Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 4, 2507–2539
https://doi.org/10.1214/20-AIHP1048

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES

Lower deviation and moderate deviation probabilities for maximum of a branching random walk

Xinxin Chen[a] and Hui He[b]

[a] *Institut Camille Jordan, C.N.R.S. UMR 5208, Universite Claude Bernard Lyon 1, 69622 Villeurbanne Cedex, France. E-mail: xchen@math.univ-lyon1.fr*
[b] *School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China. E-mail: hehui@bnu.edu.cn*

Received 24 September 2018; revised 20 December 2019; accepted 3 February 2020

Abstract. Given a supercritical branching random walk on $\mathbb{R}$ started from the origin, let M_n be the maximal position of individuals at the nth generation. Under some mild conditions, it is proved in (*Ann. Probab.* **41** (2013) 1362–1426) that as $n \to \infty$, $M_n - x^* n + \frac{3}{2\theta^*} \log n$ converges in law for some suitable constants x^* and θ^*. In this work, we investigate its moderate deviation, in other words, the convergence rates of

$$\mathbb{P}\left(M_n \le x^* n - \frac{3}{2\theta^*} \log n - \ell_n\right),$$

for any positive sequence (ℓ_n) such that $\ell_n = O(n)$ and $\ell_n \uparrow \infty$. As a by-product, we obtain lower deviation of M_n; i.e., the convergence rate of $\mathbb{P}(M_n \le xn)$ for $x < x^*$ in Böttcher case where the offspring number is at least two. We also apply our techniques to study the small ball probability of the limit of the so-called derivative martingale. Our results complete those in (*Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 233–260) and (*Electron. Commun. Probab.* **23** (2018) 1–12).

Résumé. Étant donnée une marche aléatoire branchante surcritique sur $\mathbb{R}$ issue de l'origine, on note M_n la position maximale des individus à la n-ème génération. Sous des conditions raisonnables, il a été prouvé dans (*Ann. Probab.* **41** (2013) 1362–1426) que lorsque $n \to \infty$, $M_n - x^* n + \frac{3}{2\theta^*} \log n$ converge en loi pour certaines constantes appropriées x^* et θ^*. Dans cet article, nous envisageons la déviation modérée, autrement dit, les taux de convergence de

$$\mathbb{P}\left(M_n \le x^* n - \frac{3}{2\theta^*} \log n - \ell_n\right),$$

pour toute positive suite (ℓ_n) telle que $\ell_n = O(n)$ et $\ell_n \uparrow \infty$. En particulier, nous obtenons la déviation inférieure de M_n; c'est-à-dire, le taux de convergence de $\mathbb{P}(M_n \le xn)$ avec $x < x^*$ dans le cas Böttcher où le nombre d'enfants est au moins deux. Nous appliquons également ces techniques à l'étude de la petite déviation de la limite de la martingale dérivée. Notre résultats complètent ceux dans (*Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 233–260) et (*Electron. Commun. Probab.* **23** (2018) 1–12).

MSC2020 subject classifications: 60F10; 60J80; 60G50

Keywords: Branching random walk; Maximal position; Moderate deviation; Lower deviation; Schröder case; Böttcher case; Small ball probability; Derivative martingale

1. Introduction

1.1. *Branching random walk and its maximum*

We consider a discrete-time branching random walk on the real line, which, as a generalized branching process, has always been a very attractive objet in probability theory in recent years. It is closely related to many other random models, for example, random walk in random environment, random fractals and discrete Gaussian free field; see [9,26,30,33] and [3] references therein. One can refer to [37] and [38] for the recent developments.

Generally, to construct a branching random walk, we take a random point measure as the reproduction law which describes both the number of children and their displacements. Each individual produces independently its children according to the law of this random point measure. In this way, one develops a branching structure with motions.

In this work, we study a relatively simpler model which is constructed as follows. We take a Galton–Watson tree $\mathcal{T}$, rooted at ρ, with offspring distribution given by $\{p_k; k \geq 0\}$. For any $u, v \in \mathcal{T}$, we write $u \preceq v$ if u is an ancestor of v or $u = v$. Moreover, to each node $v \in \mathcal{T} \setminus \{\rho\}$, we attach a real-valued random variable X_v to represent its displacement. So the position of v is defined by

$$S_v := \sum_{\rho \prec u \preceq v} X_u.$$

Let $S_\rho := 0$ for convenience. Suppose that given the tree $\mathcal{T}$, $\{X_v; v \in \mathcal{T} \setminus \{\rho\}\}$ are i.i.d. copies of some random variable X (which is called displacement or step size). Let $|u|$ denote the generation of an individual $u \in \mathcal{T}$, i.e., the graph distance between v and ρ. Note here that the reproduction law is given by $\sum_{|u|=1} \delta_{X_u}$. Thus, $\{S_u; u \in \mathcal{T}\}$ is our branching random walk with independence between offsprings and motions. This independence will be necessary for our arguments. We will discuss this at the end of this article, in Section 6.

