Journal of Statistical Physics (2022) 188:27
https://doi.org/10.1007/s10955-022-02957-9

The Tail of the Length of an Excursion in a Trap of Random Size

Nina Gantert[1] **· Achim Klenke**[2]

Received: 4 November 2021 / Accepted: 20 June 2022 / Published online: 6 July 2022

Abstract

Consider a random walk with a drift to the right on $\{0, \ldots, k\}$ where k is random and geometrically distributed. We show that the tail $\mathbb{P}[T > t]$ of the length T of an excursion from 0 decreases up to constants like $t^{-\varrho}$ for some $\varrho > 0$ but is not regularly varying. We compute the oscillations of $t^{\varrho}\,\mathbb{P}[T > t]$ as $t \to \infty$ explicitly.

Keywords Trapping phenomena · Tails of hitting times · Excursions of random walks · Tail of the population size in a branching process in random environment

1 Introduction and Main Result

1.1 Introduction

In this paper, we study a simple object: the tail of the time a biased random walk spends in a trap of random size. Our result is very explicit and may serve as a building block in the study of trapping models. Trapping phenomena for biased walks have been investigated intensively over the last decade, we refer to [4] for a survey. As a model for transport in an inhomogeneous medium, one can study biased random walk on a supercritical percolation cluster on $\mathbb{Z}^d$ for $d \geq 2$. It turns out that for small values of the bias, the walk moves at a positive linear speed, whereas for large values of the bias, the speed vanishes. The critical value of the bias separating the two regimes is the value where the expectation of the time spent in a trap changes from being finite to being infinite. This model goes back to [3] and was investigated in [7] and [22]. Finally, Alexander Fribergh and Alan Hammond proved a sharp

---

Communicated by Li-Cheng Tsai.

✉ Nina Gantert
gantert@ma.tum.de

Achim Klenke
math@aklenke.de

1 Fakultät für Mathematik, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany

2 Institut für Mathematik, Johannes Gutenberg-Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany

transition for the positivity of the speed in [12]. Concerning limit laws for the distribution of the walker, a central limit theorem for small bias was proved in [22]. The law of the walker in the subballistic case was addressed by [12]: the authors find the polynomial order of the distance of the walker to the origin. It is conjectured that it depends on the spatial direction of the bias if there is a limit law for the distance of the walker to the origin.

Replacing the integer lattice with a tree yields a biased random walk on a supercritical Galton–Watson tree. In this case, the phase transition for the bias is easier to understand and was shown in [20]. It turns out that the distance of the walker to the origin does not satisfy a limit law, but there are subsequences converging to certain infinitely divisible laws, see [5]. The crucial object is the time T spent in traps (averaged over the size of the trap): since the tail of this random variable is not in the domain of attraction of a stable law, there is no limit law for the time the walker needs to go at a certain distance of the origin. We refer to the introduction of [5] for more explanations. If one randomizes the bias, the situation changes, see [6] and [16]. For one-dimensional random walk in random environment, limit laws for the distance of the walker to the origin have been proved in [17] under a non-lattice assumption. If the non-lattice assumption is violated, one would expect convergence of subsequences as for the aforementioned biased random walk on a Galton–Watson tree. The result of this paper can be used to confirm this in the simple case of an environment which has either a drift to the left or a reflection to the right, treated in [21] and [13].

As a toy model for the supercritical percolation cluster, one may consider a percolation on a ladder graph, conditioned to survive. This model was introduced in [2] and further investigated by [14, 15, 19]. Again, our result may be applied to show that there is no limit law for the distance of the walker to the origin, as conjectured in [19].

There is a well-known connection between hitting times of a random walk (or random walk in random environment) and the total population size in a branching process (or branching process in random environment) with geometric offspring laws. For subcritical branching processes in random environment (BPRE), a precise asymptotics for the tail of the total population size under a non-lattice assumption was given in [1]. See also [10] for an upper bound on the same tail without non-lattice assumption. Again, our result can serve as an example that the precise asymptotics fails in the lattice case, at least in a particular case of a degenerate environment. More precisely, consider a subcritical BPRE where in each generation the law of the offspring is either geometric with expectation > 1 or the Dirac measure at 0. Denote by T the total population size in this BPRE. Then, while the probability $\mathbb{P}[T > t]$ satisfies, for positive constants c_1 and c_2 and a certain exponent ϱ,

$$c_1 t^{-\varrho} \leq \mathbb{P}[T > t] \leq c_2 t^{-\varrho}, \tag{1.1}$$

it is not regularly varying. More precisely, we show that $\mathbb{P}[T > t]t^{\varrho}$ is asymptotically equivalent to a nonconstant, multiplicatively periodic function, see (1.11). In our setup, with T denoting the time spent in a trap of random size, (1.1) was proved in [19] and it was conjectured that the tail is not regularly varying. This is confirmed by our result. Similar tail asymptotics for various quantities are known in the context of branching processes, see for instance [8, 9, 23].

1.2 Main Result

Let us now give precise definitions and state our main result, Theorem 1.1. Let $\beta > 1$ be a fixed parameter. Let $k \in \mathbb{N}_0$ and consider discrete time random walk X on $\{0, \dots, k\}$ with edge weight $C(l, l+1) = \beta^l$ along the edge $\{l, l+1\}$ and started in $X_0 = 0$. That is, if

$X_n = l \in \{1, \dots, k-1\}$ then it jumps to $l+1$ with probability $\beta/(1+\beta)$ and to $l-1$ with probability $1/(1+\beta)$. There is reflection at the boundaries: If $X_n = 0$, then it jumps to 1. If $X_n = k$, then it jumps to $k-1$. Of course, for $k = 0$, the random walk is trivial. Let $\mathbf{P}_k$ denote the probabilities with respect to fixed k and let $\mathbb{P}$ denote the probabilities with respect to a random geometrically distributed k with parameter $1-\alpha$, that is,

$$\mathbb{P} = \sum_{k=0}^{\infty} (1-\alpha)\alpha^k \, \mathbf{P}_k. \tag{1.2}$$

Also let $\mathbf{E}_k$ and $\mathbb{E}$ be the corresponding expectations, respectively. Here $\alpha \in (0, 1)$ is a fixed parameter. Let

$$T := \inf\left\{t > 0 : \ X_t = 0\right\} \quad \text{if } \ k \geq 1 \tag{1.3}$$

and $T = 0$ if $k = 0$, be the length of an excursion from 0. Let

$$\varrho := -\frac{\log(\alpha)}{\log(\beta)}. \tag{1.4}$$

Our random walk X is a special case of a random walk in an irreducible electrical network, see, e.g., [18, Chap. 19], on a finite graph (V, E) with edge weights $C(e)$, $e \in E$. Denote by $C(x)$ the sum of $C(e)$ for all edges incident to the vertex $x \in V$, and let $C := \sum_x C(x)$. The transition probabilities are given by $p(x, y) = C(\{x, y\})/C(x)$, $x, y \in V$. It is easy to check that $\pi(x) := C(x)/C$ defines the unique invariant measure. By [18, Theorem 17.52], the expected time to return to x (when started in x) equals $1/\pi(x) = C/C(x)$.

We use this fact to compute, for fixed k the expectation of T:

$$\mathbf{E}_k[T] = \frac{2}{C(0,1)} \sum_{l=0}^{k-1} C(l, l+1) = 2 \sum_{l=0}^{k-1} \beta^l = 2\frac{\beta^k - 1}{\beta - 1}. \tag{1.5}$$

Hence

$$\mathbb{E}[T] = (1-\alpha) \sum_{k=1}^{\infty} \alpha^k \, 2 \, \frac{\beta^k - 1}{\beta - 1} = \begin{cases} \frac{2\alpha}{1-\alpha\beta} < \infty, & \text{if } \varrho > 1, \\ \infty, & \text{if } \varrho \leq 1. \end{cases} \tag{1.6}$$

A similar but more involved computation shows that

$$\mathbb{E}\left[T^2\right] < \infty \text{ if and only if } \varrho > 2. \tag{1.7}$$

In order to describe the tail of T, we introduce the function g defined by

$$g(t) := \frac{\beta - 1}{\beta} \frac{(1-\alpha)\Gamma(\varrho)}{\log(\beta)} \left(\frac{2\beta}{(\beta-1)^2}\right)^{\varrho} \left[1 + \sum_{\ell=1}^{\infty} c_\ell \cos\left(2\pi\ell \frac{\log(t)}{\log(\beta)} - d_\ell\right)\right] \tag{1.8}$$

with

$$c_\ell = 2\frac{\left|\Gamma\left(\varrho + \frac{2\pi i \, \ell}{\log(\beta)}\right)\right|}{\Gamma(\varrho)} \quad \text{and} \quad d_\ell = \arg\left(\Gamma\left(\varrho + \frac{2\pi i \, \ell}{\log(\beta)}\right)\right). \tag{1.9}$$

Here, Γ is Euler's Gamma function and $\arg(a + bi) \in (-\pi/2, \pi/2)$ denotes the angle of $a + bi$ for $a > 0$ and $b \in \mathbb{R}$. Note that the c_ℓ decrease quickly with ℓ and hence the constant and the $\ell = 1$ mode are dominant.

