epsilon_n} \leq s_0 \leq 0} I_n^{s_0}(lk_n) \leq -\frac{\pi^2\sigma^2}{2} \int_0^1 \frac{1}{(a^{-1/2}b(\frac{N}{l} - t)^{\frac{1}{3}} + C(N/\ell)^{1/3}a^{-1/2})^2} \, dt$$

$$\leq -\frac{\pi^2\sigma^2}{2}\Big(\frac{l}{N}\Big)^{2/3} \frac{a}{(b+C)^2}.$$

Moreover, as $n \to \infty$,

$$\epsilon_n^{\frac{1}{2}}(lk_n)^{\frac{1}{3}} = l^{\frac{1}{3}}\Big(\epsilon_n^{\frac{3}{2}}\lfloor (a/\epsilon_n)^{\frac{3}{2}}N^{-1} \rfloor\Big)^{\frac{1}{3}} \to l^{\frac{1}{3}}\frac{a^{\frac{1}{2}}}{N^{\frac{1}{3}}}. \qquad \square$$

Proof of the upper bounds in Theorems 1.1 and 1.2. Since $\theta(p, f)$ is non-decreasing in retention probability p and attachment rule f, it suffices to consider the asymptotic behaviour along a discrete subsequence. As before, for Theorem 1.1 we take any discrete sequence of retention probabilities $p_n \downarrow p_c$ and for Theorem 1.2 we take any discrete sequence $t_n \uparrow \infty$. We make use of the notation introduced in

and before Lemma 3.3. In particular, we fix $N \in \mathbb{N}$, $a > 0$ and $b > 0$ and let $k_n := \lfloor (a/\epsilon_n)^{\frac{3}{2}} N^{-1} \rfloor$ and $b_i = b_i(s_0) = b_i^n = a^{-1/2} b (N k_n - i)^{\frac{1}{3}} - s_0$. Write

$$
\zeta_{s_0}^n = \begin{cases} \zeta_{(s_0,\ell)}(p_n, f) & \text{for Theorem 1.1,} \\ \zeta_{(s_0,\ell)}(1, f_{t_n}) & \text{for Theorem 1.2.} \end{cases}
$$

Applying Lemma 3.2 with $\alpha = \alpha_n^*$, we obtain for every $N \in \mathbb{N}$, $a > 0$, $b > 0$, $C > 0$

$$
\log \Big(\sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} \zeta_{s_0}^n \Big) - \log(N+1)
$$

$$
\le \max_{l \in [N-1]} \Big\{ \alpha_n^* C \epsilon_n^{-1/2} + k_n N \log \rho_n(\alpha_n^*) + \log \sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} I_n^{s_0}(k_n N),
$$

$$
\log k_n - \alpha_n^*(b_{k_n}^n + s_0) + k_n \log \rho_n(\alpha_n^*),
$$

$$
\log k_n - \alpha_n^*(b_{(l+1)k_n}^n + s_0) + (l+1)k_n \log \rho_n(\alpha_n^*) + \log \sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} I_n^{s_0}(l k_n) \Big\}
$$

$$
= \max_{l \in [N-1]} \Big\{ \alpha_n^* C \epsilon_n^{-1/2} + k_n N \epsilon_n + \log \sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} I_n^{s_0}(k_n N),
$$

$$
\log k_n - \alpha_n^* a^{-1/2} b (N k_n - k_n)^{\frac{1}{3}} + k_n \epsilon_n,
$$

$$
\log k_n - \alpha_n^* a^{-1/2} b (N k_n - (l+1)k_n)^{\frac{1}{3}} + (l+1)k_n \epsilon_n + \log \sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} I_n^{s_0}(l k_n) \Big\}.
$$

Recall that the choice of parameters implies that

$$
\epsilon_n^{\frac{3}{2}} k_n \to \frac{a^{\frac{3}{2}}}{N}, \quad \text{and} \quad \sqrt{\epsilon_n} \log k_n \to 0 \quad \text{as } n \to \infty.
$$

Hence, by Lemma 3.3,

$$
\limsup_{n \to \infty} \epsilon_n^{\frac{1}{2}} \log \Big(\sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} \zeta_{s_0}^n \Big)
$$

$$
\le \max_{l \in [N-1]} \Big\{ C\alpha^* + a^{\frac{3}{2}} - a^{\frac{3}{2}} \frac{\pi^2 \sigma^2}{2(b+C)^2}, 0 - \alpha^* b \frac{1}{N^{\frac{1}{3}}} (N-1)^{\frac{1}{3}} + \frac{a^{\frac{3}{2}}}{N},
$$

$$
0 - \alpha^* b \frac{1}{N^{\frac{1}{3}}} (N - (l+1))^{\frac{1}{3}} + (l+1) \frac{a^{\frac{3}{2}}}{N} - a^{\frac{3}{2}} \frac{\pi^2 \sigma^2}{2} \frac{1}{(b+C)^2} \Big(\frac{l}{N} \Big).
$$

Taking N to infinity, we deduce

$$
\limsup_{n \to \infty} \epsilon_n^{\frac{1}{2}} \log \big(\sup_{-C/\sqrt{\epsilon_n} \le s \le 0} \zeta_s^n \big) \le \max \Big\{ C\alpha^* + a^{\frac{3}{2}} \Big(1 - \frac{\pi^2 \sigma^2}{2} \frac{1}{(b+C)^2} \Big), -\alpha^* b,
$$

$$
\sup_{x \in (0,1]} \Big\{ -\alpha^* b (1-x)^{\frac{1}{3}} + a^{\frac{3}{2}} \Big(1 - \frac{\pi^2 \sigma^2}{2} \frac{1}{(b+C)^2} \Big) x \Big\}
$$

$$
= \max \Big\{ C\alpha^* + a^{\frac{3}{2}} \Big(1 - \frac{\pi^2 \sigma^2}{2} \frac{1}{(b+C)^2} \Big), \sup_{x \in [0,1]} \Big\{ -\alpha^* b (1-x)^{\frac{1}{3}} + a^{\frac{3}{2}} \Big(1 - \frac{\pi^2 \sigma^2}{2} \frac{1}{(b+C)^2} \Big) x \Big\}
$$

Now, we consider any $C < \sqrt{\frac{\pi^2 \sigma^2}{2}}$ and we choose b such that

$$
b < \sqrt{\frac{\pi^2 \sigma^2}{2}} - C.
$$

For these choices, we can take a sufficiently large to see that the first term in the above maximisation can be ignored, while in the second one the supremum is achieved at $x = 0$. This gives the bound

$$
\limsup_{n \to \infty} \epsilon_n^{\frac{1}{2}} \log \Big(\sup_{-C/\sqrt{\epsilon_n} \le s_0 \le 0} \zeta_{s_0}^n \Big) \le -\alpha^* b.
$$

Now, we can let $b \uparrow \sqrt{\frac{\pi^2\sigma^2}{2}} - C$ to see that for any $0 < C < \sqrt{\frac{\pi^2\sigma^2}{2}}$,

$$\limsup_{n\to\infty} \epsilon_n^{\frac{1}{2}} \log \Big(\sup_{-C/\sqrt{\epsilon_n} \leq s_0 \leq 0} \zeta_{s_0}^n \Big) \leq -\alpha^* \Big(\sqrt{\frac{\pi^2\sigma^2}{2}} - C \Big). \tag{3.5}$$

To complete the proof take $\delta = \frac{1}{2}(1 - \alpha^*) > 0$. Choose $\ell = \lfloor \alpha^* \frac{\pi\sigma}{\sqrt{2}\delta} \rfloor$ and note that $\ell\delta \leq \alpha^* \frac{\pi\sigma}{\sqrt{2}} < \frac{\pi\sigma}{\sqrt{2}}$. Then, we obtain from (3.5) using the monotonicity of $s \mapsto \zeta_s^n$,

$$
\begin{aligned}
\int_0^\infty e^{-s} \zeta_{-s}^n \, ds &\leq \sum_{k=0}^{\ell-1} \int_{k/\sqrt{\epsilon_n}}^{(k+1)/\sqrt{\epsilon_n}} e^{-s} \zeta_s^n \, ds + e^{-\ell\delta/\sqrt{\epsilon_n}} \\
&\leq \sum_{k=0}^{\ell-1} e^{1/\sqrt{\epsilon_n}(-\delta k - \alpha^* \frac{\pi\sigma}{\sqrt{2}} + \alpha^*(k+1)\delta)(1+o(1))} + e^{-(\alpha^* \frac{\pi\sigma}{2}\alpha^* - \delta)/\sqrt{\epsilon_n}} \\
&\leq e^{-1/\sqrt{\epsilon_n}(\alpha^* \frac{\pi\sigma}{\sqrt{2}} - 2\delta)(1+o(1))},
\end{aligned}
$$

where we used that $\alpha^* < 1$. Thus, we can deduce by first taking the limit $n \to \infty$ and then $\delta \downarrow 0$ that

$$\limsup_{n\to\infty} \sqrt{\epsilon_n} \log \int_0^\infty e^{-s} \zeta_{-s}^n \, ds \leq -\alpha^* \sqrt{\frac{\pi^2\sigma^2}{2}}.$$

This immediately implies Theorem 1.2 since $\sigma^2 = \rho''(\alpha^*)$. For Theorem 1.1, $p_c = 1/\rho(\alpha^*)$ implies

$$\epsilon_n = \log(p_n \rho(\alpha^*)) = \log(p_n/p_c) = \log \Big(1 + \frac{p_n - p_c}{p_c} \Big) = \frac{p_n - p_c}{p_c}(1 + o(1)).$$

so that Theorem 1.1 follows since $\sigma^2 = \rho''(\alpha^*)p_c$. $\qquad\square$

4 Proofs: Lower bound

The general strategy of the lower bound is to identify a subtree of the IBRW that has the same distribution as a Galton-Watson tree. For this carefully chosen subtree we then lower bound the survival probability, which in return gives the required lower bound on the survival probability of the IBRW. In Section 4.1 we collect some general facts about Galton-Watson trees, which we will then use in Section 4.2 to carry out the proof of the lower bound.