For any $n \in \mathbb{N}$, let M_n be the maximal position at the nth generation, in other words,

$$M_n := \sup_{|v|=n} S_v.$$

The asymptotics of M_n have been studied by many authors, both in the subcritical/critical case and in supercritical case. One can refer to [29,35] and [37] for more details.

We are interested in the supercritical case where $\sum_{k\geq 0} k p_k > 1$ and the system survives with positive probability. Let $(S_n)_{n\geq 0}$ be a random walk started from 0 with i.i.d. increments distributed as X. Observe that for any individual u of the nth generation, S_u is distributed as S_n. If $\mathbb{E}[|X|] < \infty$, classical law of large number tells us that $S_n \sim \mathbb{E}[X]n$ almost surely. However, as there are an exponentially large number of individuals in this supercritical system, the asymptotical behavior of M_n is not as that of S_n.

Conditionally on survival, under some mild conditions, it is known from [5,24,28] that

$$\frac{M_n}{n} \to x^* > \mathbb{E}[X], \quad \text{a.s.,}$$

where x^* is a constant depending on both offspring law and step size. Later, the logarithmic order of $M_n - x^* n$ is given by [1,27] in different ways. Aïdékon in [2] showed that $M_n - x^* n + \frac{3}{2\theta^*} \log n$ converges in law for some suitable $\theta^* \in \mathbb{R}_+^*$, which is an analogue of Bramson's result for branching Brownian motion in [8]; see also [10]. More details on these results will be given in Section 1.2.

For maximum of branching Brownian motion, Chauvin and Rouault [12] first studied the large deviation probability. Recently, Derrida and Shi [15–17] considered both the large deviation and the lower deviation. They established precise estimates. On the other hand, for branching random walks, Hu in [25] studied the moderate deviation for $M_n - x^* n + \frac{3}{2\theta^*} \log n$; i.e.; $\mathbb{P}(M_n \leq x^* n - \frac{3}{2\theta^*} \log n - \ell_n)$ with $\ell_n = o(\log n)$. Later, Gantert and Höfelsauer [22] and Bhattacharya [4] investigated the large deviation probability $\mathbb{P}(M_n \geq xn)$ for $x > x^*$. In the same paper [22], Gantert and Höfelsauer also studied the lower deviation probability $\mathbb{P}(M_n \leq xn)$ for $x < x^*$ mainly in the Schröder case when $p_0 + p_1 > 0$. In fact, branching random walk in the Schröder case can be viewed as a generalized version of branching Brownian motion. Some other related works include Rouault [36] and Buraczewski and Maślanka [11].

Motivated by [22,25] and [16], the goal of this article is to study the moderate deviation $\mathbb{P}(M_n \leq x^* n - \frac{3}{2\theta^*} \log n - \ell_n)$ with $\ell_n = O(n)$. As a by-product of our main results, in the Böttcher case where $p_0 = p_1 = 0$, we also obtain the lower deviation of M_n, i.e., $\mathbb{P}(M_n \leq xn)$ for $x < x^*$, which completes the work [22]. We shall see that the lower deviation of M_n in the Böttcher case turns to be very different from that in the Schröder case. In fact, Gantert and Höfelsauer [22] proved that in the Schröder case $\mathbb{P}(M_n \leq xn)$ decays exponentially. On contrast, in the Böttcher case, we shall show that $\mathbb{P}(M_n \leq xn)$ may decay double-exponentially or sub/super-exponentially depending on the tail behaviors of step size X. We will consider three typical left tail distributions of the step size X and obtain the corresponding decay rates and rate functions. Finally, we also apply our techniques to study the small ball probability for the limit of the so-called derivative martingale. The corresponding problem was also considered in [25] for a class of Mandelbrot's cascades in the Böttcher case with bounded step size and in the Schröder case; see also [31] and [32] for more backgrounds. Let us state the theorems in the following subsection.

In this paper, we use $(c_i)_{i\geq 0}$ and $(C_i)_{i\geq 0}$ to present positive constants. And we write $C(x)$ for constant depending on x. As usual, $f_n = O(g_n)$ or $f_n = O(1)g_n$ means that $|f_n| \leq C g_n$ for all $n \geq 1$. And $f_n = \Theta(1) g_n$ means that f_n is bounded above and below by a positive and finite constant multiple of g_n for all $n \geq 1$. $f_n = o(g_n)$ or $f_n = o_n(1) g_n$ means $\lim_{n\to\infty} \frac{f_n}{g_n} = 0$.