Note that g is a nonconstant multiplicatively periodic function, that is

$$g(\beta t) = g(t) \quad \text{for all } t > 0. \tag{1.10}$$

In particular, g is not slowly varying.

Theorem 1.1 *For g defined in (1.8), we have*

$$\lim_{t\to\infty} \frac{t^{\varrho}\,\mathbb{P}[T > t]}{g\left(\frac{(\beta-1)^2}{2\beta}t\right)} = 1\,. \tag{1.11}$$

1.3 Outline

The strategy of the proof is as follows: We first consider the event A where X reaches k before returning to 0. On the complement of this event, T is very small and hence this case can be neglected for the tail of T (Lemma 2.3). On the event A, we split the time T into three parts:

(1) the time T_{in} needed to reach k,
(2) the time T_{exc} spent in excursions from k to k that do not reach 0, and
(3) the length T_{out} of the last excursion from k to 0.

We will show that the contributions from (1) and (3) can be neglected (Lemmas 2.4 and 2.5). Finally, we consider (2). The number of excursions is geometrically distributed and the length of the single excursion has exponential moments. We infer that the tail of T is governed by the number of excursions multiplied by their expected lengths (Proposition 2.17). The number of excursions is geometrically distributed with a parameter that depends on k. We use a very detailed analysis to determine the tail averaged over k.

2 Proofs

2.1 The Time to Get In and Out

Let

$$T_{\text{in}} := \inf\left\{t > 0 : X_t = k\right\} \tag{2.1}$$

be the time it takes to hit the right end of the interval. Let

$$T_{\text{last}} := \sup\left\{t < T : X_t = k\right\} \tag{2.2}$$

be the last visit (if any) of the right end of the interval before returning to 0. Let

$$T_{\text{out}} := T - T_{\text{last}} \tag{2.3}$$

denote the time it takes for this last excursion from k to hit 0. Finally, let

$$T_{\text{exc}} := T_{\text{last}} - T_{\text{in}}, \tag{2.4}$$

denote the time, the random walk spends in excursions from k before the last excursion from k starts. The random times T_{exc}, T_{last} and T_{out} are well-defined on the event

$$A := \left\{T_{\text{in}} < T\right\} \tag{2.5}$$

In fact, on A, we have $T_{\text{last}} < \infty$.

Lemma 2.1

$$\mathbf{P}_k[A] = \frac{\beta - 1}{\beta - \beta^{1-k}} \geq \frac{\beta - 1}{\beta}. \tag{2.6}$$

Proof Considering $\{0, \ldots, k\}$ as an electrical network with resistances $R(l, l+1) = \beta^{-l}$, we get the effective resistances $R_{\text{eff}}(0, 1) = 1$ and

$$R_{\text{eff}}(0, k) = 1 + \beta^{-1} + \cdots + \beta^{-k+1} = \frac{1 - \beta^{-k}}{1 - 1/\beta}.$$

Now (compare, e.g., [18, (19.9)])

$$\mathbf{P}_k[A] = \frac{R_{\text{eff}}(0, 1)}{R_{\text{eff}}(0, k)} = \frac{\beta - 1}{\beta - \beta^{1-k}}.$$

□

On A^c, until time T, X is a random walk conditioned to return to 0 before hitting k. Now let U be such a random walk started in $U_0 = 0$. Let $T^U := \inf\{t > 0 : U_t = 0\}$. Then

$$\mathbf{P}_k\big[T^U = t\big] = \mathbf{P}_k\big[T = t \,\big|\, A^c\big] \quad \text{for all } t. \tag{2.7}$$

The transition probabilities of U can be computed via Doob's h-transforms. Let $h_k(l) = \beta^{-l} - \beta^{-k}$ be a harmonic (on $\{1, \ldots, k-1\}$) function for X with $h_k(k) = 0$ and $h_k(0) > 0$. Then for $l = 1, \ldots, k-1$, we have

$$\mathbf{P}_k\big[U_{t+1} = l+1 \,\big|\, U_t = l\big] = \frac{h_k(l+1)}{h_k(l)} \frac{\beta}{1+\beta} = \frac{1}{\beta+1}\left(1 - \frac{\beta - 1}{\beta^{k-l} - 1}\right). \tag{2.8}$$

We compare U to the random walk $\check{Y}$ on $\mathbb{Z}$ with conductances β^{-l} along the edge $\{l, l+1\}$. That is, $\check{Y}$ makes a jump to the right with probability $1/(1+\beta)$ and to the left with probability $\beta/(1+\beta)$. Also, let Y be the random walk on $\mathbb{Z}$ with conductances β^l along the edge $\{l, l+1\}$. That is, $-Y$ has the same jump probabilities as $\check{Y}$. Let

$$T^Y := \inf\{t > 0 : Y_t = 0\} \quad \text{and} \quad T^{\check{Y}} := \inf\{t > 0 : \check{Y}_t = 0\}.$$

Clearly, if $Y_0 = \check{Y}_0 = 0$, then T^Y and $T^{\check{Y}}$ have the same distribution. By (2.8), we see that T^U is stochastically bounded by $T^{\check{Y}}$. More precisely, we have

$$\mathbf{P}_k[T^U > t] \le \mathbf{P}\big[T^{\check{Y}} > t \,\big|\, \check{Y}_1 = 1\big]. \tag{2.9}$$

Lemma 2.2 *We have*

$$\mathbf{E}\big[T^{\check{Y}} \,\big|\, \check{Y}_1 = 1\big] = \frac{2\beta}{\beta - 1}, \qquad \mathbf{Var}\big[T^{\check{Y}} \,\big|\, \check{Y}_1 = 1\big] = \frac{4\beta(\beta+1)}{(\beta-1)^3} \tag{2.10}$$

and

$$\mathbf{E}\big[e^{\lambda T^{\check{Y}}} \,\big|\, \check{Y}_1 = 1\big] = \frac{1}{2}\left(\beta + 1 - \sqrt{(\beta+1)^2 - 4\beta e^{2\lambda}}\right) \quad \text{for all } \lambda < \log \frac{\beta+1}{2\sqrt{\beta}}. \tag{2.11}$$

By symmetry, the statements also hold for Y instead of $\check{Y}$ conditioned on $Y_1 = -1$.

Proof Define

$$\tau := \inf\{t \ge 1 : \check{Y}_t = 1\}.$$

Define the function ψ by

$$\psi(\lambda) = \mathbf{E}\big[e^{\lambda T^{\check{Y}}} \,\big|\, \check{Y}_0 = 1\big] = e^{\lambda}\, \mathbf{E}\big[e^{\lambda T^{\check{Y}}} \,\big|\, \check{Y}_1 = 1\big].$$

Decomposing according to the position of $\check{Y}$ at time 1 and using the strong Markov property at time τ (in the fourth line) yields

$$\begin{aligned}\psi(\lambda) &= e^{\lambda}\,\mathbf{E}\big[e^{\lambda T^{\check{Y}}}\mathbf{1}_{\{\check{Y}_1=0\}}\,\big|\,\check{Y}_0=1\big] + e^{\lambda}\,\mathbf{E}\big[e^{\lambda T^{\check{Y}}}\mathbf{1}_{\{\check{Y}_1=2\}}\,\big|\,\check{Y}_0=1\big]\\
&= \frac{\beta}{1+\beta}e^{2\lambda} + e^{\lambda}\,\mathbf{E}\big[e^{\lambda(T^{\check{Y}}-\tau)}\mathbf{1}_{\{\check{Y}_1=2\}}\,\big|\,\check{Y}_0=1\big]\mathbf{E}\big[e^{\lambda\tau}\,\big|\,\check{Y}_1=2\big]\\
&= \frac{\beta}{1+\beta}e^{2\lambda} + \frac{1}{1+\beta}e^{\lambda}\,\mathbf{E}\big[e^{\lambda(T^{\check{Y}}-\tau)}\,\big|\,\check{Y}_1=2\big]\mathbf{E}\big[e^{\lambda\tau}\,\big|\,\check{Y}_1=2\big]\\
&= \frac{\beta}{1+\beta}e^{2\lambda} + \frac{1}{1+\beta}e^{\lambda}\mathbf{E}\big[e^{\lambda T^{\check{Y}}}\,\big|\,\check{Y}_0=1\big]\mathbf{E}\big[e^{\lambda\tau}\,\big|\,\check{Y}_1=2\big]\\
&= \frac{\beta}{1+\beta}e^{2\lambda} + \frac{1}{1+\beta}\psi(\lambda)^2.\end{aligned}$$

This quadratic equation has two solutions which at $\lambda=0$ take the values 1 and β, respectively. The relevant one takes the value 1 and is given in (2.11). Taking the derivatives at $\lambda=0$ gives (2.10). □