4.1 Galton-Watson lemmas

In order to show the lower bound we construct a Galton-Watson tree, whose particles are a subfamily of the killed branching random walk. To estimate the survival probability of this Galton-Watson tree, we will use the following general lemma due to [8] and we also recall the proof for completeness.

Lemma 4.1. *Let* GW *be a Galton-Watson tree and denote by X the number of children in the first generation and by q its extinction probability. Then, for all $r \leq \min\{\frac{1}{8}, q\}$,*

$$\mathbb{P}(\mathrm{GW} \neq \emptyset) \geq \mathbb{P}(X \neq 0) - 2r^{-2}\mathbb{P}(1 \leq X \leq r^{-2}) - 2r.$$

Proof. Denote by q the extinction probability of GW and by $s \mapsto g(s) = \mathbb{E}[s^X]$ the generating function of GW. Then, for every $r \in [0, q]$,

$$
\begin{aligned}
q = g(q) &= g(0) + \int_0^q g'(s) \, ds \\
&= g(0) + \int_0^{q-r} g'(s) \, ds + \int_{q-r}^q g'(s) \, ds \\
&\leq \mathbb{P}(X = 0) + g'(1 - r) + r, \tag{4.1}
\end{aligned}
$$

where we used that $g'(s)$ is increasing, and $g'(s) \leq 1$ for all $s \in [0, q_{n,\epsilon}]$. We continue by estimating

$$g'(1-r) = \mathbb{E}\big[X(1-r)^{X-1}\big] = \frac{1}{1-r}E\big[X(1-r)^X\big]$$

$$\leq \frac{1}{1-r}\mathbb{E}\big[Xe^{-rX}\big]$$

$$= \frac{1}{1-r}\Big(\mathbb{E}\big[Xe^{-rX}; X \leq r^{-2}\big] + \mathbb{E}\big[Xe^{-rX}; X > r^{-2}\big]\Big)$$

$$\leq \frac{1}{1-r}\Big(r^{-2}\mathbb{P}(1 \leq X \leq r^{-2}) + r^{-2}e^{-1/r}\Big), \tag{4.2}$$

where we used for the second summand that $u \mapsto ue^{-ru}$ is decreasing on $[r^{-1}, \infty)$ and $[r^{-2}, \infty) \subseteq [r^{-1}, \infty)$. Then, for all $r \in (0, \frac{1}{8}]$,

$$\frac{1}{1-r} \leq 2 \qquad \text{and} \qquad \frac{1}{1-r}r^{-2}e^{-1/r} \leq r. \tag{4.3}$$

Combining (4.2) and (4.3), we deduce

$$g'(1-r) \leq 2r^{-2}\mathbb{P}(1 \leq X \leq r^{-2}) + r \tag{4.4}$$

Rearranging (4.1) and (4.4), we conclude that for all $r \leq \min\{\frac{1}{8}, q\}$,

$$1 - q \geq \mathbb{P}(X_n \neq 0) - 2r^{-2}\mathbb{P}(1 \leq X_n \leq r^{-2}) - 2r, \tag{4.5}$$

as required. $\qquad\square$

We will also need an estimate which guarantees that a supercritical Galton-Watson process grows exponentially fast on survival with large probability.

Lemma 4.2. *For all $\theta_1 > 1 > \theta_2 > 0$ there exists $\delta > 0$ such that for any Galton-Watson process $(X_n \colon n \in \mathbb{N})$ with $X_0 = 1$ where mean offspring m, offspring distribution $(p_k)_{k \in \mathbb{N}_0}$, and extinction probability q satisfy $q, p_1 < \delta$ and $m > 1/\delta$, we have, for sufficiently large n,*

$$\mathbb{P}(X_n \leq \theta_1^n \mid X_n \geq 1) \leq \theta_2^n.$$

Proof. Denote by g the generating function of $(X_n \colon n \in \mathbb{N})$. By pruning the tree, i.e. removing all finite subtrees, we obtain a tree which on survival of the original process equals a Galton-Watson process $(\tilde{X}_n \colon n \in \mathbb{N})$ with $\tilde{p}_0 = 0$, $\tilde{p}_1 = f'(q) \leq p_1 + \frac{2q}{(1-q)^2}$, and the same mean as the original process. Hence $\delta > 0$ can be chosen such that the pruned process has arbitrarily small $\tilde{p}_1$ and arbitrarily large mean.

Under this assumption we now calculate

$$\mathbb{P}(\tilde{X}_n \leq \theta_1^n) \leq \sum_{j=1}^{\lceil n/2 \rceil} \mathbb{P}\big(\tilde{X}_n \leq \theta_1^n, \text{ first branching in generation } j\big)$$

$$+ \mathbb{P}\big(\text{ no branching up to generation } \lceil n/2 \rceil\big)$$

$$\leq \sum_{j=1}^{\lceil n/2 \rceil} \big(\mathbb{P}(\tilde{X}_{n-j} \leq \theta_1^n)\big)^2 \tilde{p}_1^{j-1} + \tilde{p}_1^{n/2}$$

$$\leq \tilde{p}_1^{-1}\mathbb{P}(\tilde{X}_n \leq \theta_1^n)\mathbb{P}(\tilde{X}_{\lceil n/2 \rceil} \leq \theta_1^n) + \tilde{p}_1^{n/2}.$$

Hence

$$\mathbb{P}(\tilde{X}_n \leq \theta_1^n) \leq \frac{\tilde{p}_1^{n/2}}{1 - \tilde{p}_1^{-1}\mathbb{P}(\tilde{X}_{\lceil n/2 \rceil} \leq \theta_1^n)}.$$

Choosing $\delta > 0$ so small that $\mathbb{E}\tilde{X}_1 > \theta_1^2$ and $\sqrt{\tilde{p}_1} < \theta_2$, for sufficiently large n, the denominator is larger than $1/2$ and the numerator is smaller than $\frac{1}{8}\theta_2^n$, so that

$$\mathbb{P}(\tilde{X}_n \leq \theta_1^n) \leq \frac{1}{4}\theta_2^n.$$

Note that if we condition X_n on extinction in distribution it is equal to a Galton-Watson process X_n^* with mean $f'(q) = \tilde{p}_1$. Therefore, by Markov's inequality

$$\mathbb{P}(X_n \geq 1 \,|\, \text{extinction}) = \mathbb{P}(X_n^* \geq 1) \leq \mathbb{E}[X_n^*] = \tilde{p}_1^n \leq \frac{1}{4}\theta_2^n,$$

by the same assumptions on δ as above for n large. We can also assume that δ is sufficiently small, so that the extinction probability q is less than $1/2$. Hence, combining the above estimates,

$$\mathbb{P}(X_n \leq \theta_1 \,|\, X_n \geq 1) \leq \frac{1}{\mathbb{P}(X_n \geq 1)}\Big\{\mathbb{P}(X_n \leq \theta_1 \,|\, \text{survival}) + \mathbb{P}(X_n \geq 1 \,|\, \text{extinction})\Big\}$$
$$\leq \frac{1}{1-q}\Big(\frac{1}{4}\theta_2^n + \frac{1}{4}\theta_2^n\Big) \leq \theta_2^n,$$

which completes the proof. $\qquad\square$

4.2 The lower bound

Throughout, we use the same notation as in the upper bound: For the proof of Theorem 1.1, let $(p_n)_{n\in\mathbb{N}}$ be a sequence of retention probabilities with $p_n \downarrow p_c$. For Theorem 1.2, let $(t_n)_{n\in\mathbb{N}}$ be a sequence of parameters with $t_n \uparrow \infty$. We denote by $S^{(n)}$ the positions either in the percolated IBRW or in the IBRW with attachment rule f_{t_n}. If the context is clear, we will omit the superscript. Also, we write,

$$\rho_n(\cdot) = \begin{cases} p_n\rho(\cdot) & \text{for Theorem 1.1} \\ \rho_{t_n}(\cdot) & \text{for Theorem 1.2} \end{cases} \quad\text{and}\quad \alpha_n^* := \begin{cases} \alpha^* & \text{for Theorem 1.1} \\ \alpha_{t_n}^* & \text{for Theorem 1.2}. \end{cases}$$

Moreover, we denote by v_n the eigenfunction for $\rho_n(\alpha_n^*)$ from Lemma 2.2.

Finally, we introduce for all $n \in \mathbb{N}$,

$$\epsilon_n := \log \rho_n(\alpha_n^*) \quad\text{and}\quad \sigma^2 := \lim_{n\to\infty} \frac{\rho_n''(\alpha_n^*)}{\rho_n(\alpha_n^*)}. \tag{4.6}$$

Notice, that in the situation of Theorem 1.1, $\sigma^2 = \rho''(\alpha^*)p_c$. The choice $\alpha = \alpha_n^*$ guarantees by Proposition 2.3 that for both theorems, $\epsilon_n \downarrow 0$ as $n \to \infty$.