1.2. *Main results*

Suppose that we are in the supercritical case where the tree $\mathcal{T}$ survives with positive probability. Formally, we assume that for the offspring law $\{p_k\}_{k\geq 0}$:

$$m := \sum_{k\geq 0} k p_k \in (1,\infty) \quad \text{and} \quad \sum_{k\geq 0} k^{1+\xi} p_k < \infty, \quad \text{for some } \xi > 0, \tag{1.1}$$

and for the step size X,

$$\mathbb{E}[X] = 0, \quad \text{and} \quad \psi(t) := \mathbb{E}\big[e^{tX}\big] < \infty, \quad \text{for some } t > 0. \tag{1.2}$$

We define the rate function of large deviation for the corresponding random walk $(S_n)_{n\geq 0}$ with i.i.d. step sizes X by

$$I(x) := \sup_{t\in\mathbb{R}} \big\{tx - \log\psi(t)\big\}, \quad \forall x \in \mathbb{R}.$$

Then it is known from Theorem 3.1 in [6] that under (1.1) and (1.2), on the survival set $\{\mathcal{T} = \infty\}$, a.s.,

$$\frac{M_n}{n} \to x^*,$$

where $x^* = \sup\{x \geq 0 : I(x) \leq \log m\} \in (0,\infty)$. Note that if $x^* < \text{ess sup} X \in (0,\infty]$, then $I(x^*) = \log m$ since I is continuous in $(0, \text{ess sup} X)$, and

$$\exists \text{ unique } \theta^* \in (0,\infty) \text{ such that } I\big(x^*\big) = \theta^* x^* - \log\psi\big(\theta^*\big) = \log m. \tag{1.3}$$

According to Theorem 4.1 in [6], it further follows from (1.3) that $\mathbb{P}$-a.s.,

$$M_n - nx^* \to -\infty.$$

This is typical behaviour. But whether this holds or not is the explosion problem for the general branching random walk. One can refer to [6] and [23] for more details. Besides (1.1), (1.2) and (1.3), if we further suppose that

$$\psi(t) < \infty, \quad \forall t \in \big(0, \theta^* + \delta\big) \text{ for some } \delta > 0, \tag{1.4}$$

then it is shown in [1] and [27] that $M_n = m_n + o_{\mathbb{P}}(\log n)$, where

$$m_n := x^* n - \frac{3}{2\theta^*}\log n, \quad \forall n \geq 1.$$

Define the so-called derivative martingale by

$$D_n := \sum_{|u|=n} \theta^*\big(nx^* - S_u\big)e^{\theta^*(S_u - nx^*)}, \quad n \geq 1.$$

It is known from [7] and [2] that under assumptions (1.1), (1.2), (1.3) and (1.4), there exists a non-negative random variable D_∞ such that

$$D_n \overset{\mathbb{P}-a.s.}{\longrightarrow} D_\infty, \quad \text{as } n \to \infty,$$

with $\{D_\infty > 0\} = \{\mathcal{T} = \infty\}$ a.s. Next, assume that

$$\text{the distribution of } X \text{ is non-lattice.} \tag{1.5}$$

Under (1.1), (1.2), (1.3), (1.4) and (1.5), Aïdékon [2] proved the convergence in law of $M_n - m_n$ as follows: for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}(M_n \leq m_n + x) = \mathbb{E}\big[e^{-C_* e^{-x} D_\infty}\big], \tag{1.6}$$

where $C_* > 0$ is a constant. In this work, we are going to study the asymptotic of $\mathbb{P}(M_n \le m_n - \ell_n)$ for $1 \ll \ell_n = O(n)$, as well as that of $\mathbb{P}(0 < D_\infty < \varepsilon)$. Let us introduce the minimal branching/offspring number for $\mathcal{T}$:

$$b := \min\{k \ge 0 : p_k > 0\}.$$

We first present the main results in the Böttcher case where $b \ge 2$.

Theorem 1.1 (Böttcher case, bounded step size: lower deviation). *Assume* (1.1), (1.2) *and* $b \ge 2$. *Suppose that ess inf* $X = -L$ *for some* $0 < L < \infty$, *then for* $x \in (-L, x^*)$,

$$\lim_{n\to\infty} \frac{1}{n} \log\bigl[-\log \mathbb{P}(M_n \le xn)\bigr] = \frac{x^* - x}{x^* + L} \log b. \tag{1.7}$$

If $\mathbb{P}(X = -L) > 0$, *then* (1.7) *holds also for* $x = -L$.

Remark 1.1. Note that the assumptions (1.1) and (1.2) do not imply the second logarithmic order of M_n. For example, if $x^* = \text{ess sup } X \in \mathbb{R}$ and $\mathbb{P}(X = x^*) > \frac{1}{m}$, then with some extra conditions, it is shown in [1] that $\mathbb{E}[M_n - nx^*]$ is bounded.

In the following results, we shall work in the regime that (1.6) holds.

