

Large deviations for the maximum of a branching random walk

Nina Gantert and Thomas Höfelsauer

April 24, 2018

Abstract

We consider real-valued branching random walks and prove a large deviation result for the position of the rightmost particle. The position of the rightmost particle is the maximum of a collection of a random number of dependent random walks. We characterise the rate function as the solution of a variational problem. We consider the same random number of independent random walks, and show that the maximum of the branching random walk is dominated by the maximum of the independent random walks. For the maximum of independent random walks, we derive a large deviation principle as well. It turns out that the rate functions for upper large deviations coincide, but in general the rate functions for lower large deviations do not.

Keywords: branching random walk, large deviations

AMS 2000 subject classification: 60F10, 60J80, 60G50.

1 Introduction

We study the maximum of a branching random walk in discrete time. The branching random walk can be described as follows. Given are two ingredients, an offspring distribution on \mathbb{N}_0 with weights $(p(k))_{k \in \mathbb{N}_0}$ and a (centred) step size distribution on \mathbb{R} . At time 0 the process starts with one particle at the origin. At every time $n \in \mathbb{N}$ each particle repeats the following two steps, independently of everything else. First, it produces offspring according to the offspring distribution $(p(k))_{k \in \mathbb{N}_0}$, and then it dies. Afterwards, the offspring particles move according to the step size distribution. Let m be the reproduction mean of the offspring distribution. We always assume $m > 1$. We are interested in the position of the rightmost particle at time n . Therefore, let D_n be the set of particles at time n . For $v \in D_n$ denote by S_v the position of particle v at time n . The maximum of the branching random walk is defined as

$$M_n = \max_{v \in D_n} S_v. \quad (1)$$

The asymptotics of the maximum M_n are by now well-understood. Let I be the rate function of the random walk. Conditioned on survival, it is well-known, see Biggins [6], Hammersley [15] and Kingman [17], that M_n grows at linear speed x^* , where

$$x^* = \sup\{x \geq 0: I(x) \leq \log m\}. \quad (2)$$

Addario-Berry and Reed [1] as well as Hu and Shi [16] obtain a logarithmic second term. In [2] Aïdékon finally proves that $M_n - x^*n - c \log n$ converges in distribution, where $c > 0$ is an explicit constant. Furthermore, let \tilde{M}_n be the maximum of $|D_n|$ independent random walks. One can check that the speed of $(\tilde{M}_n)_n$ also equals x^* , see [22, Theorem 1] or (18). However, compared to the branching random walk, there is a larger logarithmic correction term, see e.g. [22, Theorem 1]. Similar results were proved for branching Brownian motion by Bramson in [7].

We investigate the exponential decay rates of the probabilities $\mathbb{P}(\frac{M_n}{n} \geq x)$ for $x \geq x^*$ and $\mathbb{P}(\frac{M_n}{n} \leq x)$ for $x \leq x^*$. Our main result, Theorem 3.2, characterises these exponential decay rates. We consider the same question for \tilde{M}_n and determine the exponential decay rates, see Theorem 3.1. Interestingly, the rate functions coincide for $x \geq x^*$, but in general they do not coincide for $x < x^*$.

Similar questions have been studied before. Large deviation estimates for the maximum of branching Brownian motion have first been investigated by Chauvin and Rouault in [9] and very recently by Derrida and Shi in [14] and [13]. See also [12] and [21] for extensions with coalescence and selection or immigration, respectively. Note that [13] also treats continuous time branching random walks. The difference to our setup is that in the time-continuous case, the strategies can involve the exponential waiting times, while in our setup, they can involve the branching mechanism given by the offspring distribution. Upper large deviations for the maximum of discrete time branching random walks have been investigated in [20] in the case where every particle has at least one offspring. Recently large deviation results for the empirical distribution of the branching random walk have been obtained in [18], [10], [19], but they do not seem to imply our result. We also mention that in the case of a fixed number of offspring, much more precise results (describing not only the exponential decay rates) were derived in [8].

Our strategy of proof is rather direct. We compare M_n and \tilde{M}_n and show that \tilde{M}_n stochastically dominates M_n for all $n \in \mathbb{N}$, see Lemma 5.2.

Let us now introduce the model in a more formal way and fix some notation. Let $(Z_n)_{n \in \mathbb{N}_0}$ be a Galton-Watson process with one initial particle and offspring law given by the weights $(p(k))_{k \in \mathbb{N}_0}$. Let $m = \sum_{k=1}^{\infty} kp(k)$ be the reproduction mean. The associated Galton-Watson tree is denoted by $\mathcal{T} = (V, E)$, where V is the set of vertices and E is the set of edges. Further, for $n \in \mathbb{N}$ let D_n be the set of vertices in the n -th generation of the tree. Then, $|D_n| = Z_n$. For $v \in D_n$ the set of descendants of v in the $(l+n)$ -th generation is denoted by D_l^v . Note that $|D_l^v|$ equals $|D_l|$ in distribution. The root of \mathcal{T} is called $o \in V$. For $v, w \in V$ define $[v, w]$ as the set of edges along the unique path from v to w . We now define the locations of the particles. Let $(X_e)_{e \in E}$ be a collection of i.i.d. random variables, i.e. every edge of \mathcal{T} is labelled with a random variable. For $v \in D_n$ the position of the particle v at time n is defined as $S_v = \sum_{w \in [o, v]} X_w$. For $n \in \mathbb{N}$ the position of the rightmost particle at time n is $M_n = \max_{v \in D_n} S_v$ and if $D_n = \emptyset$, we set $M_n = -\infty$. We refer to $(M_n)_{n \in \mathbb{N}}$ as the maximum of the branching random walk. For $v \in D_n$, the rightmost descendant of v at time $l+n$ is defined as $M_l^v = \max_{w \in D_l^v} S_w$.