Lemma 2.3 *There exists an* $\varepsilon>0$ *such that*

$$\mathbb{P}[T>t\,|\,A^c]\le e^{-\varepsilon t},\qquad t\ge 1. \tag{2.12}$$

Proof This is a direct consequence of (2.7), (2.9) and the existence of exponential moments (Lemma 2.2). □

Lemma 2.4 *There exists an* $\varepsilon>0$ *such that*

$$\mathbb{P}[T_{\mathrm{in}}>t\,|\,A]\le e^{-\varepsilon t},\qquad t\ge 1. \tag{2.13}$$

Proof By Lemma 2.1, it is enough to show

$$\mathbb{P}[T_{\mathrm{in}}>t]\le e^{-\varepsilon t},\qquad t\ge 1. \tag{2.14}$$

Note that X and Y can be coupled such that $X_t\ge Y_t$ for all $t\le T_{\mathrm{in}}$. Hence

$$\mathbf{P}_k[T_{\mathrm{in}}>t]\le \mathbf{P}[Y_t<k\,|\,Y_0=0].$$

Now Y_t is a sum of t i.i.d. random variables and $\mathbf{E}[Y_t]=\frac{\beta-1}{\beta+1}t>0$, hence by Cramér's theorem, there exists an $\varepsilon>0$ such that $\mathbf{P}[Y_t<k]\le e^{-\varepsilon t}$ for $t\,\frac{\beta-1}{\beta+1}\ge 2k$.

Hence

$$\mathbf{P}_k[T_{\mathrm{in}}>t]\le e^{-\varepsilon t}\text{ for }t\,\frac{\beta-1}{\beta+1}\ge 2k.$$

Concluding, we have

$$\mathbb{P}[T_{\mathrm{in}}>t]\le e^{-\varepsilon t}+(1-\alpha)\sum_{2k>t(\beta-1)/(\beta+1)}\alpha^k\le e^{-\varepsilon t}+\alpha^{t(\beta-1)/2(\beta+1)}.$$

□

Lemma 2.5 *There exists an* $\varepsilon>0$ *such that*

$$\mathbb{P}[T_{\mathrm{out}}>t\,|\,A]\le e^{-\varepsilon t},\qquad t\ge 1. \tag{2.15}$$

Furthermore, for each $k\in\mathbb{N}$,

$$\mathbf{E}_k\big[T_{\mathrm{out}}\,|\,A\big]\le k\,\frac{\beta+1}{\beta-1}. \tag{2.16}$$

Proof Let V be a random walk on $\{0, \dots, k\}$ with the same transition probabilities as U (see (2.8)) but started at k. Define $T^V := \inf\{t > 0 : V_t = 0\}$.

Note that V can be coupled with $\check{Y}$ (started in $\check{Y}_0 = k$) such that $V_t \le \check{Y}_t$ for $t \le T^V$. Arguing as in the proof of Lemma 2.4, we get an $\varepsilon > 0$ such that

$$\mathbb{P}\big[T_{\text{out}} > t \,|\, A\big] = \mathbb{P}\big[T^V > t\big] \le \mathbf{P}\big[\check{Y}_t > 0\big] \le e^{-\varepsilon t}, \qquad t \ge 1.$$

Let $T_0^{\check{Y}} := \inf\{t > 0 : \check{Y}_t = 0\}$. Note that $\check{Y}$ has a drift $\frac{\beta-1}{\beta+1}$ to the left. Hence, the average time it takes to visit the point left of the starting point is $\frac{\beta+1}{\beta-1}$. Now $T_0^{\check{Y}}$ is the time it takes to visit the kth point left of the staring point. Hence, again by stochastic domination,

$$\mathbf{E}_k\big[T_{\text{out}}|A\big] \le \mathbf{E}_k\Big[T_0^{\check{Y}}\Big] = k\frac{\beta+1}{\beta-1}.$$

□

2.2 The Time Spent in Excursions

Recall that $T = T_{\text{in}} + T_{\text{exc}} + T_{\text{out}}$. We have dealt with T_{in} and T_{out}. Now we turn to the time T_{exc} the random walk X spends in excursions from k before it hits 0. These excursions are pieces of the random walk conditioned not to hit 0. Let N denote the number of these excursions if A occurs and $N = 0$ on A^c. Note that N is geometrically distributed with respect to the conditional probability $\mathbf{P}_k[\,\cdot\,|A]$.

Our strategy is

- to compute the parameter of N (depending on k),
- to estimate expectation and exponential moments of the lengths of the excursions and
- to show that for the tail of T, it is good enough to replace the lengths of the excursions by their expected value.

Hence, the tail of N rules the game, see Proposition 2.17.

Finally, we will compute the tail of N with an involved analysis using Mellin transforms.

Let $\bar{X}$ be the random walk on $\{0, \dots, k\}$ started in $\bar{X}_0 = k$. Let

$$\bar{T}_0 := \inf\left\{t > 0 : \bar{X}_t = 0\right\}$$

and

$$\bar{T}_k := \inf\left\{t > 0 : \bar{X}_t = k\right\}.$$

Let

$$B := \{\bar{T}_k < \bar{T}_0\} = \{\bar{X} \text{ returns to} k \text{before hitting} 0\}.$$

Lemma 2.6 *We have*

$$\mathbf{P}_k[B] = 1 - \frac{\beta-1}{\beta^k - 1}.$$

Proof This is similar to the proof of Lemma 2.1. □

Let $\check{X}$ be the random walk on $\{0, \dots, k\}$ started at $\check{X}_0 = k$ and conditioned to return to k before hitting 0. This means the transition probabilities of $\check{X}$ are given by Doob's h-transform with the harmonic function $h_0(l) = 1 - \beta^{-l}$. Explicitly, we have

$$\mathbf{P}_k\big[\check{X}_{t+1} = l+1 \,\big|\, \check{X}_t = l\big] = \frac{\beta}{\beta+1}\frac{h(l+1)}{h(l)} = \frac{\beta}{\beta+1}\frac{\beta^{l+1}-1}{\beta^{l+1}-\beta} > \frac{\beta}{\beta+1}. \tag{2.17}$$

Let

$$T_k^{\check{X}} := \inf\left\{t > 0 : \check{X}_t = k\right\}. \tag{2.18}$$

Then

$$\mathbf{P}_k\left[T_k^{\check{X}} = t\right] = \mathbf{P}_k\left[\bar{T}_k = t \,\middle|\, B\right].$$

Let $N, T^{(1)}, T^{(2)}, \dots$ be independent random variables with respect to $\mathbf{P}_k$ and such that N is geometrically distributed with parameter $\mathbf{P}_k[B^c] = \frac{\beta-1}{\beta^k-1}$ and

$$\mathbf{P}_k[T^{(i)} = l] = \mathbf{P}_k[\bar{T}_k = l \,|\, B], \qquad l \in \mathbb{N}_0,\ i = 1, 2, \dots.$$

Also let

$$\widetilde{T} := \sum_{i=1}^{N} T^{(i)}. \tag{2.19}$$

Lemma 2.7 *We have*

$$\mathbf{P}_k\left[\widetilde{T} = t\right] = \mathbf{P}_k\left[T_{\mathrm{exc}} = t \,\middle|\, A\right], \qquad t \in \mathbb{N}_0.$$

Proof This is a simple application of the strong Markov property. □

Lemma 2.8

$$\mathbf{E}_k[N] \le \mathbf{E}_k[N \,|\, A] = \frac{\mathbf{P}_k[B]}{\mathbf{P}_k[B^c]} = \frac{\beta^k - \beta}{\beta - 1} \tag{2.20}$$

and (since $N = 0$ if $k = 0$)

$$\mathbb{E}[N] \le \mathbb{E}[N \,|\, A] = (1-\alpha)\sum_{k=1}^{\infty} \alpha^k\, \mathbf{E}_k[N] = \frac{\alpha^2\beta}{1-\alpha\beta}. \tag{2.21}$$

Proof This is a direct computation. □

While $\widetilde{T}$ is the quantity we have to study, it is more convenient to get rid of the randomness inherent in the lengths of the excursions and to replace them by their expected value. Hence, as a substitute for $\widetilde{T}$, we introduce

$$\hat{T} := N \cdot \mathbf{E}_k\left[T^{(1)}\right]. \tag{2.22}$$

In order to show that $\widetilde{T}$ and $\hat{T}$ are in fact close, we estimate the exponential moments of $T^{(1)}$ and use Markov's inequality. As a direct computation of the exponential moments is a bit tricky, we make a little detour and use a comparison argument for branching processes. We prepare for this comparison argument with some considerations on the convex ordering of geometric distributions. Note that for the case $\varrho < 2$, a simpler estimate based on variances would be good enough for our purposes. In fact, the variances exist for any fixed k and give estimates of order t^{-2} which is good enough compared with the leading order term $t^{-\varrho}$ if $\varrho < 2$.