Theorem 1.2 (Böttcher case, bounded step size: moderate deviation). *Assume* (1.1), (1.2), (1.3), (1.4), (1.5) *and* $b \ge 2$. *Suppose that ess inf* $X = -L$ *for some* $0 < L < \infty$. *Then for any positive increasing sequence* ℓ_n *such that* $\ell_n \uparrow \infty$ *and* $\limsup_{n\to\infty} \frac{\ell_n}{n} < x^* + L$,

$$\mathbb{P}(M_n \le m_n - \ell_n) = e^{-e^{\ell_n \beta(1+o_n(1))}}, \tag{1.8}$$

where $\beta := \frac{\log b}{x^* + L} \in (0, \theta^*)$.

Remark 1.2. Hu [25] obtained this moderate deviation (1.8) for $\ell_n = o(\log n)$ in a more general setting with bounded step size and without assuming independence between offsprings and motions. One could check that $\beta = \sup\{a > 0 : \mathbb{P}(\sum_{|u|=1} e^{-a(x^* - X_u)} \ge 1) = 1\} = \frac{\log b}{x^* + L}$ is coherent with that defined in (1.10) of [25].

Remark 1.3. Suppose that all assumptions in Theorem 1.2 hold. Then Theorem 1.3 in [25] gives

$$\mathbb{P}(D_\infty < \varepsilon) = e^{-\varepsilon^{-\frac{\beta}{\theta^* - \beta} + o(1)}}, \quad \text{as } \varepsilon \to 0+.$$

Theorem 1.3 (Böttcher case, Weibull left tail). *Assume* (1.1), (1.2), (1.3), (1.4), (1.5) *and* $b \ge 2$. *Suppose* $\mathbb{P}(X \le -z) = \Theta(1) e^{-\lambda z^\alpha}$ *as* $z \to +\infty$ *for some constant* $\alpha > 0$ *and* $\lambda > 0$. *Then for any positive increasing sequence* ℓ_n *such that* $\ell_n \uparrow \infty$ *and* $\ell_n = O(n)$,

$$\lim_{n\to\infty} \frac{1}{\ell_n^\alpha} \log \mathbb{P}(M_n \le m_n - \ell_n) = -\lambda C(b, \alpha), \tag{1.9}$$

where $C(b, \alpha) := (b^{\frac{1}{\alpha-1}} - 1)^{\alpha-1}$ *for* $\alpha > 1$ *and* $C(b, \alpha) := b$ *for* $\alpha \in (0, 1]$. *In particular, for any* $x < x^*$,

$$\lim_{n\to\infty} \frac{1}{n^\alpha} \log \mathbb{P}(M_n \le xn) = -\lambda C(b, \alpha)\bigl(x^* - x\bigr)^\alpha. \tag{1.10}$$

Remark 1.4. In fact, our arguments in Section 3.2 for $\alpha \in (0, 1]$ also work for X with polynomial left tails. For brevity, we only state Weibull left tail in this theorem.

The weak convergence (1.6) shows the link between $M_n - m_n$ and D_∞. In fact, we will see that $\mathbb{P}(M_n \le m_n - \ell_n)$ and $\mathbb{P}(D_\infty < \varepsilon)$ are closely related, obtained from some similar rare events. So inspired by the proof of Theorem 1.3, we get the following result.

Proposition 1.4 (Böttcher case, Weibull left tail). *Suppose that all assumptions in Theorem* 1.3 *hold. Then*

$$\lim_{\varepsilon\to 0+} \frac{1}{(-\log \varepsilon)^\alpha} \log \mathbb{P}(D_\infty < \varepsilon) = -\frac{\lambda}{(\theta^*)^\alpha} C(b, \alpha). \tag{1.11}$$

The following theorem considers the case of Gumbel left tail.

Theorem 1.5 (Böttcher case, Gumbel left tail). *Assume* (1.1), (1.2), (1.3), (1.4), (1.5) *and* $b \geq 2$. *Suppose* $\mathbb{P}(X \leq -z) = \Theta(1)\exp(-e^{z^\alpha})$ *as* $z \to +\infty$ *for some constant* $\alpha > 0$. *Then for any positive increasing sequence* ℓ_n *such that* $\ell_n \uparrow \infty$ *and* $\ell_n = O(n)$,

$$\lim_{n\to\infty} \ell_n^{-\frac{\alpha}{\alpha+1}} \log\bigl[-\log \mathbb{P}(M_n \leq m_n - \ell_n)\bigr] = \left(\frac{1+\alpha}{\alpha}\log b\right)^{\frac{\alpha}{\alpha+1}}. \tag{1.12}$$

In particular, for any $x < x^*$,

$$\lim_{n\to\infty} n^{-\frac{\alpha}{\alpha+1}} \log\bigl[-\log \mathbb{P}(M_n \leq xn)\bigr] = \left(\frac{1+\alpha}{\alpha}\log b\right)^{\frac{\alpha}{\alpha+1}} \bigl(x^* - x\bigr)^{\frac{\alpha}{\alpha+1}}. \tag{1.13}$$

Again, inspired by the proof of Theorem 1.5, we obtain the following result.