We also introduce the collection of i.i.d. random variables $(X_i^j)_{i, j \in \mathbb{N}}$, where X_1^1 has the same distribution as X_e for some $e \in E$. Moreover, for $j, n \in \mathbb{N}$ define the random walk $S_n^j = \sum_{i=1}^n X_i^j$ and the maximum of independent random walks as

$$\tilde{M}_n = \max_{1 \leq j \leq Z_n} S_n^j. \quad (3)$$

In analogy to the maximum of the branching random walk, we set $\tilde{M}_n = -\infty$, if $D_n = \emptyset$. Furthermore, for $i \in \mathbb{N}$ let X_i be an independent copy of X_1 and define $S_n = \sum_{i=1}^n X_i$. Note that for every time n the number of particles in the branching random walk equals the number of random walks considered for \tilde{M}_n . However, the positions of the particles in the branching random walk are not independent. Indeed, this dependence is such that the maximum of independent random walks stochastically dominates the maximum of the branching random walk, see Lemma 5.2. We introduce the measure

$$\mathbb{P}^*(\cdot) = \mathbb{P}(\cdot | Z_n > 0 \forall n \in \mathbb{N}). \quad (4)$$

The associated expectation is denoted by \mathbb{E}^* . Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers and let $c \in (0, \infty]$ be a constant. With a slight abuse of notation for $c = \infty$, we write $a_n = \exp(-cn + o(n))$, if $\lim_{n \rightarrow \infty} \frac{1}{n} \log a_n = -c$. Note that a_n decays faster than exponentially in n if $c = \infty$. The stochastic processes considered in this paper are discrete time processes. However, to increase the readability of the paper, we omit integer parts if no confusion arises.

The paper is organised as follows. In Section 2 we first introduce the rate function of the random walk and two rate functions concerning the Galton-Watson process. We further describe some general assumptions. Then we state our main results in Section 3. We collect some preliminary results in Section 4 and prove the main results in Section 5.

2 Rate functions and assumptions

In this section we introduce the rate functions of the random walk and the Galton-Watson process, which are needed to state our main results.

2.1 Rate function of the random walk

For $x \in \mathbb{R}$ the rate function of the random walk $(S_n)_{n \in \mathbb{N}}$ is defined as

$$I(x) = \sup_{\lambda \in \mathbb{R}} (\lambda x - \log \mathbb{E}[e^{\lambda X_1}]). \quad (5)$$

Assumption 1. *There exists $\varepsilon > 0$ such that $\mathbb{E}[e^{\lambda X_1}] < \infty$ for all $\lambda \in (-\varepsilon, \varepsilon)$. Furthermore, for simplicity suppose that $\mathbb{E}[X_1] = 0$.*

Note that $\mathbb{E}[X_1] = 0$ is not necessary for the results in Section 3, since we could derive similar results for the collection of random variables $(X_e - \mathbb{E}[X_1])_{e \in E}$. Assumption 1 ensures that $I(x) > 0$ for all $x \neq 0$ and $I(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. If Assumption 1 is satisfied, Cramér's theorem implies that the probabilities $\mathbb{P}(S_n \geq xn)$ decay exponentially in n with rate $I(x)$ for $x > 0$. A proof can e.g. be found in [11, Theorem 3.7.4].

2.2 Rate functions of the Galton-Watson process

First, we need to introduce some more assumptions before we can state the large deviation results.

Assumption 2. *The Galton-Watson process is supercritical, i.e. $m > 1$.*

Let $q = \inf\{s \in [0, 1]: \mathbb{E}[s^{Z_1}] = s\}$. Note that q is the extinction probability of the process $(Z_n)_{n \in \mathbb{N}_0}$. Assumption 2 implies that $q < 1$.

Assumption 3. *The Galton-Watson process satisfies $\mathbb{E}[Z_1 \log Z_1] < \infty$.*

First of all, note that Assumption 3 implies $m < \infty$. Together with Assumption 1, this yields that the speed of the branching random walk x^* defined in (2) is finite.

Due to the well-known Kesten-Stigum Theorem, Assumption 3 implies that the Galton-Watson process grows like its expectation, see Theorem 4.1.

Assumption 4. *Every particle has less than two offsprings with positive probability, i.e. $p(0) + p(1) > 0$.*

Assumption 4 is often referred to as Schröder case, whereas the case $p(0) + p(1) = 0$ is called Böttcher case.

If Assumptions 2 and 3 are satisfied, there is also a large deviation result for the probability that $(Z_n)_{n \in \mathbb{N}_0}$ grows at most subexponentially. A sequence $(a_n)_{n \in \mathbb{N}}$ is called subexponential, if $a_n e^{-\varepsilon n} \rightarrow 0$ as $n \rightarrow \infty$ for all $\varepsilon > 0$. Define

$$\rho := -\log \mathbb{E}[Z_1 q^{Z_1 - 1}] \in (0, \infty]. \quad (6)$$

Note that $\rho = -\log p(1)$ if $p(0) = 0$ (and therefore also $q = 0$). In particular, $\rho < \infty$ if and only if Assumption 4 is satisfied. Consider the set

$$A = \{l \in \mathbb{N}: \exists n \in \mathbb{N} \text{ such that } \mathbb{P}(Z_n = l) > 0\} \quad (7)$$

containing all positive integers l such that there are l particles at some time n with positive probability.

Theorem 2.1. *Let Assumption 2 hold. Then, for every $k \in A$ we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^*(Z_n = k) = -\rho.$$

Moreover, for every subexponential sequence $(a_n)_{n \in \mathbb{N}}$ such that $a_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^*(Z_n \leq a_n) = -\rho.$$

A proof of the first statement in Theorem 2.1 can be found in [3, Chapter 1, Section 11, Theorem 3]. The second statement is a consequence of [5, Theorem 3.1].

For $x \in [0, \log m]$ define the rate function of the Galton-Watson process as

$$I^{\text{GW}}(x) = \rho(1 - x(\log m)^{-1}) \quad (8)$$

Note that $I^{\text{GW}}(x) > 0$ for all $x < \log m$.

Theorem 2.2. *Under Assumptions 2 and 3 we have for $x \in [0, \log m]$*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^*(Z_n \leq e^{xn}) = -I^{\text{GW}}(x).$$

This theorem is a consequence of [5, Theorem 3.2]. Note that there is also an upper large deviation result for $\mathbb{P}^*(Z_n \geq e^{xn})$, where $x > \log m$, see e.g. [4, Theorem 1].