Lemma 2.9 *We can define a family $(W_r)_{r\in(0,1]}$ of geometrically distributed random variables with parameters r, such that*

$$W_r \mathit{and} W_q - W_r \mathit{are\ independent\ if}\ 0 < q \le r \le 1.$$

We have

$$\mathbf{P}[W_q - W_r = k] = \begin{cases} q\left(1 - \frac{q}{r}\right)(1-q)^{k-1}, & \text{if } k = 1, 2, \dots, \\ \frac{q}{r}(r - q), & \text{if } k = 0. \end{cases} \tag{2.23}$$

Proof Let $(U_n)_{n\in\mathbb{N}_0}$ be i.i.d. random variables uniformly distributed on $[0, 1]$. Let

$$W_r := \inf\{n : U_n \le r\}.$$

It is easy to check that the (W_r) have the desired properties. $\square$

Lemma 2.10 *Let $0 < q \le r \le 1$ and let W_q and W_r be geometrically distributed with parameters q and r, respectively. Let $\varphi : \mathbb{R} \to [0, \infty)$ be a convex function. Then*

$$\mathbf{E}\left[\varphi(W_r - \mathbf{E}[W_r])\right] \le \mathbf{E}\left[\varphi(W_q - \mathbf{E}[W_q])\right] \tag{2.24}$$

Proof By Lemma 2.9, we may and will assume that W_r and $W_q - W_r$ are independent. Hence

$$W_r - \mathbf{E}[W_r] = \mathbf{E}\big[W_r - \mathbf{E}[W_r]\,\big|\, W_r\big] = \mathbf{E}\big[W_q - \mathbf{E}[W_q]\,\big|\, W_r\big].$$

By Jensen's inequality, we get

$$\begin{aligned}\mathbf{E}\big[\varphi\big(W_r - \mathbf{E}[W_r]\big)\big] &= \mathbf{E}\big[\varphi\big(\mathbf{E}[W_q - \mathbf{E}[W_q]\,|\, W_r]\big)\big]\\ &\le \mathbf{E}\big[\mathbf{E}\big[\varphi(W_q - \mathbf{E}[W_q])\,\big|\, W_r\big]\big] \\ &= \mathbf{E}\big[\varphi(W_q - \mathbf{E}[W_q])\big].\end{aligned} \tag{2.25}$$

$\square$

Corollary 2.11 *For $\lambda \in \mathbb{R}$, $\kappa \ge 1$ and $0 < q \le r \le 1$, we have*

$$\mathbf{E}\left[e^{\lambda(W_r - \mathbf{E}[W_r])}\kappa^{W_r}\right] \le \mathbf{E}\left[e^{\lambda(W_q - \mathbf{E}[W_q])}\kappa^{W_q}\right]. \tag{2.26}$$

Proof Let $\varphi(x) := e^{\lambda x}\kappa^x$. Since $\mathbf{E}[W_q] \ge \mathbf{E}[W_r]$, we get by Lemma 2.10

$$\begin{aligned}\mathbf{E}\left[e^{\lambda(W_r - \mathbf{E}[W_r])}\kappa^{W_r}\right] &= \mathbf{E}\left[\varphi(W_r - \mathbf{E}[W_r])\right]\,\kappa^{\mathbf{E}[W_r]}\\ &\le \mathbf{E}\left[\varphi(W_q - \mathbf{E}[W_q])\right]\,\kappa^{\mathbf{E}[W_q]}\\ &= \mathbf{E}\left[e^{\lambda(W_q - \mathbf{E}[W_q])}\kappa^{W_q}\right].\end{aligned}$$

$\square$

Lemma 2.12 *Let $Z^{(1)}$ and $Z^{(2)}$ be two Galton–Watson branching processes with generation dependent offspring laws and $Z_0^{(1)} = Z_0^{(2)} = 1$. Let*

$$\check{Z}^{(i)} := \sum_{n=0}^{\infty} Z_n^{(i)}, \qquad i = 1, 2,$$

be the total population sizes. The offspring law of $Z^{(i)}$ in generation n is assumed to be geometric with parameter $p^{i,n}$, $i = 1, 2$, $n \in \mathbb{N}_0$. We also assume that $p^{1,n} \le p^{2,n}$ for all $n \in \mathbb{N}_0$ and

$$\mathbf{E}\big[(\check{Z}^{(1)})^2\big] < \infty.$$

Then, we have

$$\mathbf{E}\big[\check{Z}^{(2)}\big] \le \mathbf{E}\big[\check{Z}^{(1)}\big] < \infty \quad \text{and} \quad \mathbf{Var}\big[\check{Z}^{(2)}\big] \le \mathbf{Var}\big[\check{Z}^{(1)}\big] < \infty. \tag{2.27}$$

For all $\lambda \in \mathbb{R}$ with $\mathbf{E}[e^{\lambda \check{Z}^{(1)}}] < \infty$, we have

$$\mathbf{E}\left[\exp\left(\lambda(\check{Z}^{(2)} - \mathbf{E}[\check{Z}^{(2)}])\right)\right] \le \mathbf{E}\left[\exp\left(\lambda(\check{Z}^{(1)} - \mathbf{E}[\check{Z}^{(1)}])\right)\right]. \tag{2.28}$$

In particular, for $\lambda \geq 0$,

$$\mathbf{E}\left[\exp\left(\lambda \check{Z}^{(2)}\right)\right] \leq \mathbf{E}\left[\exp\left(\lambda \check{Z}^{(1)}\right)\right]. \tag{2.29}$$

Proof First assume that

$$p^{1,n} = 1 \quad \text{for } n \geq n_0 \text{for some} n_0. \tag{2.30}$$

Hence $\check{Z}^{(i)} = Z_0^{(i)} + \cdots + Z_{n_0}^{(i)}$, $i = 1, 2$. For $n_0 = 1$, the statement follows from the expectation and variance formula for the geometric distribution. The induction step from $n_0 - 1$ to n_0 is a simple application of Wald's formula and the Blackwell-Girshick formula. In order to get rid of assumption (2.30), take monotone limits.

For the exponential inequalities we proceed similarly. Consider first the case (2.30) and $n_0 = 1$. In this case the assertion is a direct consequence of Corollary 2.11. For the induction step, we assume that the statement is true for $n_0 - 1$ and we show it for n_0. Define

$$\kappa^{(i)} := \mathbf{E}[\exp(\lambda(Z_2^{(i)} + \cdots + Z_{n_0}^{(i)} - \mathbf{E}[Z_2^{(i)} + \cdots + Z_{n_0}^{(i)}])) \,\big|\, Z_1^{(i)} = 1].$$

By the induction hypothesis, applied to the branching processes started at time 1 instead of 0, we have

$$1 \leq \kappa^{(2)} \leq \kappa^{(1)}.$$

By decomposing according to the value of $Z_1^{(i)}$, we infer (again for the processes started at time 0)

$$\begin{aligned} \mathbf{E}\left[\exp\left(\lambda\left(\check{Z}^{(2)} - \mathbf{E}[\check{Z}^{(2)}]\right)\right)\right] &= \mathbf{E}\left[\exp\left(\lambda\left(Z_1^{(2)} + \cdots + Z_{n_0}^{(2)} - \mathbf{E}[Z_1^{(2)} + \cdots + Z_{n_0}^{(2)}]\right)\right)\right] \\ &= \mathbf{E}\left[\exp\left(\lambda\left(Z_1^{(2)} - \mathbf{E}[Z_1^{(2)}]\right)\right)(\kappa^{(2)})^{Z_1^{(2)}}\right] \\ &\leq \mathbf{E}\left[\exp\left(\lambda\left(Z_1^{(2)} - \mathbf{E}[Z_1^{(2)}]\right)\right)(\kappa^{(1)})^{Z_1^{(2)}}\right] \\ &\leq \mathbf{E}\left[\exp\left(\lambda\left(Z_1^{(1)} - \mathbf{E}[Z_1^{(1)}]\right)\right)(\kappa^{(1)})^{Z_1^{(1)}}\right] \\ &= \mathbf{E}\left[\exp\left(\lambda\left(\check{Z}^{(1)} - \mathbf{E}[\check{Z}^{(1)}]\right)\right)\right], \end{aligned} \tag{2.31}$$

where in the fourth line we used Corollary 2.11 and the assumption $p^{1,1} \leq p^{2,1}$. □

Lemma 2.13 *We have*

$$\frac{2\beta}{\beta - 1} - \frac{2\beta(\beta+1)}{\beta - 1} k\,\beta^{-k} \leq \mathbf{E}_k\left[T^{(1)}\right] \leq \frac{2\beta}{\beta - 1} \quad \textit{for all } k \geq 2 \tag{2.32}$$

and

$$\mathbf{Var}_k\left[T^{(1)}\right] \leq \frac{4\beta(\beta+1)}{(\beta-1)^3} \quad \textit{for all } k \geq 2. \tag{2.33}$$