Given any starting point $s \geq 0$ and initial type τ, we will write $P = P_{(s,\tau)} = P_{(s,\tau)}^{p_n}$ in the percolation case and $P = P_{(s,\tau)} = P_{(s,\tau)}^1$ in the case of Theorem 1.2.

In view of Lemma 4.1, we will now choose a Galton-Watson tree GW_n as a subtree of the killed IBRW in the following way, where we denote by $X^{(n)}$ the number of children in the first generation of GW_n:

(a) $P(X^{(n)} \neq 0) \approx$ the survival probability of the Galton-Watson process. That is, when there are offspring, then the process usually survives.

(b) $P(X^{(n)} \neq 0)$ is close to the survival probability of the killed IBRW. That means that we chose the subpopulation as a good approximation of the BRW and that the first inequality in (4.5) is a good estimate.

(c) $P(1 \leq X^{(n)} \leq r^{-2})$ has to be small to beat r^2.

The Galton-Watson tree is obtained by a coarse-graining procedure, which we now describe. It involves positive parameters b, λ, θ, M which will be chosen carefully at a later stage of the proof. We group together the first $N + o(N)$ generations in the IBRW to form the first generation in GW_n. It turns that we have to choose $N = N_n$ such that

$$N_n = \lfloor (b/\epsilon_n)^{3/2} \rfloor,$$

and for the first N steps we only choose particles whose positions are in the interval

$$I_{i,n} = [-\tfrac{\theta b}{N_n^{2/3}} i - \lambda N_n^{1/3}, -\tfrac{\theta b}{N_n^{2/3}} i].$$

To be precise, let $L(N) = N + \lfloor N^{1/3} \rfloor$ and

$$\mathcal{C}_n = \{x \colon |x| = L(N_n), S(x_i) - S(x_0) \in I_{i,n}, \text{ for } i = 1, \dots, N_n,$$
$$S(x_i) - S(x_{i-1}) \le M, \text{ for } i = N_n + 1, \dots, L(N_n)\}.$$

Then we define $\mathcal{C}_n$ to be the particles in the first generation of GW_n. We include the last $\lfloor N^{1/3} \rfloor$ generations to make sure that if we survive until time N, then we will have many particles by time $L(N)$.

We iterate the procedure, i.e. the children of $y \in \mathcal{C}_n$ will be

$$\mathcal{C}_n(y) = \{x \colon |x| = 2L(N_n), y \le x, S(x_i) - S(y) \in I_{i,n}, \text{ for } i = L(N_n) + 1, \dots, L(N_n) + N_n,$$
$$S(x_i) - S(x_{i-1}) \le M, \text{ for } i = L(N_n) + N_n + 1, \dots, 2L(N_n)\}.$$

and $\bigcup_{y \in \mathcal{C}_n} \mathcal{C}_n(y)$ will form the individuals of the second generation of GW_n and we continue in a similar way. Note that by the construction of $I_{i,n}$, we only include children in the second generation of the original IBRW that are to the left of the position of their ancestor. In particular, their distribution does not depend on the type of the parent. Moreover, all the conditions on the spatial positions are relative to $S(\emptyset)$. Therefore, the distribution of $\mathcal{C}_n$ does not depend on either the type of the root $\emptyset$ nor the initial position $S(\emptyset)$. Similarly, the distribution of $\mathcal{C}_n(y)$ does not depend on either type nor position of y. Hence, the number of individuals in the different generations really do form a Galton-Watson process.

Moreover, if we assume that

$$M < \theta b, \tag{4.7}$$

then we have that all particle positions satisfy $S(x_i) - S(x_0) \le 0$ for all $i \le |x|$ and $x \in \mathcal{C}_n$ and GW_n is really a subset of the killed branching random walk.

Coupling with a Galton-Watson process

To control the contribution of the last $N^{1/3}$ steps of the branching random walk, we use the following coupling: We can couple the IBRW with a modified IBRW, where in generations $kL(N) + N, \dots, kL(N) + L(N) - 1$, for $k \in \mathbb{N}_0$, the particles place their offspring according to the following rules relative to their own position:

(i) to the right, the positions of the offspring follow the jump times of the birth process $Z_t^{(f)}$ started in 0, which jumps from k to $k + 1$ at rate $f(k)$.

(ii) to the left, the positions of the offspring are given by a Poisson point process with intensity

$$e^t E[f(Z_{-t}^{(f)})] \mathbb{1}_{(-\infty, 0]}(t) \, dt,$$

where $(Z_t^{(f)})$ is the birth process with jump rate f started in 0.

(iii) Also, in these generations types do not play a role, so we define all particles to have the same type as $S(\emptyset)$ in the original process.

Now, we define the Galton-Watson $\underline{\mathrm{GW}}_n$ similarly to above in terms of the modified IBRW, where we again denote by $\underline{\mathcal{C}}_n$ the individuals in the first generation. Since (f_{t_k}) is decreasing, we can couple the processes such that if $\underline{\mathrm{GW}}_n$ survives then GW_n survives and also such that $|\underline{\mathcal{C}}_n| \le |\mathcal{C}_n|$.

For the lower bound on $P(\mathcal{C}_n \neq \emptyset)$ it turns out that it is enough to control the probability of the set

$$\mathcal{C}_n^{(\text{asym})} = \{x \colon |x| = N_n, S(x_i) - S(x_0) \in I_{i,n}, \text{ for } i = 1, \dots, N_n\}.$$

being non-empty. Note that for both, the original process and the modification, the set $\mathcal{C}_n^{(\text{asym})}$ is the same, however in the modified process, the next generations $N + 1, \dots, L(N)$ have the distribution of a single-type branching random walk. In particular, the number of particles in each generation form a Galton-Watson process, which we will denote by $(\underline{X}_k)_{k \in \mathbb{N}_0} = (\underline{X}_k(M))_{k \in \mathbb{N}_0}$. Moreover, we will denote by $\underline{q} = \underline{q}_M$ its extinction probability when started with a single particle. Note that $\underline{q}$ does not depend on n and we will use that by increasing M we have that $\lim_{M \to \infty} \underline{q}_M = 0$ and also $\lim_{M \to \infty} \mathbb{E}\underline{X}_1(M) = \infty$.

By the Markov property, the survival of the two subsets of the killed branching branching walk are related by

$$P(\mathcal{C}_n \neq \emptyset) \geq P(\underline{\mathcal{C}}_n \neq \emptyset) \geq (1 - \underline{q}_M)\mathbb{P}(\mathcal{C}_n^{(\text{asym})} \neq \emptyset). \tag{4.8}$$

The next result is the key step in the overall lower bound, where we will bound the probability that $\mathcal{C}_n^{(\text{asym})}$ is non-empty.

Proposition 4.3. *For any $\alpha^*\theta < 1$ and any $\lambda^2 > \frac{\pi^2 \sigma^2}{2b(1 - \alpha^*\theta)}$, we have for any initial position $s \geq 0$ and initial type $\tau \in \mathcal{T}$,*

$$\liminf_{n \to \infty} N_n^{-1/3} \log P_{(s,\tau)}\big(\mathcal{C}_n^{(\text{asym})} \neq \emptyset\big) \geq -\lambda\alpha^*.$$

Proof. By construction the probability of the event $\mathcal{C}_n^{(\text{asym})} \neq \emptyset$ does not depend on the starting point of the initial point nor its type, so we can assume that $S^{(n)}(\emptyset) = 0$ and $\tau(\emptyset) = 0$.

For the first part of the proof, we will omit the indices n, whenever the context is clear and we are not dealing with asymptotic statements. In particular, we will write $N = N_n, S = S^{(n)}, \alpha = \alpha_n$, etc. Also, in this proof $\rho = \rho^{p_n}$ in the percolation case and $\rho = \rho_{t_n}$ if the attachment rule is varying.

The first step is to carefully, select the relevant particles in $\mathcal{C}_n^{(\text{asym})}$. For $R_n = e^{N^{1/4}}$ and $M_n = N^{1/5}$ and an individual x, we write

$$\Delta S(x_i) = S(x_i) - S(x_{i-1}), \quad i \leq |x|.$$

Recall that

$$\nu_x = \sum_{y \colon y- = x} \delta_{\Delta S(y)}.$$

For any $|x| = N$, let

$$\bar{\nu}_{x_{i-1}} := \nu_{x_{i-1}}\big((-\infty, M_n)\big).$$

Define

$$Y_n = \#\{x \colon |x| = N, S(x_i) \in I_{i,n}, \bar{\nu}_{x_{i-1}} \leq R_n, \Delta S(x_i) \leq M_n \, \forall i \in [N]\}.$$

The Paley-Zygmund inequality yields

$$P(\mathcal{C}_n^{(\text{asym})} \neq \emptyset) \geq P(Y_n \geq 1) \geq \frac{E[Y_n]^2}{E[Y_n^2]}. \tag{4.9}$$

The remaining proof proceeds as follows: in the first step we find an easier upper bound on $\mathbb{E}[Y_n^2]$, which we can estimate using the many-to-one lemma in the second step. In Step 3, we find a lower bound on $\mathbb{E}[Y_n]$, which we will then combine with the other steps to obtain our claim.