Proposition 1.6 (Böttcher case, Gumbel left tail). *Suppose that all assumptions in Theorem* 1.5 *hold. Then*

$$\lim_{\varepsilon\to 0+} \frac{1}{(-\log\varepsilon)^{\frac{\alpha}{\alpha+1}}} \log\bigl[-\log \mathbb{P}(D_\infty < \varepsilon)\bigr] = \left(\frac{1+\alpha}{\theta^*\alpha}\log b\right)^{\frac{\alpha}{\alpha+1}}. \tag{1.14}$$

Remark 1.5. The assumption $\mathbb{P}(X > z) = \Theta(1)e^{-\lambda z^\alpha}$ (or $\mathbb{P}(X > z) = \Theta(1)e^{-e^{z^\alpha}}$) in Theorem 1.3 and Proposition 1.4 (or in Theorem 1.5 and Proposition 1.6) is made for convenience and the results hold under the assumption that $\mathbb{P}(X > z) = e^{-(\lambda+o(1))z^\alpha}$ (or $\mathbb{P}(X > z) = e^{-e^{(1+o(1))z^\alpha}}$) as $z \to \infty$. The proofs carry through, albeit with some extra epsilons and deltas.

Next theorem concerns the Schröder case where $p_0 + p_1 > 0$. Let $q := \mathbb{P}(\mathcal{T} < \infty) \in [0, 1)$ be the extinction probability and $f(s) := \sum_{k\geq 0} p_k s^k$, $s \in [0, 1]$ be the generating function of its offspring. Let $\mathbb{P}^s(\cdot) := \mathbb{P}(\cdot|\mathcal{T} = \infty)$. Note that (1.6) also holds under $\mathbb{P}^s$. Denote $\max\{a, 0\}$ by a_+ for any real number $a \in \mathbb{R}$. For the step size X, we make a stronger assumption than (1.2):

$$\mathbb{E}[X] = 0, \text{ and there exists } t_0 > 0 \text{ such that } \psi(t) := \mathbb{E}\bigl[e^{tX}\bigr] < \infty, \quad \text{for all } |t| < t_0. \tag{1.2a}$$

If (1.2a) fails, one could also consider $\mathbb{P}(M_n \leq m_n - \ell_n)$. For instance, if we suppose that X has Weibull left tail with $\alpha \in (0, 1)$, i.e., $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ as $z \to \infty$, by use of Theorem 3 in [21], one can show that $\mathbb{P}(M_n \leq m_n - \ell_n) = e^{-(\lambda+o_n(1))\ell_n^\alpha}$. In the following theorem, we work under (1.2a).

Theorem 1.7 (Schröder case). *Assume* (1.1), (1.2*a*), (1.3), (1.4), (1.5) *and* $0 < p_0 + p_1 < 1$. *Then for any positive sequence* (ℓ_n) *such that* $\ell_n \uparrow \infty$ *and that* $\ell^* := \lim_{n\to\infty} \frac{\ell_n}{n}$ *exists with* $\ell^* \in [0, \infty)$, *we have*

$$\lim_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s(M_n \leq m_n - \ell_n) = H\bigl(\ell^*, \gamma\bigr), \tag{1.15}$$

where $\gamma = \log f'(q)$ *and*

$$H\bigl(\ell^*, \gamma\bigr) = \sup_{a \geq \ell^* \vee x^*} \frac{\gamma - I(x^* - a)}{a}. \tag{1.16}$$

In particular, we have for any $x < x^*$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^s(M_n \leq xn) = \bigl(x^* - x\bigr) \sup_{a \leq x \wedge 0} \frac{-I(a) + \gamma}{x^* - a}. \tag{1.17}$$

Remark 1.6. (1.17) was obtained first by Gantert and Höfelsauer in [22]. In fact, it is shown in [22] that for any $x < x^*$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^s(M_n \leq xn) = -\inf_{t\in(0,1]} \bigl\{-t\gamma + tI\bigl(\bigl(x - (1-t)x^*\bigr)/t\bigr)\bigr\}.$$

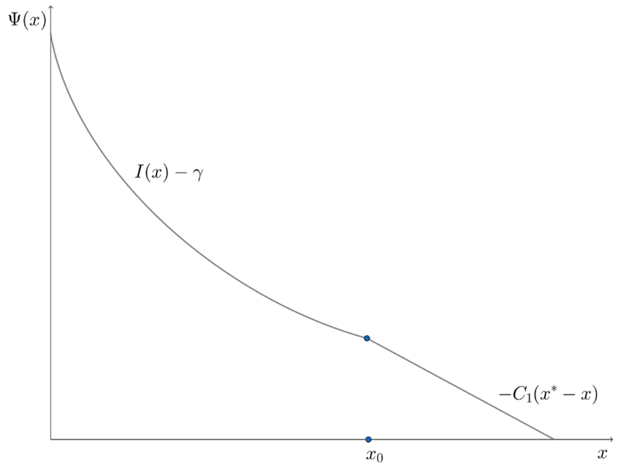

Fig. 1. Ψ may be not analytical at x_0.