Furthermore, there is a $\delta > 0$ *such that*

$$\mathbf{E}_k\left[e^{\lambda(T^{(1)} - \mathbf{E}_k[T^{(1)}])}\right] \leq 1 + \frac{4\beta(\beta^2+1)}{(\beta-1)^3}\lambda^2 \quad \textit{for all } \lambda \in [-\delta, \delta],\ k \geq 2. \tag{2.34}$$

Proof Let Y be the random walk on $\mathbb{Z}$ that jumps to the right with probability $\beta/(1+\beta)$ and to the left with probability $1/(1+\beta)$ starting in $k-1$. Let

$$T_l^Y := \inf\left\{t > 0 : Y_t = l\right\}, \qquad l = 0, \ldots, k.$$

Recall $T_k^{\check{X}}$ from (2.18). By the basic connection between the occupation times of excursions of random walks and Galton–Watson processes with geometric offspring distributions, we see that $\frac{1}{2}(T_k^Y+1)$ has the same distribution as $\check{Z}^{(1)}$ from Lemma 2.12 with $p^{1,n} \equiv \frac{\beta}{\beta+1}$. Similarly, using (2.17), we see that $\frac{1}{2}T_k^{\check{X}}$ has the same distribution as $\check{Z}^{(2)}$ with

$$p^{2,n} = \frac{\beta}{\beta+1}\frac{\beta^{k-n+1}-1}{\beta^{k-n+1}-\beta} > p^{1,n}, \quad n = 0, \dots, k-1.$$

By Lemmas 2.12 and 2.2, we infer

$$\mathbf{E}_k\big[T^{(1)}\big] = \mathbf{E}_k\big[T_k^{\check{X}}\big] \le 1 + \mathbf{E}_k\big[T_k^Y\big] = \frac{2\beta}{\beta-1} \tag{2.35}$$

and

$$\mathbf{Var}_k\big[T^{(1)}\big] = \mathbf{Var}_k\big[T_k^{\check{X}}\big] \le \mathbf{Var}\big[T_k^Y\big] = \frac{4\beta(\beta+1)}{(\beta-1)^3}. \tag{2.36}$$

On the other hand,

$$\begin{aligned}\mathbf{E}_k\big[T^{(1)}\big] &= 1 + \mathbf{E}\big[T_k^Y \,\big|\, T_k^Y < T_0^Y\big] \ge 1 + \mathbf{E}\big[T_k^Y \mathbf{1}_{\{T_k^Y < T_0^Y\}}\big] \\ &= 1 + \mathbf{E}\big[T_k^Y\big] - \mathbf{P}_k\big[T_k^Y > T_0^Y\big]\mathbf{E}\big[T_k^Y \,\big|\, T_k^Y > T_0^Y\big].\end{aligned}$$

By Lemma 2.5, we get

$$\mathbf{E}\big[T_0^Y \,\big|\, T_k^Y > T_0^Y\big] = \mathbf{E}_k\big[T_{\mathrm{out}} \,\big|\, A\big] \le \frac{\beta+1}{\beta-1}k.$$

Using the Markov property and arguing as in Lemma 2.5, we get

$$\mathbf{E}\big[T_k^Y - T_0^Y \,\big|\, T_k^Y > T_0^Y\big] = \frac{\beta+1}{\beta-1}k.$$

Summing up and using Lemma 2.6 to get $\mathbf{P}_k[T_k^Y > T_0^Y] = \frac{\beta-1}{\beta^k-1}$, we have

$$\mathbf{E}_k\big[T^{(1)}\big] \ge \frac{2\beta}{\beta-1} - \frac{\beta-1}{\beta^k-1}k\,2\frac{\beta+1}{\beta-1}.$$

Now we turn to the proof of (2.34). Again by Lemmas 2.12 and 2.2, we get for $\lambda < \log\frac{\beta+1}{2\sqrt{\beta}}$

$$\begin{aligned}\mathbf{E}_k\Big[e^{\lambda(T^{(1)}-\mathbf{E}_k[T^{(1)}])}\Big] &\le \mathbf{E}\Big[e^{\lambda(T^Y-\mathbf{E}[T^Y])}\Big] \\ &= F(\lambda) := \frac{1}{2}\left(\beta+1-\sqrt{(\beta+1)^2-4\beta e^{2\lambda}}\right)\cdot e^{-(2\beta/(\beta-1))\lambda}.\end{aligned} \tag{2.37}$$

The first and second derivatives at zero are

$$F'(0) = 0 \quad \text{and} \quad F''(0) = \frac{4\beta(\beta^2+1)}{(\beta-1)^3}.$$

Hence, by Taylor's theorem, there exists a $\delta > 0$ such that

$$F(\lambda) \le F(0) + F''(0)\lambda^2 = 1 + \frac{4\beta(\beta^2+1)}{(\beta-1)^3}\lambda^2 \quad \text{for all } \lambda \in [-\delta, \delta]. \tag{2.38}$$

Combining (2.37) and (2.38) gives (2.34). □

Recall $\widetilde{T}$ and $\hat{T}$ from (2.19) and (2.22), respectively. We now use the exponential moment estimates on $T^{(1)}$ to get that $\widetilde{T}$ and $\hat{T}$ are close.

Lemma 2.14 *There is a constant $c > 0$, such that for all $t > 0$, we have*

$$\mathbb{P}\big[|\widetilde{T} - \hat{T}| > t\big] \le c\, t^{-2\varrho}. \tag{2.39}$$

Proof By Markov's inequality and Lemma 2.13, there are $\delta > 0$ and $C < \infty$ such that for $\lambda \in [0, \delta]$,

$$\begin{aligned} \mathbf{P}_k\big[|\widetilde{T} - \hat{T}| > t \,\big|\, N\big] &\le e^{-\lambda t}\, \mathbf{E}_k\big[\exp(\lambda|\widetilde{T} - \hat{T}|) \,\big|\, N\big] \\ &= e^{-\lambda t}\, \mathbf{E}_k\big[\exp(\lambda|T^{(1)} - \mathbf{E}_k[T^{(1)}]|)\big]^N \\ &\le 2e^{-\lambda t}\, (1 + C\lambda^2)^N \\ &\le 2e^{-\lambda t}\, e^{C\lambda^2 N}. \end{aligned} \tag{2.40}$$

We need to make a good choice for λ to make this inequality effective. Recall that N is geometric with parameter $r_k := \frac{\beta-1}{\beta^k-1}$ under the conditional probability $\mathbf{P}_k[\,\cdot\,|\,A]$. Define

$$\lambda_k := \sqrt{\frac{1}{C}\log\left(\frac{1 - r_k/2}{1 - r_k}\right)}, \qquad k = 2, 3, \ldots \tag{2.41}$$

Then we have

$$\mathbf{E}_k\big[e^{C\lambda_k^2 N} \,\big|\, A\big] = \frac{r_k}{1 - (1 - r_k)e^{C\lambda_k^2}} = \frac{r_k}{1 - (1 - r_k/2)} = 2, \quad k \ge 2, \tag{2.42}$$

and for $l > k$,

$$\mathbf{E}_k\big[e^{C\lambda_l^2 N} \,\big|\, A\big] \le \mathbf{E}_k\big[e^{C\lambda_k^2 N} \,\big|\, A\big] = 2. \tag{2.43}$$

Note that $\frac{\beta-1}{\beta^k-\beta} < \frac{1}{2}$ for all $k \ge 2$. Hence (using the fact that $\log(1 + x) \ge x/2$ for $x \in [0, 1/2]$),

$$\lambda_k = \sqrt{\frac{1}{C}\log\left(1 + \frac{1}{2}\frac{\beta - 1}{\beta^k - \beta}\right)} \ge \sqrt{\frac{\beta - 1}{4C}}\, \beta^{-k/2} \quad \text{for all } k \ge 2. \tag{2.44}$$

Let $C' := \sqrt{\frac{\beta-1}{4C}}$. Note that $\lambda_k \downarrow 0$ and let $k_0 \in \mathbb{N}$ be large enough such that $\lambda_k < \delta$ for all $k \ge k_0$.

Then (using Lemma 2.18 with $\sqrt{\beta}$ instead of β and hence 2ϱ instead of ϱ in the last step) there is a constant $\tilde{C} < \infty$ such that

$$\begin{aligned} \mathbb{P}\big[|\widetilde{T} - \hat{T}| > t\big] &\le 2(1-\alpha)\sum_{k=1}^{\infty} \alpha^k e^{-\lambda_{k\vee k_0} t}\, \mathbf{E}_k\Big[e^{C\lambda_{k\vee k_0}^2 N} \,\Big|\, A\Big] \\ &\le 4e^{-\lambda_{k_0} t} + 2(1-\alpha)\sum_{k=k_0+1}^{\infty} \alpha^k e^{-\lambda_k t}\, \mathbf{E}_k\Big[e^{C\lambda_k^2 N} \,\Big|\, A\Big] \\ &\le 4e^{-\lambda_{k_0} t} + 4(1-\alpha)\sum_{k=k_0+1}^{\infty} \alpha^k e^{-C'\beta^{-k/2} t} \\ &\le 4e^{-\lambda_{k_0} t} + \tilde{C} t^{-2\varrho}. \end{aligned} \tag{2.45}$$

Since $\lambda_{k_0} > 0$ is a constant, the claim follows. □

It is still a bit inconvenient to work with $\hat{T}$ as the expectation of $T^{(1)}$ depends on k, though only slightly. The next step is to replace $\mathbf{E}_k[T^{(1)}]$ in the definition of $\hat{T}$ by its limit $\lim_{k\to\infty} \mathbf{E}_k[T^{(1)}] = \frac{2\beta}{\beta-1}$.