Step 1: Upper bound on $\mathbb{E}[Y_n^2]$. First write

$$E[Y_n^2] = E\Big[\sum_{|x|=N} \sum_{|y|=N} \mathbb{1}\{S(x_i) \in I_{i,n}, \bar{\nu}_{x_{i-1}} \leq R_n, \Delta S(x_i) \leq M_n \, \forall i \leq N\}$$
$$\times \mathbb{1}\{S(y_i) \in I_{i,n}, \bar{\nu}_{y_{i-1}} \leq R_n, \Delta S(y_i) \leq M_n \, \forall i \leq N\}\Big]$$

25

We split the sum over y according to the last time $j \in \{0, \dots, N\}$ that the ancestors of y agree with x to obtain

$$E[Y_n^2] = E\Big[\sum_{|x|=N} \mathbb{1}\{S(x_i) \in I_{i,n}, \bar{\nu}_{x_{i-1}} \le R_n, \Delta S(x_i) \le M_n \, \forall i \in [N]\}$$

$$\times \sum_{j=0}^{N} \sum_{|y|=N} \mathbb{1}\{y_i = x_i \forall i \le j, y_{j+1} \neq x_{j+1}, S(y_i) \in I_{i,n}, \bar{\nu}_{y_{i-1}} \le R_n, \Delta S(y_i) \le M_n \, \forall i \in [N]\}\Big].$$

Conditioning on the $j+1$ first generations and using the independence of the branching process and the fact that on $\{\bar{\nu}_{x_j} \le R_n\}$, we only have to consider at most R_n relevant siblings of x_{j+1}, we obtain the upper bound

$$E[Y_n^2] \le E\Big[\sum_{|x|=N} \mathbb{1}\{S(x_i) \in I_{i,n}, \bar{\nu}_{x_{i-1}} \le R_n, \Delta S(x_i) \le M_n \, \forall i \in [N]\}\Big]\Big(1 + \sum_{j=0}^{N-1} R_n h_{j+1,n}\Big),$$

where $h_{N,n} = 1$ and for $j \le N - 2$,

$$h_{j,n} := \sup_{s_0 \in I_{j,n}, \tau_0 \in \mathcal{T}} E_{(s_0, \tau_0)}\Big[\sum_{|y'|=N-j} \mathbb{1}\{S(y_i') \in I_{i+j,n} \, \forall i \le N - j)\}\Big].$$

In particular, we have shown that

$$E[Y_n^2] \le E[Y_n]\Big(1 + \sum_{j=0}^{n-1} h_{j+1,n} R_n\Big),$$

leading to

$$P(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset) \ge \frac{E[Y_n]}{1 + \sum_{j=0}^{n-1} h_{j+1,n} R_n}. \tag{4.10}$$

Step 2: Upper bound on $h_{j,n}$. First, we note since the left, resp. right, end point of $I_{k,n}$ are to the left of the left, resp. right, end point of $I_{j,n}$ for $k \ge j$, we can apply the monotonicity in the initial types to deduce that

$$h_{j,n} \le \sup_{s_0 \in I_{j,n}} E_{(0,0)}\Big[\sum_{|y'|=N-j} \mathbb{1}\{S(y_i') + s_0 \in I_{i+j,n}, \, \forall i \le N - j\}\Big]$$

Thus, we obtain by the many-to-one Lemma 2.4

$$h_{j,n} \le \sup_{s_0 \in I_{j,n}} E_{(0,0)}\Big[\sum_{|y'|=N-j} \mathbb{1}\{S(y_i') + s_0 \in I_{i+j,n} \, \forall i \le N - j\}\Big]$$

$$= \sup_{s_0 \in I_{j,n}} \mathbb{E}_{(0,0)}^\alpha\Big[e^{\alpha S_{N-j}} \frac{v_\alpha(\tau_0)}{v_\alpha(\tau_N)} \rho(\alpha)^{N-j} \mathbb{1}\{S_i + s_0 \in I_{i+j,n} \, \forall i \le N - j\}\Big]$$

By the monotonicity in types, Lemma 2.2, we have that $v_\alpha(\tau_0)/v_\alpha(\tau_N) \le v_\alpha(0)/v_\alpha(\ell)$ and further by Proposition 2.3, $C := \sup_n \frac{v_{\alpha_n}^{(n)}(0)}{v_{\alpha_n}^{(n)}(\ell)} < \infty$, so that

$$h_{j,n} \le C \sup_{s_0 \in I_{j,n}} \mathbb{E}_{(0,0)}^\alpha\Big[e^{\alpha S_{N-j}} \rho(\alpha)^{N-j} \mathbb{1}\{S_i + s_0 \in I_{i+j,n} \, \forall i \le N - j\}\Big]$$

$$\le C \sup_{u \in [-\lambda N^{1/3}, 0]} e^{-\alpha \theta \epsilon_n (N-j) + \lambda \alpha N^{1/3}} \rho(\alpha)^{N-j} \mathbb{P}_{(0,0)}^\alpha\big(S_i + u - \tfrac{\theta b}{N^{2/3}} j \in I_{i+j,n} \, \forall i \le N - j\}\big), \tag{4.11}$$

We now would like to apply Mogulskii's theorem to estimate the probability, but we need a bound that works uniformly for all j (since we will be summing over j) and uniformly in u. We thus approximate

26

the sum over j by a finite sum and also split up the interval $[-\lambda N^{1/3}, 0]$ into smaller subintervals. So fix $\kappa \in \mathbb{N}$ and define
$$K := K_n := \lfloor N/\kappa \rfloor.$$

For the next estimate, suppose that $j \in [(j'-1)K, j'K - 1]$ for some $j' \in \{1, \ldots, \kappa - 1\}$ and assume $u \in [-q\lambda N^{1/3}/\kappa, -(q-1)\lambda N^{1/3}/\kappa]$ for some $q \in \{1, \ldots, \kappa\}$. Then,

$$
\begin{aligned}
\big\{ S_i + u &- \tfrac{\theta b}{N^{2/3}} j \in I_{i+j,n} \, \forall i \le N - j \} \} \\
&= \big\{ -\tfrac{\theta b}{N^{2/3}}(i+j) - \lambda N^{1/3} \le S_i + u - \tfrac{\theta b}{N^{2/3}} j \le -\tfrac{\theta b}{N^{2/3}}(i+j) \, \forall i \le N - j \big\} \\
&\subseteq \big\{ -\tfrac{\theta b}{N^{2/3}} i - \lambda N^{1/3} \le S_i + u \le -\tfrac{\theta b}{N^{2/3}} i \, \forall i \le (\kappa - j')K \big\} \\
&\subseteq \big\{ (q-1)\lambda N^{1/3}/\kappa - \tfrac{\theta b}{N^{2/3}} i - \lambda N^{1/3} \le S_i \le -\tfrac{\theta b}{N^{2/3}} i + q\lambda N^{1/3}/\kappa \, \forall i \le (\kappa - j')K \big\} \\
&\subseteq \big\{ (q-1)\lambda/\kappa - \tfrac{\theta b}{N} i - \lambda \le \tfrac{S_i}{N^{1/3}} \le -\tfrac{\theta b}{N} i + q\lambda/\kappa \, \forall i \le (\kappa - j')K \big\} \\
&\subseteq \Big\{ \tfrac{(q-1-\kappa)\lambda}{\kappa} - \theta b \tfrac{\kappa - j'}{\kappa} \tfrac{i}{(\kappa-j')K} \le \tfrac{S_i}{N^{1/3}} \le -\theta b \tfrac{\kappa-j'-1}{\kappa} \tfrac{i}{(\kappa-j')K} + \tfrac{q\lambda}{\kappa} \, \forall i \le (\kappa-j')K \Big\} \\
&=: E_{j',n},
\end{aligned}
$$

where we assumed that n is sufficiently large so that

$$
\frac{i}{N} \le \frac{\kappa - j'}{\kappa} \frac{i}{(\kappa - j')K} \quad \text{and} \quad \frac{i}{N} \ge \frac{\kappa - j'}{\kappa} \frac{i}{(\kappa - j')K} \frac{K}{K+1} \ge \frac{\kappa - j'}{\kappa} \Big(1 - \frac{1}{\kappa - j'}\Big) \frac{i}{(\kappa - j')K}.
$$

By Lemma 2.12 we can apply Mogulskii's theorem, Theorem 2.10 with $g_1(t) = \frac{(q-1-\kappa)\lambda}{\kappa} - \theta b \frac{\kappa-j'}{\kappa} t$ and $g_2(t) = -\theta b \frac{\kappa-j'-1}{\kappa} t + \frac{q\lambda}{\kappa}$.

$$
\lim_{n\to\infty} \frac{N_n^{2/3}}{(\kappa - j')K_n} \log \mathbb{P}_{(0,0)}^{\alpha_n}(E_{j',n}) = -\frac{\pi^2 \sigma^2}{2} \int_0^1 \frac{1}{(g_2(t) - g_1(t))^2} \, dt =: -C(\kappa),
$$

where σ is defined in (4.6) and we noticed that the integral does not depend on q nor j'.