3 Results

After defining the rate functions of the random walk and the Galton-Watson process we are now able to state our main results.

Note that $I(x^*) = \log m$ if $I(x) < \infty$ for some $x > x^*$. On the other hand, $I(x^*) < \log m$ already implies $\mathbb{P}(X_1 > x^*) = 0$. This case leads to a different shape of the rate functions, see Figure 1. Let

$$k^* = \inf\{k \geq 1: p(k) > 0\}. \quad (9)$$

Note that $k = k^*$ is the smallest positive integer, such that $\mathbb{P}(Z_n = k) > 0$ for some $n \in \mathbb{N}$. Define the rate function for the maximum of independent random walks as

$$I^{\text{ind}}(x) = \begin{cases} I(x) - \log m & \text{for } x > x^*, \\ 0 & \text{for } x = x^*, \\ \rho\left(1 - \frac{I(x)}{\log m}\right) & \text{for } 0 \leq x < x^*, \\ k^*I(x) + \rho & \text{for } x \leq 0. \end{cases} \quad (10)$$

Note that $\rho\left(1 - \frac{I(x)}{\log m}\right) = I^{\text{GW}}(I(x))$ for $0 \leq x < x^*$. Recall the maximum \tilde{M}_n of a random number of independent walks, defined in (3).

Theorem 3.1. *Suppose that Assumptions 1, 2 and 3 are satisfied. Then, the laws of $\frac{\tilde{M}_n}{n}$ under \mathbb{P}^* satisfy a large deviation principle with rate function I^{ind} .*

In the Böttcher case ($p(0) + p(1) = 0$) we have $\rho = \infty$ and therefore $I^{\text{ind}}(x) = \infty$ for all $x < x^*$. Hence, in this case the lower deviation probabilities $\mathbb{P}^*(\tilde{M}_n \leq xn)$ for $x < x^*$ decay faster than exponentially in n .

Let us now give some intuition for the rate function I^{ind} and describe the large deviation event $\{\tilde{M}_n \geq xn\}$ for some $x > x^*$, respectively $\{\tilde{M}_n \leq xn\}$ for some $x < x^*$.

For $x > x^*$, the number of particles should be larger or equal than expected, i.e. $Z_n \geq e^{nt}$ for some $t \geq \log m$. The probability of such an event is of order $\exp(-I^{\text{GW}}(t)n + o(n))$. If there are e^{nt} particles at time n , the probability that at least one particle reaches xn is of order $\exp(-I(x)n + tn + o(n))$ for $t < I(x)$. Therefore, we need to maximize the product of these two probabilities, which amounts to minimize $I^{\text{GW}}(t) + I(x) - t$, where t runs over the interval $[\log m, I(x))$. It turns out that the optimal value is $t = \log m$. This argument will go through for the maximum of the branching random walk.

If $0 \leq x < x^*$, the probability that one particle reaches xn is of order $\exp(-I(x)n + o(n))$. Hence, for every $\varepsilon > 0$, if there are less than $e^{(I(x)-\varepsilon)n}$ particles, the probability that none of these particles reaches xn is close to 1. However, if there are more than $e^{(I(x)+\varepsilon)n}$ particles, this probability decays exponentially in n .

If $x < 0$, already the probability that a single particle is below xn at time n decays exponentially fast in n . Hence, if the number of particles Z_n grows exponentially, the probability that all particles are below xn at time n decays faster than exponentially. Therefore, the number of particles needs to grow subexponentially. Since ρ does not depend on the choice of k in Theorem 2.1, there have to be only k^* particles at time n (provided that $\rho < \infty$).

Next, we consider the maximum of the branching random walk. For $x < x^*$ let

$$H(x) = \inf_{t \in (0,1]} \left\{ t\rho + tI(t^{-1}(x - (1-t)x^*)) \right\}. \quad (11)$$

Note that for $x > 0$ it suffices to take the infimum over $t \in (0, 1 - \frac{x}{x^*}]$. Define the rate function of the branching random walk as

$$I^{\text{BRW}}(x) = \begin{cases} I(x) - \log m & \text{for } x > x^*, \\ 0 & \text{for } x = x^*, \\ H(x) & \text{for } x < x^*. \end{cases} \quad (12)$$

Theorem 3.2. *Suppose that Assumptions 1, 2, 3 and 4 are satisfied. Then, the laws of $\frac{M_n}{n}$ under \mathbb{P}^* satisfy a large deviation principle with rate function I^{BRW} .*

In contrast to the case of independent random walks we only consider the Schröder case (Assumption 4) for the branching random walk.

Remark. *Let us comment on the shape of the rate functions and on Assumption 4.*

- a) *One can check that the rate function I^{BRW} is convex. Furthermore, also I^{ind} is convex, if $I(x^*) = \log m$, see also the following discussion about the continuity of $I^{\text{ind}}(x)$ at $x = x^*$.*
- b) *Assumption 4 is only needed for the lower deviations ($x < x^*$) in Theorem 3.2. In the Böttcher case, i.e. if Assumption 4 is not satisfied, the strategy for lower deviations is different and will be investigated in a follow-up paper.*

For $x > x^*$ we have $I^{\text{BRW}}(x) = I^{\text{ind}}(x)$. In this case the strategy is the same as for independent random walks. The strategy in the case $x < x^*$ goes as follows. At time tn there are only k^* particles, and the position of one of those particles is smaller than its expectation. All other $k^* - 1$ particles are killed at time tn . Note that by Assumption 4 either $k^* = 1$ or particles may have no offspring with positive probability. Afterwards, every particle moves and branches according to its usual behaviour.

Further notice, that in contrast to the case of independent random walks, the number of particles can also grow exponentially if $x < 0$. It suffices to have a small number of particles at time tn for some $t \in [0, 1]$.