Lemma 2.15 *There is a constant $c > 0$, such that for all $t > 0$, we have*

$$\mathbb{P}\left[\left|\hat{T} - N\frac{2\beta}{\beta-1}\right| > t\right] \le e^{-c\sqrt{t}}. \tag{2.46}$$

Proof By Lemma 2.13, and by the fact that $\hat{T} = 2N$ if $k = 1$, we know that

$$\left|\hat{T} - N\frac{2\beta}{\beta-1}\right| \le N\frac{2\beta(\beta+1)}{\beta-1}k\beta^{-k} \quad \text{for all } k \ge 1.$$

Hence for any $k_0 \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\left[\left|\hat{T} - N\frac{2\beta}{\beta-1}\right| > t\right] &\le \sum_{k=1}^{\infty}(1-\alpha)\alpha^k\, \mathbf{P}_k\left[N\frac{2\beta(\beta+1)}{\beta-1}k\beta^{-k} > t\right] \\
&\le \alpha^{k_0+1} + \sum_{k=1}^{k_0}(1-\alpha)\alpha^k\left(1 - \frac{\beta-1}{\beta^k-1}\right)^{t\beta^k k^{-1}\frac{\beta-1}{2\beta(\beta+1)}} \\
&\le \alpha^{k_0+1} + \exp\left(-\frac{(\beta-1)^2}{2\beta(\beta+1)}k_0^{-1}t\right).
\end{aligned}$$

Now choose $k_0 = \sqrt{t}$ to get the result. □

In order to see that the error terms are smaller than the main term, that is the tail of N, we need a lower bound for the tail of N. Since we give a more detailed analysis later, here we only make a very rough assertion.

Lemma 2.16 *There exists a constant $c > 0$ such that*

$$\mathbb{P}[N > t] \ge ct^{-\varrho} \quad \textit{for all } t \ge 1.$$

Proof For $t \in [1, \beta^2]$, the statement holds with $c = \mathbf{P}[N > \beta^2]$. Now assume $t \ge \beta^2$ and let $c = \frac{\beta-1}{\beta}(1-\alpha)e^{-2\beta^2}$. Let $k \in \mathbb{N}$, $k \ge 2$ be such that $\beta^k \le t \le \beta^{k+1}$. Then (recall Lemma 2.1 and note that $1 - x \ge e^{-2x}$ for $x \in [0, 1/2]$)

$$\begin{aligned}
\mathbb{P}[N > t] &\ge \frac{\beta-1}{\beta}\,\mathbb{P}[N > t\,|\,A] \\
&\ge \frac{\beta-1}{\beta}(1-\alpha)\alpha^k\left(1 - \frac{\beta-1}{\beta^k-1}\right)^t \\
&\ge \frac{\beta-1}{\beta}(1-\alpha)\exp\left(-2\frac{\beta-1}{\beta^k-1}\beta^{k+1}\right)\alpha^k \\
&\ge c\,\alpha^k \ge c\,t^{-\varrho}.
\end{aligned}$$

□

We summarize the above discussion in the following proposition.

Proposition 2.17 *We have*

$$\lim_{t\to\infty}\frac{\mathbb{P}\left[N > \frac{\beta-1}{2\beta}t\right]}{\mathbb{P}[T > t]} = 1.$$

Proof By Lemma 2.15, the tails of $\frac{2\beta}{\beta-1}N$ and $\hat{T}$ coincide in our scale, given by Lemma 2.16. By Lemma 2.14, the tails of $\widetilde{T}$ and $\hat{T}$ coincide. Finally, by Lemmas 2.4, 2.5 and 2.7 the tails of T and $\widetilde{T}$ coincide. □

2.3 The Tail of a Geometric Random Variable with Random Parameter

In order to compute the tail of N, it is convenient to replace the geometrically distributed random variable with parameter $\frac{\beta-1}{\beta^k-1}$ by an exponentially distributed random variable N' with parameter β^{-k}. Note that we neglected the factor $\beta - 1$ and we will re-introduce it by a scaling of t. The tail of N' is given by

$$\mathbb{P}[N' > t] = f(t) := \sum_{k=0}^{\infty} (1-\alpha)\alpha^k \exp\left(-\beta^{-k}t\right), \qquad t > 0. \tag{2.47}$$

Lemma 2.18 *There are constants* $0 < C_1 < C_2 < \infty$ *such that*

$$C_1 t^{-\varrho} \le f(t) \le C_2 t^{-\varrho} \quad \textit{for all } t > 1.$$

Proof Let $k \in \mathbb{N}_0$ be chosen such that $\beta^{k-1} \le t < \beta^k$. Recall that $\varrho = -\log(\alpha)/\log(\beta)$. Then

$$f(t) \ge f(\beta^k) \ge (1-\alpha)\alpha^k e^{-1} = (1-\alpha)e^{-1}(\beta^k)^{-\varrho} \ge (1-\alpha)e^{-1}\beta^{-\varrho} t^{-\varrho}. \tag{2.48}$$

Let

$$C_2 := \alpha^{-1} \sum_{k=-\infty}^{\infty} (1-\alpha)\alpha^k \exp\left(-\beta^{-k}\right).$$

Note that f is decreasing and hence for $l \in \mathbb{Z}$ and $\beta^{l+1} > t \ge \beta^l$, we have

$$\begin{aligned} f(t) \le f(\beta^l) &= \sum_{k=0}^{\infty} (1-\alpha)\alpha^k \exp\left(-\beta^{l-k}\right) \\ &= \alpha^l \sum_{k=-l}^{\infty} (1-\alpha)\alpha^k \exp\left(-\beta^{-k}\right) \le C_2 \alpha^{l+1} \le C_2 t^{-\varrho}. \end{aligned}$$

□

Lemma 2.19 *We have*

$$\lim_{t\to\infty} \frac{\mathbb{P}[N > t]}{f\left((\beta-1)t\right)} \frac{\beta}{\beta - 1} = 1. \tag{2.49}$$

Proof By Lemmas 2.16 and 2.18, all error terms of order $o(t^{-\varrho})$ can be neglected. We use this first to show that the summands of $f(t)$ with $\beta^k \le t^{2/3}$ are negligible:

$$\begin{aligned} (1-\alpha) \sum_{k:\beta^k \le t^{2/3}} \alpha^k (1-\beta^{-k})^t &\le (1-\alpha) \sum_{k:\beta^k \le t^{2/3}} \alpha^k \exp\left(-\beta^{-k} t\right) \\ &\le (1-\alpha) \sum_{k=0}^{\infty} \alpha^k \exp\left(-t^{1/3}\right) \\ &= \exp\left(-t^{1/3}\right). \end{aligned} \tag{2.50}$$

For N note that

$$\mathbf{P}_k[N > t] \le \mathbf{P}_k[N > t \,|\, A] = \left(1 - \frac{\beta - 1}{\beta^k - 1}\right)^t \le \exp\left(-\beta^{-k}(\beta-1)t\right). \tag{2.51}$$

We use this first to show as in (2.50) that the summands of $\mathbf{P}[N > t]$ with $\beta^k \le t^{2/3}$ are negligible:

$$(1-\alpha) \sum_{k:\beta^k \le t^{2/3}} \alpha^k \, \mathbf{P}_k[N > t] \le \quad \exp\left(-((\beta-1)t)^{1/3}\right). \tag{2.52}$$

Recall from Lemma 2.1 that $\mathbf{P}_k[A] = \frac{\beta-1}{\beta-\beta^{1-k}}$. Let $\varepsilon > 0$ and choose t large enough such that $\mathbf{P}_k[A] \le (1+\varepsilon)\frac{\beta-1}{\beta}$ for all k such that $\beta^k > t^{2/3}$. Then

$$\begin{aligned}
\mathbb{P}[N > t] &\le (1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \, \mathbf{P}_k[N > t] + \quad \exp\left(-((\beta-1)t)^{1/3}\right) \\
&\le (1+\varepsilon)\frac{\beta-1}{\beta}(1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \, \mathbf{P}_k[N > t \,|\, A] + \quad \exp\left(-((\beta-1)t)^{1/3}\right) \\
&\le (1+\varepsilon)\frac{\beta-1}{\beta}(1-\alpha) \sum_{k=0}^{\infty} \alpha^k \, \exp\left(-\beta^{-k}(\beta-1)t\right) + \quad \exp\left(-((\beta-1)t)^{1/3}\right) \\
&= (1+\varepsilon)\frac{\beta-1}{\beta} f\left((\beta-1)t\right) + \quad \exp\left(-((\beta-1)t)^{1/3}\right).
\end{aligned} \tag{2.53}$$

This shows

$$\limsup_{t\to\infty} \frac{\mathbb{P}[N > t]}{f\left((\beta-1)t\right)} \frac{\beta}{\beta-1} \le 1.$$

Now we come to the complementary estimate for the lim inf.