Hence, applying the display (4.11), where we use the definition of $\epsilon_n = \rho(\alpha)$, and also that $\alpha_n \theta < 1$ (since $\alpha^* \theta < 1$, for all n sufficiently large, uniformly in $j \in [(j'-1)K, j'K - 1]$, we find that

$$
\begin{aligned}
\log h_{j,n} &\le (-\alpha_n \theta + 1)\epsilon_n(N_n - j) + \lambda \alpha_n N_n^{1/3} - \frac{(\kappa - j')K_n}{N_n^{2/3}} C(\kappa)(1 + o(1)) \\
&\le (-\alpha_n \theta + 1) \frac{b}{(N_n + 1)^{2/3}}(N_n - j) + \alpha_n \lambda N_n^{1/3} - \frac{(\kappa - j')K_n}{N_n^{2/3}} C(\kappa)(1 + o(1)) \\
&\le N^{1/3}(-\alpha^* \theta b + b + \alpha^* \lambda - C(\kappa))(1 + o(1)) + j'K_n(-b(1 - \alpha^* \theta) + C(\kappa)))(1 + o(1)).
\end{aligned}
$$

Note that $C(\kappa) \to \frac{\pi^2 \sigma^2}{2\lambda^2}$ as $\kappa \to \infty$. However, by assumption we have that $\frac{\pi^2 \sigma^2}{2\lambda^2} < b(1 - \alpha^* \theta)$, so we can assume that κ is large enough such that $C(\kappa) < b(1 - \alpha^* \theta)$. In this case, we can deduce that

$$\log h_{j,n} \le N^{1/3}(-\alpha^* \theta b + b + \alpha^* \lambda - C(\kappa))(1 + o(1)). \tag{4.12}$$

Moreover, if $j \in [(\kappa - 1)K + 1, N]$, then we can use the trivial bound

$$\log(h_{j,n}/C) \le \alpha_n \lambda N^{1/3} + (N - j)\epsilon_n \le \alpha_n \lambda N^{1/3} + \frac{b}{N^{2/3}} 2K \le N^{1/3}(\alpha^* \lambda + 2b/\kappa)(1 + o(1)). \tag{4.13}$$

Consequently, if we combine (4.12) and (4.13) we obtain

$$\limsup_{n\to\infty} N_n^{-1/3} \log \sum_{j=1}^N h_{j,n} \le b(1 - \alpha^* \theta + 2/\kappa) + \alpha^* \lambda - C(\kappa).$$

Finally, letting $\kappa \to \infty$ gives

$$\limsup_{n\to\infty} N_n^{-1/3} \log \sum_{j=1}^N h_{j,n} \leq -\alpha^*\theta b + b + \alpha^*\lambda - \frac{\pi^2\sigma^2}{2\lambda^2}.$$

Step 3: Lower bounding $E[Y_n]$. Let $\theta_1 > \theta$, then if we set $\Delta S_i = S_i - S_{i-1}$ and $\bar{\nu}_i = \nu_i((-\infty, M_n))$, we have by Lemma 2.5 that

$$E[Y_n] = \mathbb{E}^\alpha_{(0,0)}\Big[e^{\alpha S_N}\rho(\alpha)^N \frac{v_\alpha(0)}{v_\alpha(\tau_N)} \mathbb{1}_{\{S_i \in I_{i,n}, \bar{\nu}_{i-1} \leq R_n, \Delta S_i \leq M_n \forall i \leq N\}}\Big]$$

$$\geq \mathbb{E}^\alpha_{(0,0)}[e^{\alpha S_N}\rho(\alpha)^N \mathbb{1}_{\{S_i \in I_{i,n}, \bar{\nu}_{i-1} \leq R_n, \Delta S_i \leq M_n \forall i \leq N, S_N \geq -\theta_1 b N^{1/3}\}}]$$

$$\geq e^{-\alpha\theta_1 b N^{1/3}}\rho(\alpha)^N \mathbb{P}^\alpha_{(0,0)}\big(S_i \in I_{i,n}, \bar{\nu}_{i-1} \leq R_n, \Delta S_i \leq M_n \forall i \leq N, S_N \geq -\theta_1 b N^{1/3}\big),$$

where we also used that by the monotonicity of types, see Lemma 2.2, $v_\alpha(0) \geq v_\alpha(s)$ for any $s \in \mathcal{T}$. Define the Markov chain $(\tilde{S}_k, \tilde{\tau}_k, \tilde{\nu}_{k-1})_{k\in\mathbb{N}}$ with associated filtration $(\mathcal{F}_k)$ and transitions given for any measurable F by

$$\mathbb{E}^\alpha[F(\tilde{S}_k - \tilde{S}_{k-1}, \tilde{\tau}_{\tau_k}, \tilde{\nu}_{k-1}) \,|\, \mathcal{F}_{k-1}] = \mathbb{E}^\alpha_{(0,\tau_{k-1})}[F(S_1, \tau_1, \nu_0) \,|\, \nu_0((-\infty, M_n)) \leq R_n, S_1 \leq M_n].$$

Then, we can continue the previous display as

$$E[Y_n] \geq e^{-\alpha\theta_1 b N^{1/3}}\rho(\alpha)^N \mathbb{P}^\alpha_{(0,0)}\big(\tilde{S}_i \in I_{i,n} \forall i \leq N, \tilde{S}_N \geq -\theta_1 b N^{1/3}\big) \\ \times \inf_{\tau\in\mathcal{T}} \mathbb{P}^\alpha_{(0,\tau)}(\bar{\nu}_0 \leq R_n, \Delta S_1 \leq M_n)^N. \tag{4.14}$$

Note that

$$\mathbb{P}^\alpha_{(0,0)}\big(\tilde{S}_i \in I_{i,n} \forall i \in \leq N, \tilde{S}_N \geq -\theta_1 b N^{1/3}\big) = \mathbb{P}^\alpha_{(0,0)}\Big(-\frac{\theta b}{N}i - \lambda \leq \frac{\tilde{S}_i}{N^{1/3}} \leq -\frac{\theta b}{N}i \forall i \leq N, \tilde{S}_N \geq -\theta_1 b N^{1/3}\Big)$$

By Lemma 2.13 we can apply Mogulskii's Theorem 2.10 also to $(\tilde{S}_k)_{k\in\mathbb{N}_0}$ with $g_1(t) = -\theta bt - \lambda$, $g_2(t) = -\theta bt$, so that we get

$$\lim_{n\to\infty} N_n^{-1/3} \log \mathbb{P}^{\alpha_n}_{(0,0)}\Big(-\frac{\theta b}{N}i - \lambda \leq \frac{\tilde{S}_i^{(n)}}{N^{1/3}} \leq -\frac{\theta b}{N}i \,\forall i \leq N, \tilde{S}_N \geq -\theta_1 b N^{1/3}\Big) = -\frac{\pi^2\sigma^2}{2\lambda^2}. \tag{4.15}$$

We estimate the last term in (4.14) by deducing from Corollary 2.9 that there exist $\tilde{C}, \tilde{\gamma} > 0$ such that

$$\inf_{\tau\in\mathcal{T}} \mathbb{P}^\alpha_{(0,\tau)}(\bar{\nu}_0 \leq R_1, \Delta S_1 \leq M_n)^N \geq (1 - \tilde{C}e^{-\tilde{\gamma}N^{1/5}})^N,$$

so that in particular,

$$\lim_{N\to\infty} N^{-1/3} \log\Big(\inf_{\tau\in\mathcal{T}} \mathbb{P}^\alpha_{(0,\tau)}(\bar{\nu}_0 \leq R_n, \Delta S_1 \leq M_n)^N\Big) = 0. \tag{4.16}$$

Combining (4.15) and (4.16), we get from (4.14) if we recall the definition of ϵ_n that

$$\liminf_{n\to\infty} N_n^{-1/3} \log E[Y_n] \geq -\alpha\theta_1 b + b - \frac{\pi^2\sigma^2}{2\lambda^2}.$$

Step 4: Combining the estimates.

Combining the upper bound on $\sum_{j=1}^N h_{j,n}$ with the lower bound on $E[Y_n]$ as well as the fact that $R_n = e^{N^{1/4}}$, we obtain from the second moment bound (4.10),

$$\liminf_{n\to\infty} N_n^{-1/3} \log P\big(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset\big) \geq -\alpha\theta_1 b + b - \frac{\pi^2\sigma^2}{2\lambda^2} + \alpha b\theta - b - \alpha\lambda + \frac{\pi^2\sigma^2}{2\lambda^2} = -\alpha\lambda + \alpha b(\theta - \theta_1).$$

Finally, we can let $\theta_1 \downarrow \theta$ to obtain the claim of the proposition. $\qquad\square$

Proof of Theorems 1.1 and 1.2 – lower bounds. By the same argument as at the end of the proof of the upper bound, we have completed the proofs of Theorems 1.1 and 1.2 if we can show that if ζ_n is either $\theta(p_n, f)$ or $\theta(1, f_{t_n})$, then

$$\lim_{n \to \infty} \sqrt{\epsilon_n} \log \zeta_n \geq -\sqrt{\frac{\pi^2 \sigma^2}{2}} \alpha^*. \tag{4.17}$$

In our above construction, we first of all choose the constant M large enough such that the Galton-Watson process $(\underline{X}_k(M))_{k \in \mathbb{N}_0}$ satisfies the assumptions of Lemma 4.2 for $\theta_1 = 2$ and $\theta_2 = 1/2$. Also, we can assume that its survival probability satisfies $1 - \underline{q}_M \geq 1/2$. Let θ be such that $\alpha^* \theta < 1$. Then, choose b large enough such that $\theta b > M$, so that in particular GW_n is a subset of the killed IBRW, cf. (4.7). Additionally, we require that

$$b > \frac{\pi^2 \sigma^2}{(1 - \alpha^* \theta)} \left(\frac{2\alpha^*}{\log 2} \right)^2. \tag{4.18}$$