To compare the rate functions, note that the maximum of independent random walks stochastically dominates the maximum of the branching random walk, see Lemma 5.2. Therefore, $I^{\text{ind}}(x) \leq I^{\text{BRW}}(x)$ for $x > x^*$, respectively $I^{\text{ind}}(x) \geq I^{\text{BRW}}(x)$ for $x < x^*$. For $x < x^*$, the inequality is in general strict. For $x > x^*$, the rate functions coincide, see the argument above.

Let us now comment on the shape of the rate functions. If $I(x) = \infty$ for some $x > x^*$, also $I^{\text{ind}}(x) = I^{\text{BRW}}(x) = \infty$. More precisely, $I(x) = \infty$ already implies $\mathbb{P}(X_1 \geq x - \varepsilon) = 0$ for some $\varepsilon > 0$ and therefore $M_n \leq x^*n$, respectively $\tilde{M}_n \leq x^*n$ almost surely.

If $I(x^*) = \log m$, then the rate functions $I^{\text{ind}}(x)$ and $I^{\text{BRW}}(x)$ are continuous at $x = x^*$. However, if $I(x^*) < \log m$, the rate functions $I^{\text{ind}}(x)$ and $I^{\text{BRW}}(x)$ are infinite for $x > x^*$, since $I(x) = \infty$. Therefore, they are not continuous from the right at $x = x^*$. The rate function $I^{\text{BRW}}(x)$ is continuous from the left at $x = x^*$, since $I^{\text{BRW}}(x) \leq \rho(1 - \frac{x}{x^*})$ for $x < x^*$. However, $I^{\text{ind}}(x)$ is also not continuous from the left at $x = x^*$. In particular, $\lim_{x \nearrow x^*} I^{\text{ind}}(x) \in (0, \infty)$ if $\rho < \infty$, which implies that I^{ind} is not convex in this case.

An intuitive explanation of this discontinuity is the following. If there are at least $\exp(I(x^*)n)$ particles at time n , then $\tilde{M}_n = x^*n + o(n)$ with high probability. For a

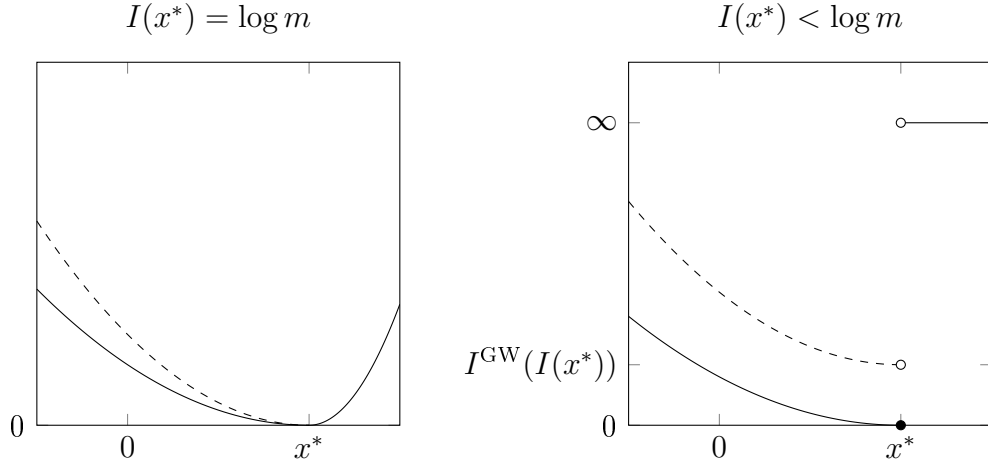


Figure 1: The figure shows the qualitative behaviour of the rate function of the branching random walk (—) and the rate function of independent random walks (---).

smaller linear term there have to be less particles, hence for all $x < x^*$ the probability $\mathbb{P}^*(\tilde{M}_n \leq xn)$ is bounded from below by the probability to have at most $\exp(I(x^*)n)$ particles at time n , which decays exponentially. Note that for the branching random walk, in contrast, it suffices to have a small number of particles at the beginning.

4 Preliminaries

Before we prove the main results, we collect some preliminaries which are needed throughout the proofs. As already mentioned in the introduction, conditioned on survival, the linear speed of the branching random walk equals x^* , i.e.

$$\lim_{n \rightarrow \infty} \frac{M_n}{n} = x^* \quad \mathbb{P}^*\text{-a.s.} \quad (13)$$

We often use the following simple inequalities.

(U1) We have $1 - x \geq e^{-ex}$ for $x \in [0, e^{-1}]$.

(U2) We have $1 - (1 - x)^y \geq xy(1 - xy)$ for $x \in [0, 1]$ and $y \geq 0$.

Proof. Both inequalities follow after some elementary calculations.

(U1) The function $f(x) = 1 - x - e^{-ex}$ is increasing on $[0, e^{-1}]$. The claim follows, as $f(0) = 0$.

(U2) The function $g(x) = 1 - e^{-x} - x(1 - x)$ is increasing for $x \geq 0$. As $g(0) = 0$, we have $1 - e^{-x} \geq x(1 - x)$. Using additionally the well known inequality $1 - x \leq e^{-x}$, we get

$$1 - (1 - x)^y \geq 1 - e^{-xy} \geq xy(1 - xy). \quad \square$$

For $i \in \mathbb{N}$ let $(a_n^i)_{n \in \mathbb{N}}$ be a sequence of positive numbers and $a^i = \limsup_{n \rightarrow \infty} \frac{1}{n} \log a_n^i$. Then, for all $k \in \mathbb{N}$ it holds that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{i=1}^k a_n^i = \max_{i \in \{1, \dots, k\}} a^i. \quad (14)$$

A proof can e.g. be found in [11, Lemma 1.2.15].

We often need to estimate the number of particles at time n , which has expectation m^n . Let $W_n = \frac{Z_n}{m^n}$ and $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be the natural filtration of the Galton-Watson process, i.e. $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$. The process $(W_n)_{n \in \mathbb{N}}$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$. Therefore, $W_n \rightarrow W$ almost surely, where W is an almost surely finite random variable. The following well-known theorem shows that under our assumptions, the limit W is non-trivial, i.e. $\mathbb{P}(W = 0) < 1$.