Then one can check that

$$-\inf_{t\in(0,1]}\{-t\gamma + tI((x-(1-t)x^*)/t)\} = (x^*-x)\sup_{a\le x\wedge 0}\frac{-I(a)+\gamma}{x^*-a} =: -\Psi(x).$$

Remark 1.7. When $\ell_n = o(\log n)$, (1.15) was obtained by Hu in [25] in a more general framework. In fact, if restricted to our setting, then conditions (1.5) and (1.6) in [25] are equivalent to say that there exists a constant $t^*>0$ such that

$$\log f'(q) + t^*x^* + \log\psi(-t^*) = 0, \quad \text{and} \quad \psi(-t)<\infty \quad \text{for some } t>t^*.$$

Since $\ell_n = o(\log n)$, then $\ell^* = 0$. So conditions (1.5) and (1.6) in [25] make sure that $a^* := x^* - (\log\psi(t))'\,|_{t=-t^*}$ is exactly the arg max of $a\mapsto \frac{\gamma - I(x^*-a)}{a}$ on $[0,\infty)$; i.e.;

$$\frac{\gamma - I(x^*-a^*)}{a^*} = \sup_{a\ge 0}\frac{\gamma - I(x^*-a)}{a} = t^*.$$

Remark 1.8. Note that there is a phase transition for Ψ. When $x\ge \arg\max\{\frac{-I(a)+\gamma}{x^*-a}; a\le x^*\} = (\log\psi(t))'\,|_{t=-t^*} =: x_0$, $(x^*-x)\sup_{a\le x\wedge 0}\frac{-I(a)+\gamma}{x^*-a} = C_1(x^*-x)$ with $C_1 := \sup_{a<x^*}\frac{-I(a)+\gamma}{x^*-a} < 0$; and when $x<(\log\psi(t))'\,|_{t=-t^*}$, $(x^*-x)\sup_{a\le x\wedge 0}\frac{-I(a)+\gamma}{x^*-a} = \gamma - I(x)$. This phenomena has been detected in [17] for branching Brownian motion. See Figure 1.

Remark 1.9. If all assumptions in Theorem 1.7 hold, then by Theorem 1.3 in [25], we have

$$\mathbb{P}(0<D_\infty<\varepsilon)\asymp \varepsilon^{t^*}, \quad \text{as } \varepsilon\to 0+.$$

Remark 1.10. In all the results presented above, we assume that offsprings and motions are independent. Without this assumption of independence, we could find some examples for which the lower deviations are totally different. See Section 6 for more details.

General strategy

Let us explain our main ideas here, especially for $\mathbb{P}(M_n\le m_n-\ell_n)$ in the Böttcher case. Intuitively, to get an unusually low maximum, for the lower bound, we need to control both the size of the genealogical tree and the displacements of individuals. More precisely, we need that at the very beginning, the size of the genealogical tree is small with all individuals moving to some atypically lower place. So, we take some intermediate time/generation t_n and suppose that the genealogical tree is b-regular up to time t_n and that all individuals at time t_n are located below certain "critical" position $-c_n$. Then the system continues with b^{t_n} i.i.d. branching random walks starting from places below $-c_n$. By choosing $c_n=\Theta(\ell_n)$ and t_n in an appropriate way, we can expect that the maximum at time n stays below $m_n-\ell_n$ with high probability. For different assumptions on $\{p_k\}_{k\ge 0}$ and X, the optimal way for moving all b^{t_n} particles to position below $-c_n$ within t_n generation changes from one case to another. If the step size is bounded from below, $t_n=\Theta(\ell_n)$. In the case of Weibull tails with $\alpha \in (0, 1]$, $t_n = 1$. If the step size has Weibull tail of $\alpha > 1$ or Gumbel tail, then $t_n = \Theta(\log \ell_n)$ or $t_n = \Theta(\ell_n^{\frac{\alpha}{\alpha+1}})$, respectively.

The proof of the upper bound goes by showing that the above strategy is optimal, that is that any other strategy would have smaller (or equal) probability. In fact, if most particles at time t_n located above some certain "critical" position $-c_n$, then it would be "very difficult" that the maximum of the system starting from those particles located at higher positions stay below $m_n - \ell_n$ at time n. This would yield the upper bound in the case of bounded step size where there is no particle located below $-Lt_n$ because of ess inf$X = -L$. For other cases, we need to consider the probabilities that there are "some" particles located below $-c_n$. In the Weibull case of $\alpha \in (0, 1]$, this is done by considering b subtrees of height $t_n - 1$ such that for each subtree there is at least one individual located below $-c_n$ with $t_n = \Theta(\ell_n)$. In the Weibull case of $\alpha > 1$, via a tree-transformation, one could extract a b-regular subtree of height s_n such that $s_n \sim t_n$ as $n \to \infty$ and all particles at s_n generations are located below $-c_n$ (reminiscent of that in the lower bound). Once the genealogy is fixed, it remains to optimize for the steps taken by the particles to obtain the upper bound. The Gumbel case is somehow similar to the Weibull case of $\alpha > 1$.