By Proposition 4.3 we obtain that for any $\lambda > \sqrt{\frac{\pi^2 \sigma^2}{2b(1 - \alpha^* \theta)}}$

$$\liminf_{n \to \infty} N_n^{-1/3} \log P\big(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset\big) \geq -\lambda \alpha^*. \tag{4.19}$$

Now, by (4.18) we have that

$$\sqrt{\frac{\pi^2 \sigma^2}{2b(1 - \alpha^* \theta)}} < \frac{\log 2}{2\sqrt{2}\, \alpha^*},$$

so we can additionally assume that λ is small enough such that

$$2\lambda \alpha^* < \log 2.$$

Then, if we define

$$r := \frac{1}{16} P(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset),$$

we obtain by (4.19) that

$$r^{-2} 2^{-N^{1/3}} \leq e^{N^{1/3}(2\lambda \alpha^* - \log 2 + o(1))} \to 0,$$

We will use the following general fact (see [8, Fact 4.2], but also [9, Lemma 5.2]): let $X_1, \ldots, X_k$ be independent non-negative random variables and suppose $F : (0, \infty) \to [0, \infty)$ is non-increasing, then

$$E\Big[\sum_{i=1}^k F(X_i) \,\Big|\, \sum_{i=1}^k X_i > 0 \Big] \leq \max_{1 \leq i \leq k} E[F(X_i) \,|\, X_i > 0]. \tag{4.20}$$

We will eventually apply Lemma 4.1 to $\underline{\mathrm{GW}}_n$ and thus first estimate $P(1 \leq |\underline{\mathcal{C}}_n| \leq r^{-2})$. We write $\underline{S}$ for the positions in the modified branching random walk used in the definition of $\underline{\mathrm{GW}}_n$. For all sufficiently large n, by using (4.20) in the second step, we obtain

$$
\begin{aligned}
P(1 \leq |\underline{\mathcal{C}}_n| \leq r^{-2}) &\leq P(1 \leq |\underline{\mathcal{C}}_n| \leq 2^{N^{1/3}}) \\
&= P\Big(1 \leq \sum_{x \in \mathcal{C}_n^{(\mathrm{asym})}} \#\{y > x \,:\, |y| = L(N)\,, \Delta \underline{S}(y_i) \leq M, i = N, \ldots, L(N)\} \leq 2^{N^{1/3}} \Big) \\
&\leq P(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset) P\big(\underline{X}_{\lfloor N^{1/3} \rfloor} \leq 2^{N^{1/3}} \,\big|\, \underline{X}_{\lfloor N^{1/3} \rfloor} \geq 1 \big) \\
&\leq P(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset)\, 2^{-N^{1/3}} \\
&= P(\mathcal{C}_n^{(\mathrm{asym})} \neq \emptyset)\, r^2 o(1),
\end{aligned}
$$

where we used Lemma 4.2 for the last inequality.

Finally, we can lower bound the survival probability of the killed IBRW by the survival probability of $\underline{\mathrm{GW}}_n$, where we recall that the distribution of $\underline{\mathrm{GW}}$ does not depend on the initial position and the initial

type. Also, we deduce from the upper bound shown in Section 3 that necessarily $P(\underline{\text{GW}}_n \neq \emptyset) \to 0$, so that the assumption $r \leq \min\{\frac{1}{8}, P(|\underline{\text{GW}}_n| < \infty)\}$ is satisfied for large n and we can apply Lemma 4.1 to obtain

$$\zeta_n \geq P(\underline{\text{GW}}_n \neq \emptyset)$$
$$\geq P(\underline{\mathcal{C}}_n \neq \emptyset) - 2r^{-2}P(1 \leq |\underline{\mathcal{C}}_n| \leq r^{-2}) - 2r$$
$$\geq (1 - \underline{q}_m)P(\mathcal{C}_n^{(\text{asym})} \neq \emptyset) - o(1)P(\mathcal{C}_n^{(\text{asym})} \neq \emptyset) - 2\frac{1}{16}P(\mathcal{C}_n^{(\text{asym})} \neq \emptyset)$$
$$\geq \frac{1}{4}P(\mathcal{C}_n^{(\text{asym})} \neq \emptyset).$$

Hence, we get from (4.19) that

$$\lim_{n \to \infty} N_n^{-1/3} \log \zeta_n \geq -\lambda \alpha^*.$$

Finally, letting $\lambda \downarrow \sqrt{\frac{\pi^2 \sigma^2}{2b(1 - \alpha\theta)}}$ and noting that $\epsilon_n = bN^{-2/3}(1 + o(1))$, we obtain

$$\lim_{n \to \infty} \epsilon_n^{1/2} \log \zeta_n \geq -\alpha^* \sqrt{\frac{\pi^2 \sigma^2}{2(1 - \alpha^*\theta)}}$$

Hence, (4.17) follows by letting $\theta \downarrow 0$. $\qquad\square$

5 Proofs in the linear case

In this section, we show how to deduce Corollary 1.3 from our general result, Theorem 1.2.

Proof of Corollary 1.3. In the proof of Proposition 1.3 in [6] it was shown that for linear attachment functions, the spectral radius of A_α equals the largest eigenvalue of

$$\begin{pmatrix} \frac{\beta}{\alpha - \gamma} & \frac{\beta}{1 - \gamma - \alpha} \\ \frac{\beta + \gamma}{\alpha - \gamma} & \frac{\beta}{1 - \gamma - \alpha} \end{pmatrix}.$$

This eigenvalue is given by

$$\rho(\alpha) = \frac{1}{2(1 - \gamma - \alpha)(\alpha - \gamma)} \left[\beta(1 - 2\gamma) + \sqrt{\beta^2(1 - 2\gamma)^2 + 4\beta\gamma(1 - \gamma - \alpha)(\alpha - \gamma)} \right]. \tag{5.1}$$

In particular, $\alpha^* = \frac{1}{2}$ regardless of the choice of $\gamma \in [0, \frac{1}{2})$ and $\beta \in (0, 1]$, and

$$\rho(\alpha^*) = \frac{1}{\frac{1}{2} - \gamma} \left[\beta + \sqrt{\beta^2 + \beta\gamma} \right].$$

In order to apply Theorem 1.2, we need to determine $\rho''(\alpha^*)$. To this end, we write $\Theta(\alpha)$ for the large squared bracket in (5.1) and

$$\varphi(\alpha) = (1 - \gamma - \alpha)(\alpha - \gamma) \quad \Rightarrow \quad \varphi'(\alpha) = 1 - \gamma - \alpha - \alpha + \gamma = 1 - 2\alpha.$$

Then

$$\rho'(\alpha) = -\frac{1}{2}\frac{\varphi'(\alpha)}{\varphi(\alpha)^2}(\beta(1 - 2\gamma) + \Theta(\alpha)) + \frac{1}{2\varphi(\alpha)}\left[0 + \frac{1}{2}\left(\beta^2(1 - 2\gamma)^2 + 4\beta\gamma\varphi(\alpha)\right)^{-\frac{1}{2}}\left(0 + 4\beta\gamma(1 - 2\alpha)\right) \right]$$
$$= \frac{2\alpha - 1}{\varphi(\alpha)}\rho(\alpha) + \frac{4\beta\gamma(1 - 2\alpha)}{4\varphi(\alpha)}\left(\beta^2(1 - 2\gamma)^2 + 4\beta\gamma\varphi(\alpha)\right)^{-\frac{1}{2}}$$
$$= (2\alpha - 1)\frac{1}{\varphi(\alpha)}\left[\rho(\alpha) - \frac{\beta\gamma}{\sqrt{\beta^2(1 - 2\gamma)^2 + 4\beta\gamma\varphi(\alpha)}}\right]. \tag{5.2}$$

30

We need the second derivative only for $\alpha = \alpha^* = \frac{1}{2}$. Since after applying the product rule, any term multiplied by $(2\alpha - 1)$ vanishes, we obtain

$$\rho''(\tfrac{1}{2}) = \frac{2}{\varphi(\alpha)}\Big[\rho(\alpha) - \frac{\beta\gamma}{\sqrt{\beta^2(1-2\gamma)^2 + 4\beta\gamma\varphi(\alpha)}}\Big]\Big|_{\alpha=\frac{1}{2}}$$

$$= \frac{2}{(\frac{1}{2}-\gamma)^2}\Big[\rho(\alpha^*) - \frac{\beta\gamma}{(1-2\gamma)\sqrt{\beta^2+\beta\gamma}}\Big].$$

(5.3)

The critical values $\beta_c(\gamma)$ and $\gamma_c(\beta)$ are chosen such that $\rho(\alpha^*) > 1$ if and only if $\beta > \beta_c(\gamma)$ or $\gamma > \gamma_c(\beta)$. This implies

$$\beta_c = \beta_c(\gamma) = \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma} \quad \text{and} \quad \gamma_c = \gamma_c(\beta) = \frac{1}{2}\big(1 - \beta - \sqrt{\beta^2+2\beta}\big).$$

(5.4)

Clearly,

$$\sqrt{\beta_c^2 + \beta_c\gamma} = \frac{1}{2} - \gamma - \beta_c = \frac{1}{2}\frac{\frac{1}{2}-\gamma}{1-\gamma},$$