Theorem 4.1 (Kesten-Stigum). *If Assumption 2 and Assumption 3 are satisfied, we have*

$$\mathbb{E}[W] = 1 \quad \text{and} \quad \mathbb{P}(W = 0) = q < 1.$$

A proof can e.g. be found in [3, Chapter 1, Section 10, Theorem 1].

5 Proofs

In this section we prove the main results of the paper.

Independent random walks

Proof of Theorem 3.1. 1. Case: $x > x^$*

Following the strategy explained in Section 3, independence of the random walks and (U2) yields

$$\begin{aligned} \mathbb{P}^* \left(\frac{\tilde{M}_n}{n} \geq x \right) &= \mathbb{E}^* \left[1 - \left(1 - \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \right)^{Z_n} \right] \\ &\geq \mathbb{P}^* \left(Z_n \geq \frac{1}{2} e^{(\log m)n} \right) \cdot \mathbb{E} \left[1 - \left(1 - \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \right)^{\frac{1}{2} e^{(\log m)n}} \right] \\ &\geq \mathbb{P} \left(W_n \geq \frac{1}{2} \right) \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \frac{1}{2} e^{(\log m)n} \left(1 - \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \right)^{\frac{1}{2} e^{(\log m)n}}. \end{aligned} \quad (15)$$

By Cramér's theorem, $\mathbb{P} \left(\frac{S_n}{n} \geq x \right) \frac{1}{2} e^{(\log m)n} \rightarrow 0$ as $n \rightarrow \infty$, since $\log m < I(x)$. For the first factor on the right hand side of (15) we have $\liminf_{n \rightarrow \infty} \mathbb{P} \left(W_n \geq \frac{1}{2} \right) \geq \mathbb{P} \left(W > \frac{1}{2} \right) > 0$, since $\mathbb{E}[W] = 1$ by Theorem 4.1. Together with Theorem 2.2 and (14) we conclude

$$\mathbb{P}^* \left(\frac{\tilde{M}_n}{n} \geq x \right) \geq \exp(-I(x) - \log m)n + o(n),$$

which yields the lower bound. For the upper bound, the Markov inequality yields

$$\mathbb{P}^* \left(\frac{\tilde{M}_n}{n} \geq x \right) = \mathbb{P}^* \left(\sum_{i=1}^{Z_n} \mathbb{1}_{\{S_n^i \geq nx\}} \geq 1 \right) \leq \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \mathbb{E}^*[Z_n] = \mathbb{P} \left(\frac{S_n}{n} \geq x \right) \frac{m^n}{1 - q}, \quad (16)$$

which immediately implies the claim.

2. Case: $0 < x < x^*$

Since the rate function I is strictly increasing on the interval $[0, x^*]$, we can choose $\varepsilon > 0$ such that $\varepsilon < I(x) < \log m - \varepsilon$. We prove the upper bound first. Using the inequality $1 - y \leq e^{-y}$ and Theorem 2.2, we have for n large enough

$$\begin{aligned} \mathbb{P}^*\left(\frac{\tilde{M}_n}{n} \leq x\right) &= \mathbb{E}^*\left[\left(1 - \mathbb{P}\left(\frac{S_n}{n} > x\right)\right)^{Z_n}\right] \leq \mathbb{E}^*\left[\exp\left(-\mathbb{P}\left(\frac{S_n}{n} > x\right)Z_n\right)\right] \\ &\leq \mathbb{P}^*\left(Z_n \leq e^{(I(x)+\varepsilon)n}\right) + \exp\left(-e^{\frac{\varepsilon}{2}n}\right) = \exp\left(-I^{\text{GW}}(I(x) + \varepsilon)n + o(n)\right). \end{aligned}$$

Letting $\varepsilon \rightarrow 0$ yields the upper bound. Note that I^{GW} defined in (8) is continuous. The proof for the lower bound is similar. More precisely, since $\mathbb{P}\left(\frac{S_n}{n} > x\right) < e^{-1}$ for n large enough, (U1) yields for n large enough

$$\begin{aligned} \mathbb{P}^*\left(\frac{\tilde{M}_n}{n} \leq x\right) &= \mathbb{E}^*\left[\left(1 - \mathbb{P}\left(\frac{S_n}{n} > x\right)\right)^{Z_n}\right] \geq \mathbb{E}^*\left[\exp\left(-e \cdot \mathbb{P}\left(\frac{S_n}{n} > x\right)Z_n\right)\right] \\ &\geq \mathbb{P}^*\left(Z_n \leq e^{(I(x)-\varepsilon)n}\right) \cdot \exp\left(-e^{1-\frac{\varepsilon}{2}n}\right) = \exp\left(-I^{\text{GW}}(I(x) - \varepsilon)n + o(n)\right). \end{aligned}$$

Letting $\varepsilon \rightarrow 0$ yields the lower bound.

3. Case: $x \leq 0$

We first consider $x < 0$. For the upper bound we have for $K \in \mathbb{N}$

$$\mathbb{P}^*\left(\frac{\tilde{M}_n}{n} \leq x\right) = \mathbb{E}^*\left[\mathbb{P}\left(\frac{S_n}{n} \leq x\right)^{Z_n}\right] \leq \sum_{k=1}^K \mathbb{P}\left(\frac{S_n}{n} \leq x\right)^k \mathbb{P}^*(Z_n = k) + \mathbb{P}\left(\frac{S_n}{n} \leq x\right)^K. \quad (17)$$

By Theorem 2.1, the probability $\mathbb{P}(Z_n = k)$ is of order $\exp(-\rho n + o(n))$ for all $k \in A$ (defined in (7)) and $\mathbb{P}(Z_n = k) = 0$ otherwise. For all $K \in \mathbb{N}$, (14) yields

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^*\left(\frac{\tilde{M}_n}{n} \leq x\right) \leq \max\{-(k^*I(x) + \rho), -KI(x)\}.$$

Hence, letting $K \rightarrow \infty$ proves the upper bound. Note that $I(x) > 0$ for $x < 0$. As in the proof of (17) we have

$$\begin{aligned} \mathbb{P}^*\left(\frac{\tilde{M}_n}{n} \leq x\right) &= \mathbb{E}^*\left[\mathbb{P}\left(\frac{S_n}{n} \leq x\right)^{Z_n}\right] \geq \mathbb{P}\left(\frac{S_n}{n} \leq x\right)^{k^*} \cdot \mathbb{P}(Z_n = k^*) \\ &= \exp\left(-(k^*I(x) + \rho)n + o(n)\right), \end{aligned}$$

which shows the lower bound. For $x = 0$ the result follows from continuity of the rate function I^{ind} at 0.