Note that $\log(1-x) \ge -x - x^2$ for $x \in [0, 1/2]$. For the summands of $\mathbf{P}[N > t]$ with $\beta^k > t^{2/3}$, and for $t \ge \beta^3$, we have $k \ge 2$ (thus $\frac{\beta-1}{\beta^k-1} \le \frac{1}{\beta+1} \le \frac{1}{2}$) and hence

$$\log\left(1 - \frac{\beta-1}{\beta^k-1}\right) \ge -\frac{\beta-1}{\beta^k-1} - \left(\frac{\beta-1}{\beta^k-1}\right)^2 \ge -\frac{\beta-1}{\beta^k} - \beta^{2-2k}.$$

We infer for $C = C(\beta)$ large enough and all $t \ge 2$,

$$\begin{aligned}
\mathbb{P}[N > t] &= \mathbb{P}[N > t \,|\, A]\, \mathbb{P}[A] \ge \mathbb{P}[N > t \,|\, A] \frac{\beta-1}{\beta} \\
&\ge \frac{\beta-1}{\beta}(1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \left(1 - \frac{\beta-1}{\beta^k-1}\right)^t \\
&\ge \frac{\beta-1}{\beta}(1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \exp\left(-\beta^{-k}(\beta-1)t\right) \exp\left(-\beta^{2-2k} t\right) \\
&\ge \frac{\beta-1}{\beta}(1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \exp\left(-\beta^{-k}(\beta-1)t\right) \exp\left(-\beta^2 t^{-1/3}\right) \\
&\ge \frac{\beta-1}{\beta}\left(1 - \beta^2\, t^{-1/3}\right)(1-\alpha) \sum_{k:\beta^k > t^{2/3}} \alpha^k \exp\left(-\beta^{-k}(\beta-1)t\right) \\
&\ge \frac{\beta-1}{\beta}\left(1 - \beta^2\, t^{-1/3}\right)\left(f\left((\beta-1)t\right) - \quad \exp\left(-((\beta-1)t)^{1/3}\right)\right)
\end{aligned}$$

$$\geq \frac{\beta-1}{\beta}\left(1-C\,t^{-1/3}\right)\,f\big((\beta-1)t\big).$$

□

Remark 2.20 Our comparison of the tails of N' and T in Lemma 2.19 and Proposition 2.17 allows to recover a result of Solomon [21] which we briefly sketch here.

Let $\nu := \frac{2\beta}{(1-\beta)^2}$ and let ψ be the Laplace transform of $\nu\,N'$, that is,

$$\psi(\lambda)=\mathbf{E}\left[e^{-\lambda\,\nu N'}\right],\qquad \lambda\geq 0.$$

Using f from (2.47) and partial integration, we get

$$\psi(\lambda)=1-\nu\lambda\int_0^\infty f(t)e^{-\lambda\nu t}\,dt=1-\nu(1-\alpha)\lambda\sum_{k=0}^{\infty}\frac{(\alpha\beta)^k}{1+\lambda\nu\beta^k}. \tag{2.54}$$

If $\alpha\beta>1$, we get the asymptotics

$$\alpha^{-\ell}\left(1-\psi\big(\lambda\beta^{-\ell}\big)\right)\overset{\ell\to\infty}{\longrightarrow}\nu(1-\alpha)\lambda\sum_{k=-\infty}^{\infty}\frac{(\alpha\beta)^k}{1+\lambda\nu\beta^k} \tag{2.55}$$

uniformly in $\lambda\in[1,\beta]$. Now let φ be the Laplace transform of T, that is $\varphi(\lambda)=\mathbb{E}[e^{-\lambda T}]$, $\lambda\geq 0$. Using Lemma 2.19 and Proposition 2.17, if $\alpha\beta>1$, it is easy to show that

$$\lim_{\lambda\downarrow 0}\frac{1-\varphi(\lambda)}{1-\psi(\lambda)}=\frac{\beta-1}{\beta}. \tag{2.56}$$

In fact, assume we have two probability measures μ_1 and μ_2 on $[0,\infty)$ and $\xi\in(0,\infty)$ such that

$$\lim_{t\to\infty}\frac{\mu_1((t,\infty))}{\mu_2((t,\infty))}=\xi. \tag{2.57}$$

Denote by $\mathcal{L}_1$ and $\mathcal{L}_2$ the Laplace transforms of μ_1 and μ_2, respectively. Then

$$\frac{1-\mathcal{L}_1(\lambda)}{1-\mathcal{L}_2(\lambda)}=\frac{\int_0^\infty\mu_1((t,\infty))e^{-\lambda t}\,dt}{\int_0^\infty\mu_2((t,\infty))e^{-\lambda t}\,dt}\longrightarrow\xi,\qquad\text{as }\lambda\downarrow 0 \tag{2.58}$$

if μ_1 (and hence μ_2) have infinite first moment, that is, if

$$\int_0^\infty\mu_1((t,\infty))\,dt=\infty. \tag{2.59}$$

Note that the expectation of N' is infinite if and only if $\alpha\beta\geq 1$. Summing up, for $\alpha\beta>1$, we have

$$\alpha^{-\ell}\left(1-\varphi\big(\lambda\beta^{-\ell}\big)\right)\overset{\ell\to\infty}{\longrightarrow}\nu(1-\alpha)\frac{\beta-1}{\beta}\lambda\sum_{k=-\infty}^{\infty}\frac{(\alpha\beta)^k}{1+\lambda\nu\beta^k} \tag{2.60}$$

uniformly in $\lambda\in[1,\beta]$. A similar asymptotics was found already by Solomon [21, Lemma (2.10)(ii)] for a model of random walk in a random environment on $\mathbb{Z}$ with a drift to the left except for geometrically placed reflection points. His asymptotics is the same as ours except for an obvious factor due to the fact that (i) Solomon's "traps" have size at least one while ours start at zero and (ii) our random walk has a positive chance to exit the trap without reaching the bottom.

The usual Tauber theorems that would help to infer the tail behaviour of T from the behaviour of its Laplace transform near zero assume regular variation of the tails (and the

Laplace transforms) which is not the case here. Solomon's proof uses asymptotic equivalence of the Laplace transform ψ to the Laplace transform φ he is interested in, just as we did above. However, this is possible only in the case $\alpha\beta > 1$ which Solomon is mainly concerned with. Our approach of comparing the tails of the approximating random variable N' instead of its Laplace transform allows to deal also with the case $\alpha\beta \leq 1$.

Now we come to determining the asymptotic behavior of $f(t)$ as $t \to \infty$. The following proposition completes the proof of Theorem 1.1.

Let Γ denotes Euler's Γ function. Recall that $\Gamma(a-bi) = \overline{\Gamma(a+bi)}$ for $a, b \in \mathbb{R}$, where the overline indicates the complex conjugate. Also let $\arg(a+bi) \in (-\pi/2, \pi/2)$ denote the angle of $a+bi$ for $a > 0$ and $b \in \mathbb{R}$.