(5.5)

$$\sqrt{\beta^2 + \beta\gamma_c} = \frac{1}{2} - \gamma_c - \beta = \frac{1}{2}\big(\sqrt{\beta^2+2\beta} - \beta\big).$$

(5.6)

One easily checks that

$$\frac{(\frac{1}{2}-\gamma_c(\beta))^2}{1-\gamma_c(\beta)} = \beta,$$

i.e. $\beta_c(\gamma_c(\beta)) = \beta$. We write $\rho(\alpha;\beta)$ or $\rho(\alpha;\gamma)$ to emphasize the dependence on β or γ, respectively. By (5.3), (5.4) and (5.5),

$$\partial_{\alpha,\alpha}\rho(\alpha^*;\beta_c) = \frac{2}{(\frac{1}{2}-\gamma)^2}\Big[1 - \frac{\beta_c\gamma}{(1-2\gamma)\sqrt{\beta_c^2+\beta_c\gamma}}\Big] = \frac{2}{(\frac{1}{2}-\gamma)^2}\Big[1 - \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}\frac{\gamma}{(\frac{1}{2}-\gamma)\frac{\frac{1}{2}-\gamma}{1-\gamma}}\Big]$$

$$= \frac{2}{(\frac{1}{2}-\gamma)^2}[1-\gamma] = 2/\beta_c.$$

By the continuity of ρ and its derivatives in β and γ, we obtain $\sigma^2 = 2/\beta_c(\gamma)$ for the convergence $\beta \downarrow \beta_c(\gamma)$ and $\sigma^2 = 2/\beta$ for $\gamma \downarrow \gamma_c(\beta)$. Moreover, two Taylor expansions yield for $\beta \downarrow \beta_c$

$$\log\rho(\alpha^*;\beta) = \log(1 + \partial_\beta\rho(\alpha^*;\beta_c)(\beta - \beta_c) + o(\beta - \beta_c)) = \partial_\beta\rho(\alpha^*;\beta_c)(\beta - \beta_c)(1 + o(1)).$$

(5.7)

The derivative is given by

$$\partial_\beta\rho(\alpha^*;\beta_c) = \frac{1}{\frac{1}{2}-\gamma}\Big(1 + \frac{2\beta_c + \gamma}{2\sqrt{\beta_c^2+\beta_c\gamma}}\Big) = \frac{1}{\frac{1}{2}-\gamma}\Big(1 + \frac{\frac{1}{2}-\gamma+\gamma^2}{1-\gamma}}{\frac{\frac{1}{2}-\gamma}{1-\gamma}}\Big)$$

$$= \frac{1}{\frac{1}{2}-\gamma}\frac{1-2\gamma+\gamma^2}{\frac{1}{2}-\gamma} = \Big(\frac{1-\gamma}{\frac{1}{2}-\gamma}\Big)^2.$$

For the corresponding derivative with respect to γ we use (5.6) to derive

$$\partial_\gamma\rho(1/2;\gamma_c) = \frac{1}{\frac{1}{2}-\gamma_c}\rho(1/2;\gamma_c) + \frac{1}{\frac{1}{2}-\gamma_c}\frac{\beta}{2\sqrt{\beta^2+\beta\gamma_c}}$$

$$= \frac{1}{\frac{1}{2}-\gamma_c}\Big(1 + \frac{\beta}{\sqrt{\beta^2+2\beta}-\beta}\Big) = \frac{2}{\sqrt{\beta^2+2\beta}+\beta}\frac{\sqrt{\beta^2+2\beta}-\beta+\beta}{\sqrt{\beta^2+2\beta}-\beta}$$

$$= 2\frac{\sqrt{\beta^2+2\beta}}{\beta^2+2\beta-\beta^2} = \frac{\sqrt{\beta^2+2\beta}}{\beta}.$$

Repeating the argument of (5.7) for γ instead of β, Theorem 1.2 yields

$$\lim_{\beta\downarrow\beta_c(\gamma)}\sqrt{\beta-\beta_c}\log\theta(\gamma\cdot+\beta)=-\sqrt{\frac{\pi^2}{2}}\sqrt{\frac{\sigma^2}{\partial_\beta\rho(\alpha^*;\beta_c)}}\alpha^*=-\frac{\pi}{2}\sqrt{\frac{2\frac{2}{\beta_c}}{\frac{1-\gamma}{\beta_c}}\frac{1}{2}}=-\frac{\pi}{2\sqrt{1-\gamma}},$$

$$\lim_{\gamma\downarrow\gamma_c(\beta)}\sqrt{\gamma-\gamma_c}\log\theta(\gamma\cdot+\beta)=-\sqrt{\frac{\pi^2}{2}}\sqrt{\frac{\sigma^2}{\partial_\gamma\rho(\alpha^*;\gamma_c)}}\alpha^*=-\frac{\pi}{2}\sqrt{\frac{2\frac{2}{\beta}}{\frac{\sqrt{\beta^2+2\beta}}{\beta}}\frac{1}{2}}=-\frac{\pi}{2(\beta^2+2\beta)^{\frac{1}{4}}}. \qquad \square$$

A   Appendix: Proof of Mogulskii's Theorem

This section is devoted to the proof of Theorem 2.10.

Proof. The proof is an adaptation of the proof presented in [8]. We give a detailed proof of the upper bound. The changes for the lower bound are similar.

Let $N=\lfloor\frac{k_n}{r_n}\rfloor$ and $m_N=k_n$, $m_k=kr_n$ for $0\le k\le N-1$. The Markov property implies that

$$\mathbb{P}_{(0,\tau_0)}(E_n)=\mathbb{P}_{(0,\tau_0)}\Big(\bigcap_{k=1}^N\{g_1(\frac{i}{k_n})\le\frac{S_i^{(n)}}{a_n}\le g_2(\frac{i}{k_n})\,\forall\,i\in(m_{k-1},m_k]\cap\mathbb{N}\}\Big)$$

$$\le\prod_{k=2}^{N-1}\sup_{x\in[g_1(\frac{m_{k-1}}{k_n}),g_2(\frac{m_{k-1}}{k_n})]}\sup_{\tau\in\mathcal{T}}\mathbb{P}_{(0,\tau)}\Big(g_1(\frac{m_{k-1}+i}{k_n})\le\frac{S_i^{(n)}}{a_n}+x\le g_2(\frac{m_{k-1}+i}{k_n})\,\forall\,i\in[r_n]\Big).$$

Since g_1 and g_2 are bounded, for every $\delta>0$ there exists $K\in\mathbb{N}$ such that $[g_1(\frac{m_{k-1}}{k_n}),g_2(\frac{m_{k-1}}{k_n})]\subseteq[-K\delta,(K-1)\delta]=\bigcup_{j=-K}^{K-1}[j\delta,(j+1)\delta]$. Continuity of g_1 and g_2 further guarantees the existence of $A=A(\delta)>0$ such that

$$\sup_{0\le s,t\le1\,:\,|s-t|\le\frac{2}{A}}|g_1(t)-g_1(s)|+|g_2(t)-g_2(s)|<\delta. \tag{A.1}$$

Let $J_l=(\frac{(l-1)(N-2)}{A}+1,\frac{l(N-2)}{A}+1]\cap\mathbb{N}$ for $j\in[A]$. We show that for sufficiently large n, for all $i\in[r_n]$ and $l\in[A]$

$$\Big|\frac{i+m_{k-1}}{k_n}-\frac{l}{A}\Big|\le\frac{2}{A}\quad\text{for all }k\in\Big\{\frac{(l-1)(N-1)}{A}+1,\dots,\frac{l(N-2)}{A}+1\Big\}.$$

Since $\frac{k_n}{r_n}-(N-1)\in[1,2]$, we have

$$\frac{i+m_{k-1}}{k_n}\ge\frac{1+r_n\frac{(l-1)(N-1)}{A}}{k_n}\ge\frac{l-2}{A}\qquad\text{for all }l\in[A]$$

$$\Leftrightarrow A\ge k_n(l-2)-r_n(l-1)(N-1)=-k_n+(l-1)r_n\Big(\frac{k_n}{r_n}-(N-1)\Big)\quad\text{for all }l\in[A]$$

$$\Leftarrow A\ge-k_n+(A-1)r_n2.$$

Since $\lim_{n\to\infty}r_n/k_n=0$, this is satisfied for large n. For the other direction, we use that $\frac{r_n}{k_n}\le\frac{1}{N}$ to see that

$$\frac{i+m_{k-1}}{k_n}\le\frac{kr_n}{k_n}\le\Big(\frac{l(N-2)}{A}+1\Big)\frac{r_n}{k_n}\le\frac{l+2}{A}\qquad\text{for all }l\in[A]$$

$$\Leftrightarrow A\frac{r_n}{k_n}\le l+2-l(N-2)\frac{r_n}{k_n}=l\Big(1-(N-2)\frac{r_n}{k_n}\Big)+2\quad\text{for all }l\in[A]$$

$$\Leftarrow A\frac{r_n}{k_n}\le2.$$

The last inequality holds since $\frac{r_n}{k_n} \to 0$. The small k-values are needed later for the proof of the lower bound. For $k \in J_l$, (A.1) implies that for all $x \in [j\delta, (j+1)\delta]$

$$\sup_{\tau \in \mathcal{T}} \mathbb{P}_{(0,\tau)}\Big(g_1\Big(\frac{m_{k-1}+i}{k_n}\Big) \le \frac{S_i^{(n)}}{a_n} + x \le g_2\Big(\frac{m_{k-1}+i}{k_n}\Big) \forall i \in [r_n]\Big)$$

$$\le \sup_{\tau \in \mathcal{T}} \mathbb{P}_{(0,\tau)}\Big(g_1\Big(\frac{l}{A}\Big) - (j+2)\delta \le \frac{S_i^{(n)}}{a_n} \le g_2\Big(\frac{l}{A}\Big) - (j-1)\delta \, \forall i \in [r_n]\Big) =: q_{l,n}(j).$$