4. Case: $x = x^*$

Since $\mathbb{P}^*(\tilde{M}_n \geq xn)$ decays exponentially in n for $x > x^*$ and $\mathbb{P}^*(\tilde{M}_n \leq xn)$ decays exponentially in n for $x < x^*$, we conclude

$$\lim_{n \rightarrow \infty} \frac{\tilde{M}_n}{n} = x^* \quad \mathbb{P}^*\text{-a.s.} \quad (18)$$

and in particular $I^{\text{ind}}(x^*) = 0$. □

Branching random walk

Before proving Theorem 3.2, we first show that the maximum of independent random walks stochastically dominates the maximum of the branching random walk.

Lemma 5.1. *Let $(X_i)_{i \in \mathbb{N}}$ and $(Y_i)_{i \in \mathbb{N}}$ be independent sequences of (not necessarily independent) random variables. Furthermore, assume that the random variables $Y_i, i \in \mathbb{N}$, have the same distribution. Then we have for all $k \in \mathbb{N}$ and $x \in \mathbb{R}$*

$$\mathbb{P}\left(\max_{i \in \{1, \dots, k\}} \{X_i + Y_1\} \leq x\right) \geq \mathbb{P}\left(\max_{i \in \{1, \dots, k\}} \{X_i + Y_i\} \leq x\right).$$

Proof. Let i^* be the smallest (random) index such that $X_{i^*} = \max_{i \in \{1, \dots, k\}} X_i$. We have

$$\begin{aligned} \mathbb{P}\left(\max_{i \in \{1, \dots, k\}} \{X_i + Y_i\} \leq x\right) &\leq \mathbb{P}(X_{i^*} + Y_{i^*} \leq x) \\ &= \mathbb{P}(X_{i^*} + Y_1 \leq x). \end{aligned} \quad \square$$

As a consequence we can show that the maximum of independent random walks stochastically dominates the maximum of the branching random walk.

Lemma 5.2. *For all $n \in \mathbb{N}$ and $x \in \mathbb{R}$*

$$\mathbb{P}(M_n \leq x) \geq \mathbb{P}(\tilde{M}_n \leq x).$$

Proof. We prove this lemma by induction over n . For $n = 1$ the inequality is obviously true. Now assume that the inequality holds for some $n \in \mathbb{N}$. Let $(S_n^{i,1})_{n \in \mathbb{N}}, (S_n^{i,2})_{n \in \mathbb{N}}, \dots$ be independent copies of $(S_n^i)_{n \in \mathbb{N}}$ and define $(\tilde{M}_n^1)_{n \in \mathbb{N}}, (\tilde{M}_n^2)_{n \in \mathbb{N}}, \dots$ and $(M_n^1)_{n \in \mathbb{N}}, (M_n^2)_{n \in \mathbb{N}}, \dots$ in the same way. Furthermore, for $i \in \{1, \dots, Z_1\}$, denote by Z_n^i the number of descendants of the i -th particle of the first generation at time $n + 1$. Note that Z_n^i equals Z_n in distribution. Using the induction hypothesis and Lemma 5.1,

$$\begin{aligned} \mathbb{P}(M_{n+1} \leq x) &= \mathbb{P}\left(\max_{i \in \{1, \dots, Z_1\}} \{X_1^i + M_n^i\} \leq x\right) \\ &\geq \mathbb{P}\left(\max_{i \in \{1, \dots, Z_1\}} \{X_1^i + \tilde{M}_n^i\} \leq x\right) \\ &\geq \mathbb{P}\left(\max_{i \in \{1, \dots, Z_1\}} \max_{j \in \{1, \dots, Z_n^i\}} \{X_1^{i,j} + S_{n+1}^{i,j} - X_1^{i,j}\} \leq x\right) \\ &= \mathbb{P}(\tilde{M}_{n+1} \leq x). \end{aligned} \quad \square$$

The statement of Lemma 5.2 is also true with respect to \mathbb{P}^* .

Proof of Theorem 3.2. 1. Case: $x > x^$*

Recall that $I^{\text{BRW}}(x) = I^{\text{ind}}(x)$ for $x \geq x^*$. Therefore, the upper bound immediately follows from Theorem 3.1 and Lemma 5.2. It remains to prove the lower bound. Recall that for $v \in D_{\varepsilon n}$ the rightmost descendant of v at time n is denoted by $M_{(1-\varepsilon)n}^v$. By Lemma 5.1,

$$\begin{aligned} \mathbb{P}^*\left(\frac{M_n}{n} \geq x\right) &= \mathbb{P}^*\left(\max_{v \in D_{\varepsilon n}} \frac{M_{(1-\varepsilon)n}^v - S_v}{(1-\varepsilon)n} + \frac{S_v}{(1-\varepsilon)n} \geq \frac{x}{1-\varepsilon}\right) \\ &\geq \mathbb{P}^*\left(\max_{v \in D_{\varepsilon n}} \frac{M_{(1-\varepsilon)n}^v - S_v}{(1-\varepsilon)n} + \frac{S_{\varepsilon n}}{(1-\varepsilon)n} \geq \frac{x}{1-\varepsilon}\right) \\ &\geq \mathbb{P}^*\left(\max_{v \in D_{\varepsilon n}} \frac{M_{(1-\varepsilon)n}^v - S_v}{(1-\varepsilon)n} \geq x\right) \cdot \mathbb{P}\left(\frac{S_{\varepsilon n}}{\varepsilon n} \geq x\right). \end{aligned} \quad (19)$$