Our arguments and techniques in the Böttcher case are inspired by [13] where we studied the large deviation of the empirical distribution of branching random walk. All these ideas also work for studying the small ball probability of D_∞. The idea to explore the Schröder case is borrowed from [22], where the lower deviation for the Schröder case is given.

The rest of this paper is organized as follows. We treat the Böttcher case with bounded step size in Section 1.2. Then, Section 3 proves Theorems 1.3 and 1.5, concerning the Böttcher cases with unbounded step size. In Section 4, we study $\mathbb{P}(0 < D_\infty < \varepsilon)$ and prove Propositions 1.4 and 1.6. Finally, we prove Theorems 1.7 for Schröder case in Section 5. In Section 6, we discuss a special example.

2. Böttcher case with step size bounded from below

In this section, we always suppose that $b \geq 2$ and ess inf $X = -L$ with $L \in (0, \infty)$. Assumption (1.2) yields $M_n = x^* n + o(n)$ with $x^* \in (0, \infty)$. We are going to prove that for any $-L < x < x^*$,

$$\mathbb{P}(M_n \leq xn) = e^{-e^{(1+o(1))\beta(x^*-x)n}}, \quad \text{as } n \to \infty, \tag{2.1}$$

with $\beta = \frac{\log b}{x^*+L}$. Next, for the second order of M_n, there are several regimes. We assume (1.3), (1.4) and (1.5) to get the classical one: $M_n = m_n + O(1)$ with $m_n = x^* n - \frac{3}{2\theta^*} \log n$. In this regime, we are going to prove that for any positive sequence $\ell_n \uparrow \infty$ such that $\limsup_{n\to\infty} \frac{\ell_n}{n} < x^* + L$,

$$\mathbb{P}(M_n \leq m_n - \ell_n) = e^{-e^{(1+o(1))\beta \ell_n}}, \quad \text{as } n \to \infty. \tag{2.2}$$

The proofs of (2.1) and (2.2) basically follow the same idea. The optimal strategy for the lower bound is that up to some intermediate time t_n, we have a b-regular tree for which all individuals in the t_nth generation are positioned around $-Lt_n$ and then the process behaves typically. For the upper bound, since there are at least b^{t_n} individuals positioned above $-Lt_n$ in the t_nth generation, we see that the following event happens with probability less than $e^{-\Theta(b^{t_n})}$: there are at least b^{t_n} BRWs started from positions above $-Lt_n$ with the maximum of all BRWs at time $n - t_n$ staying below $m_n - \ell_n \approx m_{n-t_n} - Lt_n$. We use t_n^- to denote the intermediate time chosen for the lower bounds and t_n^+ for upper bounds.

For later use, let us introduce the counting measures as follows: for any $B \subset \mathbb{R}$,

$$Z_n(B) := \sum_{|u|=n} \mathbf{1}_{\{S_u \in B\}}, \quad \forall n \geq 0.$$

For simplicity, we write Z_n for $Z_n(\mathbb{R})$ to represent the total population of the nth generation. It is clear that $Z_n \geq b^n$. For any $u \in \mathcal{T}$, let

$$M_n^u := \max_{|z|=n+|u|, u \preceq z} \{S_z - S_u\}, \quad \forall n \geq 0, \tag{2.3}$$

be the maximal relative positions of descendants of u. Clearly, $(M_n^u)_{n\geq 0}$ is distributed as $(M_n)_{n\geq 0}$.

2.1. *Proof of Theorem* 1.1

In this subsection, we are going to prove (2.1) for $x \in (-L, x^*)$.

2.1.1. *Lower bound of Theorem 1.1*
We first prove that for any $x \in (-L, x^*)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log\bigl[-\log \mathbb{P}(M_n \le xn)\bigr] \le \frac{x^* - x}{x^* + L} \log b. \tag{2.4}$$