Since $\bigcup_{l=1}^A J_l = (1, N-1] \cap \mathbb{N}$, we derive

$$\mathbb{P}_{(0,\tau_0)}(E_n) \le \prod_{l=1}^A \Big[\max_{j \in [-K,K) \cap \mathcal{Z}} q_{l,n}(j)\Big]^{|J_l|}.$$

The assumed uniform convergence implies that for $n \to \infty$

$$q_{l,n}(j) \to P\Big(g_1\Big(\frac{l}{A}\Big) - (j+2)\delta \le \sqrt{\sigma^2 A} W_t \le g_2\Big(\frac{l}{A}\Big) - (j-1)\delta \, \forall t \in [0,1]\Big).$$

The right-hand side can be estimated by (see for example [10] p.31 or [8] Eq.(5.4))

$$\exp\Big(-\frac{\pi^2 \sigma^2}{2} \frac{(1-\delta)A}{[g_2(\frac{l}{A}) - g_1(\frac{l}{A}) + 3\delta]^2}\Big).$$

Since there are only finitely many $(j,l) \in [-K,K) \cap \mathbb{N} \times \{1,\ldots,A\}$, the convergence and the bound hold uniformly in (j,l). Moreover,

$$\#J_l \ge \frac{l(N-2)}{A} - \frac{(l-1)(N-2)}{A} - 1 = \frac{N-2}{A} - 1 \ge \frac{\frac{k_n}{r_n} - 1 - 2}{A} - 1 \ge \frac{k_n}{a_n^2 A^2} - \frac{3}{A} - 1.$$

Now we collected everything needed to derive

$$\limsup_{n \to \infty} \frac{a_n^2}{k_n} \log P_{(0,\tau_0)}(E_n) \le \limsup_{n \to \infty} \frac{a_n^2}{k_n} \sum_{l=1}^A \#J_l \log \max_{j \in [-K,K) \cap \mathcal{Z}} \tilde{q}_{l,n}(j)$$

$$\le \sum_{l=1}^A \limsup_{n \to \infty} \frac{a_n^2}{k_n} \Big(\frac{n}{a_n^2 A^2} - \frac{3}{A} - 1\Big) \log \max_{j \in [-K,K) \cap \mathcal{Z}} \tilde{q}_{l,n}(j)$$

$$\le \frac{1}{A^2} \sum_{l=1}^A -\frac{\pi^2 \sigma^2}{2} \frac{(1-\delta)A}{[g_2(\frac{l}{A}) - g_1(\frac{l}{A}) + 3\delta]^2}.$$

Since g_1 and g_2 are continuous functions, taking $A \to \infty$ yields

$$\limsup_{n \to \infty} \frac{a_n^2}{k_n} \log P_{(0,\tau_0)}(E_n) \le -\frac{\pi^2 \sigma^2}{2} \int_0^1 \frac{(1-\delta)}{[g_2(x) - g_1(x) + 3\delta]^2} \, dx.$$

Now we can take $\delta \to 0$ to establish the claim.

Sketch of the lower bound: Choose a continuous function $g : [0,1] \to \mathbb{R}$ such that $g_1(t) < g(t) < g_2(t)$ for all $t \in [0,1]$. Since it suffices to prove the lower bound for b small, we can assume that $g(1) \ge g_2(1) - b$. Then, let $\delta > 0$ be such that $g(t) - g_1(t) > 3\delta$ and $g_2(t) - g(t) > 9\delta$ for all $t \in [0,1]$. Moreover, A is chosen large enough that

$$\sup_{0 \le s \le t \le 1 : \, t-s \le \frac{2}{A}} \Big(|g_1(t) - g_1(s)| + |g(t) - g(s)| + |g_2(t) - g_2(s)|\Big) \le \delta.$$

Choose $N = \lfloor \frac{k_n}{r_n} \rfloor$, $m_N = k_n$ and $m_k = kr_n$ for $0 \le k \le N-1$. Writing $y_k = g(\frac{m_k}{k_n})$ for $1 \le k \le N$, the Markov property implies

$$P_{0,\tau_0}(E_n) \ge p_{1,n}(0,\tau_0) \times \prod_{k=2}^N \inf_{y \in [y_{k-1}, y_{k-1}+6\delta]} \inf_{\tau \in \mathcal{T}} p_{k,n}(y, \tau),$$

where for $1 \leq k \leq N$, $y \in \mathbb{R}$, and $\tau \in \mathcal{T}$

$$p_{k,n}(y, \tau) = P_{(0,\tau)}\Big(\alpha_{i,k,n} \leq \frac{S_i^{(n)}}{a_n} + y \leq \beta_{i,k,n}, \forall i \in [m_k - m_{k-1}]; y_k \leq S_{m_k - m_{k-1}}^{(n)} + 6\delta\Big)$$

$$\alpha_{i,k,n} := g_1\Big(\frac{i + m_{k-1}}{k_n}\Big) \quad \beta_{i,k,n} := g_2\Big(\frac{i + m_{k-1}}{k_n}\Big).$$

Now the choice of parameters implies that for n sufficiently large

$$P_{0,\tau_0}(E_n) \geq \min\{p_{N,n}^1, p_{N,n}^2\} \prod_{l=1}^{A} \Big(\min\{q_{l,n}^1, q_{l,n}^2\}\Big)^{k_n/[(Aa_n^2 - 1)A]},$$

where $p_{N,n}^1$ is the infimum over all $\tau \in \mathcal{T}$ of $p_{N,n}(y_{k-1}, \tau)$ where $\beta_{i,N,n}$ is replaced by $\beta_{i,N,n} - 3\delta$ and the same for $p_{N,n}^2$ with $\alpha_{i,N,n}$ replaced by $\alpha_{i,N,n} - 3\delta$ and $\beta_{i,N,n}$ by $\beta_{i,N,n} - 6\delta$. Moreover,

$$q_{l,n}^{(1)} = \inf_{\tau \in \mathcal{T}} P_{(0,\tau)}\Big(g_1\Big(\frac{l}{A}\Big) - g\Big(\frac{l}{A}\Big) + 2\delta \leq \frac{S_i^{(n)}}{a_n} \leq g_2\Big(\frac{l}{A}\Big) - g\Big(\frac{l}{A}\Big) - 5\delta \, \forall i \in [r_n], \delta \leq \frac{S_{r_n}^{(n)}}{a_n} \leq 2\delta\Big),$$

$$q_{l,n}^{(2)} = \inf_{\tau \in \mathcal{T}} P_{(0,\tau)}\Big(g_1\Big(\frac{l}{A}\Big) - g\Big(\frac{l}{A}\Big) - \delta \leq \frac{S_i^{(n)}}{a_n} \leq g_2\Big(\frac{l}{A}\Big) - g\Big(\frac{l}{A}\Big) - 8\delta \, \forall i \in [r_n], -2\delta \leq \frac{S_{r_n}^{(n)}}{a_n} \leq -\delta\Big)$$

The claim now follows with the same arguments as in [8] because of the assumed convergence of the probabilities. $\square$

References

[1] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

[2] B. Bollobás, S. Janson, and O. Riordan. The phase transition in the uniformly grown random graph has infinite order. *Random Structures Algorithms*, 26(1-2):1–36, 2005.

[3] B. Bollobás and O. Riordan. Slow emergence of the giant component in the growing m-out graph. *Random Structures Algorithms*, 27(1):1–24, 2005.

[4] R. Cohen, D. ben-Avraham, and S. Havlin. Percolation critical exponents in scale-free networks. *Physical Review E*, 66:036113, 2002.

[5] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:1222–1267, 2009.

[6] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: the giant component. *Ann. Probab.*, 41(1):329–384, 2013.

[7] R. Durrett. *Probability: Theory and Examples*. Wadsworth, Pacific Grove, CA, 1991.

[8] N. Gantert, Y. Hu, and Z. Shi. Asymptotics for the survival probability in a killed branching random walk. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):111–129, 2011.

[9] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2):742–789, 2009.

[10] K. Itô and H. P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974.

[11] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995.

[12] G. Last and M. Penrose. *Lectures on the Poisson Process.* IMS Textbook. Cambridge University Press, 2017.

[13] R. Lyons. A simple path to Biggins' martingale convergence for branching random walk. In *Classical and modern branching processes*, volume 84 of *IMA Vol. Math. Appl.*, pages 217–221. Springer, New York, 1997.

[14] A. A. Mogulskii. Small deviations in the space of trajectories. *Theory Probab. Appl.*, 19:726–736, 1974.

[15] R. G. Pinsky. *Positive harmonic functions and diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1995.

[16] O. Riordan. The small giant component in scale-free random graphs. *Combin. Probab. Comput.*, 14(5-6):897–938, 2005.

[17] M. I. Roberts. *Spine changes of measure and branching diffusions.* PhD thesis, University of Bath, 2009. Available at http://people.bath.ac.uk/mir20/research.html.