It remains to estimate the first probability on the right hand side of (19). Note that $(M_{(1-\varepsilon)n}^v - S_v)_{v \in D_{\varepsilon n}}$ are independent under \mathbb{P} . Therefore, we can use similar estimates as in the proof of Theorem 3.1. More precisely, by independence and (U2) we get

$$\begin{aligned} & \mathbb{P}^* \left(\max_{v \in D_{\varepsilon n}} \frac{M_{(1-\varepsilon)n}^v - S_v}{(1-\varepsilon)n} \geq x \right) \\ & \geq \mathbb{P}^* \left(Z_{\varepsilon n} \geq \frac{1}{2} m^{\varepsilon n} \right) \cdot \left(1 - \left(1 - \mathbb{P} \left(\frac{M_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right) \right)^{\frac{1}{2} m^{\varepsilon n}} \right) \\ & \geq \mathbb{P} \left(W_{\varepsilon n} \geq \frac{1}{2} \right) (1-q) \mathbb{P}^* \left(\frac{M_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right) \frac{1}{2} m^{\varepsilon n} \left(1 - \mathbb{P} \left(\frac{M_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right) \frac{1}{2} m^{\varepsilon n} \right). \end{aligned} \quad (20)$$

Analogously to (15), for the first factor on the right hand side of (20) it holds that $\liminf_{n \rightarrow \infty} \mathbb{P}(W_{\varepsilon n} \geq \frac{1}{2}) \geq \mathbb{P}(W > \frac{1}{2}) > 0$. Furthermore, analogously to (16), the Markov inequality and the choice of ε yields

$$\begin{aligned} 1 - \mathbb{P} \left(\frac{M_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right) m^{\varepsilon n} & \geq 1 - \mathbb{P} \left(\frac{S_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right) m^n \\ & \geq 1 - \exp \left(-n \left((1-\varepsilon)I(x) - \log m \right) + o(n) \right) \rightarrow 1. \end{aligned} \quad (21)$$

Combining (19), (20) and (21) shows

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^* \left(\frac{M_n}{n} \geq x \right) \geq -\varepsilon(I(x) - \log m) + (1-\varepsilon) \liminf_{n \rightarrow \infty} \frac{1}{(1-\varepsilon)n} \log \mathbb{P}^* \left(\frac{M_{(1-\varepsilon)n}}{(1-\varepsilon)n} \geq x \right).$$

This implies the lower bound.

2. Case: $x < x^*$

Following the strategy explained in Section 3, there are only k^* particles at time tn and the position of one particle is smaller than its expectation. Afterwards, all particles move and branch as usual. For the lower bound let $t \in (0, \min\{1 - \frac{x}{x^*}, 1\}]$ and fix $\varepsilon > 0$. Note that $t \in (0, 1 - \frac{x}{x^*})$ if $x > 0$ and $t \in (0, 1]$ if $x \leq 0$. We have

$$\begin{aligned} \mathbb{P}^* \left(\frac{M_n}{n} \leq x \right) & \geq \mathbb{P}^* \left(\frac{M_n}{n} \leq x \mid Z_{tn} = k^* \right) \cdot \mathbb{P}^*(Z_{tn} = k^*) \\ & \geq q^{k^*-1} \mathbb{P}^* \left(\frac{S_{tn} + M_{(1-t)n}}{n} \leq x \right) \cdot \mathbb{P}^*(Z_{tn} = k^*) \\ & \geq \mathbb{P}^* \left(\frac{M_{(1-t)n}}{(1-t)n} \leq x^* + \varepsilon \right) \cdot \mathbb{P} \left(\frac{S_{tn}}{n} \leq (x - (1-t)(x^* + \varepsilon)) \right) \cdot \mathbb{P}^*(Z_{tn} = k^*). \end{aligned} \quad (22)$$

Since the first probability on the right hand side of (22) converges to 1 almost surely as $n \rightarrow \infty$ by (13), we get

$$\mathbb{P}^* \left(\frac{M_n}{n} \leq x \right) \geq \exp \left(- \left[I \left(t^{-1} (x - (1-t)(x^* + \varepsilon)) \right) + \rho \right] tn + o(n) \right).$$

Letting $\varepsilon \rightarrow 0$ and since this inequality holds for all $t \in (0, \min\{1 - \frac{x}{x^*}, 1\}]$, we conclude

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^* \left(\frac{M_n}{n} \leq x \right) \geq \sup_{t \in (0, \min\{1 - \frac{x}{x^*}, 1\}]} -H(x) = - \inf_{t \in (0, 1]} H(x).$$

For the upper bound define

$$T_n = \inf \left\{ t \geq 0 : Z_{tn} \geq n^3 \right\}$$

and for $\varepsilon_1 > 0$ introduce the set

$$F = F(\varepsilon_1) = \left\{ \varepsilon_1, 2\varepsilon_1, \dots, \left\lceil \min \left\{ \left(1 - \frac{x}{x^*}\right), 1 \right\} \varepsilon_1^{-1} \right\rceil \varepsilon_1 \right\}.$$

By the definition of T_n we then have

$$\begin{aligned} & \mathbb{P}^* \left(\frac{M_n}{n} \leq x \right) \\ & \leq \mathbb{P}^* \left(T_n > 1 - \frac{x}{x^*} \right) + \sum_{t \in F} \mathbb{P}^* \left(\frac{M_n}{n} \leq x \mid T_n \in (t - \varepsilon_1, t] \right) \cdot \mathbb{P}^* (T_n \in (t - \varepsilon_1, t]) \\ & \leq \mathbb{P}^* (Z_{(1-\frac{x}{x^*})n} \leq n^3) + \sum_{t \in F} \mathbb{P}^* \left(\frac{M_n}{n} \leq x \mid T_n \in (t - \varepsilon_1, t] \right) \cdot \mathbb{P}^* (Z_{(t-\varepsilon_1)n} \leq n^3). \end{aligned} \quad (23)$$