Proposition 2.21 *For all $\gamma > \varrho$, as $t \to \infty$, we have*

$$f(t) = \frac{(1-\alpha)\Gamma(\varrho)}{\log(\beta)} t^{-\varrho} \left[1 + \sum_{\ell \in \mathbb{Z},\, \ell \neq 0} \frac{1}{\Gamma(\varrho)} \Gamma\left(\varrho - \frac{2\pi i\,\ell}{\log(\beta)}\right) \exp\left(2\pi i\,\ell \frac{\log(t)}{\log(\beta)}\right)\right] + O(t^{-\gamma}) \tag{2.61}$$

or equivalently

$$f(t) = \frac{(1-\alpha)\Gamma(\varrho)}{\log(\beta)} t^{-\varrho} \left[1 + \sum_{\ell=1}^{\infty} c_\ell \cos\left(2\pi\ell \frac{\log(t)}{\log(\beta)} - d_\ell\right)\right] + O(t^{-\gamma}) \tag{2.62}$$

with

$$c_\ell = 2\frac{\left|\Gamma\left(\varrho + \frac{2\pi i\,\ell}{\log(\beta)}\right)\right|}{\Gamma(\varrho)} \quad \textit{and} \quad d_\ell = \arg\left(\Gamma\left(\varrho + \frac{2\pi i\,\ell}{\log(\beta)}\right)\right). \tag{2.63}$$

Proof The proof of (2.61) uses Mellin transforms and follows the strategy outlined in [11, Example 12]. We define the Mellin transform of f by

$$f^*(z) := \int_0^\infty t^{z-1} f(t)\, dt, \qquad z \in \mathbb{C},\ \Re(z) \in (0, \rho). \tag{2.64}$$

An explicit computation shows that the integral converges for z in the strip $\Re(z) \in (0, \rho)$ and equals

$$f^*(z) = \frac{\Gamma(z)(1-\alpha)}{1-\alpha\beta^z}. \tag{2.65}$$

That is, f^* is holomorphic for $\Re(z) \in (0, \rho)$ and can be uniquely extended to a meromorphic function in $\mathbb{C}$ with poles in the nonpositive integers and in $\chi_\ell := \varrho + 2\pi i\ell/\log(\beta)$, see Fig. 1. Let

$$\sum_{n=-\infty}^{\infty} a_{\ell,n}(z-\chi_\ell)^n$$

be the Laurent series of $f^*(z)$ around the singularity at χ_ℓ. Then

$$a_{\ell,-1} = -\frac{\Gamma(\chi_\ell)}{\log(\beta)}(1-\alpha). \tag{2.66}$$

Fix an $\eta \in (0, \varrho)$. The inversion formula for Mellin transforms (see [11]) gives

$$f(t) = \frac{1}{2\pi i} \int_{\eta-\infty i}^{\eta+\infty i} f^*(z) t^{-z}\, dz. \tag{2.67}$$

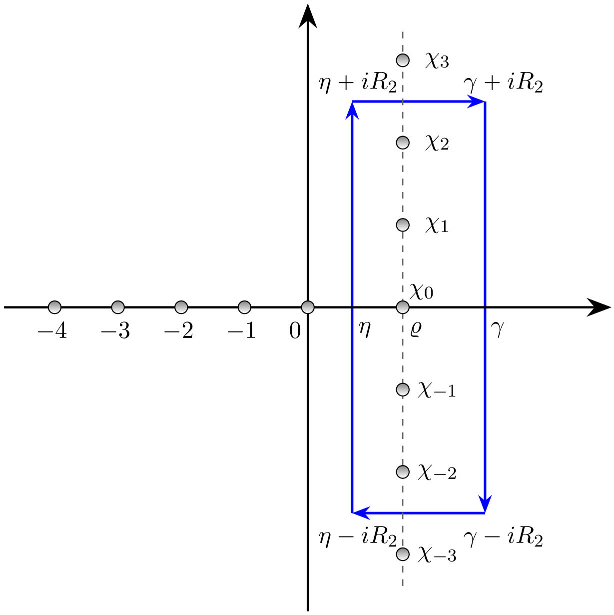

Fig. 1 Complex plane with the singularities of f^* and the integration path

Fix some $\gamma > \varrho$. We can approximate the integral by the finite integrals

$$f(t) = \frac{1}{2\pi i} \int_{\eta - R_\ell i}^{\eta + R_\ell i} f^*(z) t^{-z} \, dz \tag{2.68}$$

where $R_\ell = (2\ell + 1)\pi / \log(\beta)$. We compute this integral using residue calculus for the path consisting of the four pieces $[\eta - R_\ell i, \eta + R_\ell i]$, $[\eta + R_\ell i, \gamma + R_\ell i]$, $[\gamma + R_\ell i, \gamma - R_\ell i]$ and $[\gamma - R_\ell i, \eta - R_\ell i]$. Note that the horizontal paths do not hit the poles and hence the denominator of f^* is bounded away from 0 while the modulus of the Γ function decreases very quickly with ℓ. Thus these integrals can be neglected. The integral along the second vertical piece can be estimated by

$$\left| \int_{\gamma - R_\ell i}^{\gamma + R_\ell i} f^*(z) t^{-z} \, dz \right| \leq t^{-\gamma} \frac{1}{\alpha\beta^\gamma - 1} \int_{-\infty}^{\infty} |\Gamma(\gamma + ir)| \, dr. \tag{2.69}$$

As we integrate clockwise, $f(t)$ is minus the sum of the residues in $(\chi_\ell)_{\ell \in \mathbb{Z}}$ plus the $O(t^{-\gamma})$ term. According to (2.66) these residues are $t^{-\chi_\ell} a_{\ell,-1} = -t^{-\chi_\ell} \Gamma(\chi_\ell) \frac{1-\alpha}{\log(\beta)}$. Concluding, we get (2.61). □

Note that while (2.61) is true for all values of γ, the constant in the term $O(t^{-\gamma})$ in (2.61) is of order $\Gamma(\gamma)$, see (2.69) and thus increases quickly with γ.

Acknowledgements We would like to thank our colleague Duco van Straten from Johannes Gutenberg University Mainz for bringing the Mellin transformation to our attention. We would also like to thank the anonymous referees for their extremely careful reading and their very helpful suggestions.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availibility Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Afanasyev, V.I.: On the maximum of a subcritical branching process in a random environment. Stoch. Process. Appl. **93**(1), 87–107 (2001)
2. Axelson-Fisk, M., Häggström, O.: Conditional percolation on one-dimensional lattices. Adv. Appl. Probab. **41**(4), 1102–1122 (2009)
3. Barma, M., Dhar, D.: Directed diffusion in a percolation network. J. Phys. C **16**(8), 1451 (1983)
4. Ben Arous, G., Fribergh, A.: Biased random walks on random graphs. In: Probability and Statistical Physics in St. Petersburg, Proc. Sympos. Pure Math., vol. 91, pp. 99–153. Amer. Math. Soc., Providence (2016)
5. Ben Arous, G., Fribergh, A., Gantert, N., Hammond, A.: Biased random walks on Galton-Watson trees with leaves. Ann. Probab. **40**(1), 280–338 (2012)
6. Ben Arous, G., Hammond, A.: Randomly biased walks on subcritical trees. Commun. Pure Appl. Math. **65**(11), 1481–1527 (2012)
7. Berger, N., Gantert, N., Peres, Y.: The speed of biased random walk on percolation clusters. Probab. Theory Relat. Fields **126**(2), 221–242 (2003)
8. Biggins, J.D., Bingham, N.H.: Near-constancy phenomena in branching processes. Math. Proc. Camb. Philos. Soc. **110**(3), 545–558 (1991)
9. Biggins, J.D., Nadarajah, S.: Near-constancy of the Harris function in the simple branching process. Commun. Stat. Stoch. Models **9**(3), 435–444 (1993)
10. Dembo, A., Peres, Y., Zeitouni, O.: Tail estimates for one-dimensional random walk in random environment. Commun. Math. Phys. **181**(3), 667–683 (1996)
11. Flajolet, P., Gourdon, X., Dumas, P.: Mellin Transforms and Asymptotics: Harmonic Sums, vol. 144, pp. 3–58. Special Volume on Mathematical Analysis of Algorithms (1995)
12. Fribergh, A., Hammond, A.: Phase transition for the speed of the biased random walk on the supercritical percolation cluster. Commun. Pure Appl. Math. **67**(2), 173–245 (2014)
13. Gantert, N.: Subexponential tail asymptotics for a random walk with randomly placed one-way nodes. Ann. Inst. H. Poincaré Probab. Stat. **38**(1), 1–16 (2002)
14. Gantert, N., Meiners, M., Müller, S.: Regularity of the speed of biased random walk in a one-dimensional percolation model. J. Stat. Phys. **170**(6), 1123–1160 (2018)
15. Gantert, N., Meiners, M., Müller, S.: Einstein relation for random walk in a one-dimensional percolation model. J. Stat. Phys. **176**(4), 737–772 (2019)
16. Hammond, A.: Stable limit laws for randomly biased walks on supercritical trees. Ann. Probab. **41**(3A), 1694–1766 (2013)
17. Kesten, H., Kozlov, M.V., Spitzer, F.: A limit law for random walk in a random environment. Compos. Math. **30**, 145–168 (1975)
18. Klenke, A.: Probability Theory: A Comprehensive Course, 3rd edn. Springer, New York (2020)
19. Lübbers, J.-E., Meiners, M.: The speed of critically biased random walk in a one-dimensional percolation model. Electron. J. Probab. **24**, 1–29 (2019)
20. Lyons, R., Pemantle, R., Peres, Y.: Biased random walks on Galton-Watson trees. Probab. Theory Relat. Fields **106**(2), 249–264 (1996)
21. Solomon, F.: Random walks in a random environment. Ann. Probab. **3**, 1–31 (1975)

22. Sznitman, A.-S.: On the anisotropic walk on the supercritical percolation cluster. Commun. Math. Phys. **240**(1–2), 123–148 (2003)
23. Vatutin, V.A.: Asymptotic behavior of the probability of the first degeneration for branching processes with immigration. Theory Probab. Appl. **19**, 26–35 (1974)