Let $\varepsilon_2 > 0$. Using Lemma 5.1 again,

$$\begin{aligned} & \mathbb{P}^* \left(\frac{M_n}{n} \leq x \mid T_n \in (t - \varepsilon_1, t] \right) \\ & \leq \mathbb{P}^* \left(\max_{v \in D_{tn}} \frac{S_{tn} + M_{(1-t)n}^v}{n} \leq x \mid T_n \in (t - \varepsilon_1, t] \right) \\ & \leq \mathbb{P} \left(\frac{S_{tn}}{n} \leq -((1-t)(x^* - \varepsilon_2) - x) \right) + \mathbb{P} \left(\frac{M_{(1-t)n}}{(1-t)n} \leq x^* - \varepsilon_2 \right)^{n^2} \\ & \quad + \mathbb{P}^* (Z_{tn} \leq n^2 \mid T_n \in (t - \varepsilon_1, t]). \end{aligned} \quad (24)$$

Note that we ignored all but one particle at time tn . Further note that the second probability on the right hand side of (24) converges to q almost surely by (13) and the definition of M_n on the event of extinction. Since $q < 1$, the second term in (24) decays faster than exponentially in n . For the third term on the right hand side of (24),

$$\begin{aligned} \mathbb{P}^* (Z_{tn} \leq n^2 \mid T_n \in (t - \varepsilon_1, t]) & \leq \mathbb{P}^* (\exists k \in \mathbb{N} : Z_k \leq n^2 \mid Z_0 = n^3) \\ & \leq \binom{n^3}{n^2} q^{n^3 - n^2} \leq \exp((n^3 - n^2) \log q + 3n^2 \log n). \end{aligned} \quad (25)$$

In the second inequality we used the fact that for the event we consider, at most n^2 of the initial n^3 Galton-Watson trees may survive. Note that every initial particle produces an independent Galton-Watson tree.

Combining (23), (24) and (25) and letting $\varepsilon_1, \varepsilon_2 \rightarrow 0$, we conclude with (14) after a straightforward calculation

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}^* \left(\frac{M_n}{n} \leq x \right) \leq - \inf_{t \in (0, \min\{1-\frac{x}{x^*}, 1\}]} \left\{ t\rho + tI \left(-\frac{(1-t)x^* - x}{t} \right) \right\} = -H(x).$$

Note that we could take the limit $\varepsilon_2 \rightarrow 0$, since I is continuous from the right on $(0, \infty)$.

3. Case: $x = x^*$

This follows from (13). □

Acknowledgments We thank the referee for reading carefully the first version of this article, and for pointing out several inaccuracies.

References

- [1] L. Addario-Berry and B. Reed, *Minima in branching random walks*, Ann. Probab. **37** (2009), no. 3, 1044–1079.
- [2] E. Aïdékon, *Convergence in law of the minimum of a branching random walk*, Ann. Probab. **41** (2013), no. 3A, 1362–1426.
- [3] K. B. Athreya and P. E. Ney, *Branching processes*, Dover Publications, Inc., Mineola, NY, 2004.
- [4] V. Bansaye and C. Böinghoff, *Upper large deviations for branching processes in random environment with heavy tails*, Electron. J. Probab. **16** (2011), no. 69, 1900–1933.
- [5] ———, *Lower large deviations for supercritical branching processes in random environment*, Proc. Steklov Inst. Math **282** (2013), no. 1, 15–34.
- [6] J. D. Biggins, *The first- and last-birth problems for a multitype age-dependent branching process*, Advances in Appl. Probability **8** (1976), no. 3, 446–459.
- [7] M. D. Bramson, *Maximal displacement of branching Brownian motion*, Comm. Pure Appl. Math. **31** (1978), no. 5, 531–581.
- [8] D. Buraczewski and M. Maslanka, *Large deviation estimates for branching random walks*, (2017).
- [9] B. Chauvin and A. Rouault, *KPP equation and supercritical branching Brownian motion in the subcritical speed area. application to spatial trees*, Probab. Theory Related Fields **80** (1988), no. 2, 299–314.
- [10] X. Chen and H. He, *On large deviation probabilities for empirical distribution of branching random walks: Schröder case and Böttcher case*, (2017).
- [11] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010.
- [12] B. Derrida and Z. Shi, *Large deviations for the branching Brownian motion in presence of selection or coalescence*, J. Stat. Phys. **163** (2016), no. 6, 1285–1311.
- [13] ———, *Large deviations for the rightmost position in a branching Brownian motion*, (2017).
- [14] ———, *Slower deviations of the branching Brownian motion and of branching random walks*, J. Phys. A **50** (2017), no. 34, 344001, 13.
- [15] J. M. Hammersley, *Postulates for subadditive processes*, Ann. Probability **2** (1974), 652–680.

- [16] Y. Hu and Z. Shi, *Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees*, Ann. Probab. **37** (2009), no. 2, 742–789.
- [17] J. F. C. Kingman, *The first birth problem for an age-dependent branching process*, Ann. Probability **3** (1975), no. 5, 790–801.
- [18] O. Luidor and W. Perkins, *Large deviations for the empirical distribution in the branching random walk*, Electron. J. Probab. **20** (2015), no. 18, 19.
- [19] O. Luidor and E. Tsairi, *Large Deviations for the Empirical Distribution in the General Branching Random Walk*, (2017).
- [20] A. Rouault, *Precise estimates of presence probabilities in the branching random walk*, Stochastic Process. Appl. **44** (1993), no. 1, 27–39.
- [21] H. Sun, *A large deviation theorem for a branching Brownian motion with random immigration*, Statist. Probab. Lett. **83** (2013), no. 6, 1559–1566.
- [22] O. Zeitouni, *Branching random walks and Gaussian fields*, Probability and statistical physics in St. Petersburg, Proc. Sympos. Pure Math., vol. 91, Amer. Math. Soc., Providence, RI, 2016, pp. 437–471.

N. Gantert, FAKULTÄT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN, BOLTZMANNSTR. 3,
85748 GARCHING, GERMANY
E-mail address: gantert@ma.tum.de

T. Höfelsauer, FAKULTÄT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN, BOLTZMANNSTR. 3,
85748 GARCHING, GERMANY
E-mail address: thomas.hoefelsauer@tum.de