As $x > -L$, let $L' := L - \eta$ with sufficiently small $\eta > 0$ such that $x > -L + \eta$. Notice that ess inf $X = -L$ implies that $\mathbb{P}(X \le -L') > 0$. For some intermediate time t_n^-, whose value will be determined later, if we let every individual before the t_n^-th generation make a displacement less than $-L'$, then

$$\begin{aligned}\mathbb{P}(M_n \le xn) &\ge \mathbb{P}\bigl(Z_{t_n^-} = b^{t_n^-}; \forall |u| = t_n^-, S_u \le -L' t_n^-; M_n \le xn\bigr) \\ &\ge \mathbb{P}\Bigl(Z_{t_n^-} = b^{t_n^-}; \forall |u| = t_n^-, S_u \le -L' t_n^-; \max_{|u|=t_n^-} M^u_{n-t_n^-} \le xn + L' t_n^-\Bigr),\end{aligned}$$

where $\{M^u_{n-t_n^-}\}$ are i.i.d. copies of $M_{n-t_n^-}$. By branching property at time t_n^-, we arrive at

$$\begin{aligned}\mathbb{P}(M_n \le xn) &\ge \mathbb{P}\bigl(Z_{t_n^-} = b^{t_n^-}; \forall |u| = t_n^-, S_u \le -L' t_n^-\bigr)\mathbb{P}\bigl(M_{n-t_n^-} \le xn + L' t_n^-\bigr)^{b^{t_n^-}} \\ &\ge \mathbb{P}\bigl(Z_{t_n^-} = b^{t_n^-}; \forall 1 \le |u| \le t_n^-, X_u \le -L'\bigr)\mathbb{P}\bigl(M_{n-t_n^-} \le xn + L t_n^-\bigr)^{b^{t_n^-}} \\ &= p_b^{\sum_{k=0}^{t_n^- - 1} b^k} \mathbb{P}\bigl(X \le -L'\bigr)^{\sum_{k=1}^{t_n^-} b^k} \mathbb{P}\bigl(M_{n-t_n^-} \le xn + L' t_n^-\bigr)^{b^{t_n^-}}.\end{aligned} \tag{2.5}$$

Next, we shall estimate $\mathbb{P}(M_{n-t_n^-} \le xn + L' t_n^-)^{b^{t_n^-}}$. The sequel of this proof will be divided into two subparts depending on whether $x^* = R :=$ ess sup X or not, respectively.

Subpart 1: *the case of* $x^* = R$. Note that we have $R < \infty$ now. Take $t_n^- = \lceil \frac{(R-x)n}{R+L'} \rceil$ so that $xn + L' t_n^- \ge R(n - t_n^-)$. Thus,

$$\mathbb{P}\bigl(M_{n-t_n^-} \le xn + L' t_n^-\bigr)^{b^{t_n^-}} = 1.$$

Going back to (2.5), we end up with

$$\mathbb{P}(M_n \le xn) \ge p_b^{\frac{b^{t_n^-} - 1}{b-1}} \bigl(\mathbb{P}\bigl(X \le -L'\bigr)\bigr)^{\frac{b^{t_n^- + 1} - b}{b-1}} \ge e^{-c_1 b^{t_n^-}}. \tag{2.6}$$

It follows readily that for any $x \in (-L, x^*)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log\bigl[-\log \mathbb{P}(M_n \le xn)\bigr] \le \frac{x^* - x}{x^* + L - \eta} \log b. \tag{2.7}$$

Letting $\eta \downarrow 0$ yields (2.4).

Subpart 2: *the case of* $x^* < R \in (0, \infty]$. Now we have $I(x^*) = \log m$ because I is finite and continuous in $(0, R)$. Moreover, $I(x) < \infty$ for some $x > x^*$. For any sufficiently small $a > 0$, one has

$$\log m < I\bigl(x^* + a\bigr) < \infty, \quad \text{and} \quad \lim_{a\downarrow 0} I\bigl(x^* + a\bigr) = I\bigl(x^*\bigr) = \log m.$$

Recall that $-x < L'$. Let $t = \frac{x^*+a-x}{x^*+L'+a}$ and $t_n^- = \lceil tn \rceil$ so that $xn + L' t_n^- > (x^* + a)(n - t_n^-) \gg 1$ for all n large enough. Therefore,

$$\mathbb{P}\bigl(M_{n-t_n^-} \le xn + L' t_n^-\bigr)^{b^{t_n^-}} \ge \bigl(1 - \mathbb{P}\bigl(M_{n-t_n^-} > \bigl(x^* + a\bigr)\bigl(n - t_n^-\bigr)\bigr)\bigr)^{b^{t_n^-}}.$$

By Markov inequality and Chernoff inequality, one has

$$\begin{aligned}\mathbb{P}\big(M_{n-t_n^-} > (x^*+a)(n-t_n^-)\big) &\le \mathbb{P}\big(Z_{n-t_n^-}\big(\big((x^*+a)(n-t_n^-),\infty\big)\ge 1\big)\\ &\le \mathbb{E}[Z_{n-t_n^-}]\mathbb{P}\big(S_{n-t_n^-}\ge (x^*+a)(n-t_n^-)\big)\\ &\le e^{-(I(x^*+a)-\log m)(n-t_n^-)},\end{aligned}$$

which yields

$$\mathbb{P}\big(M_{n-t_n^-}\le xn+L't_n^-\big)^{b^{t_n^-}} \ge \big(1-e^{-(I(x^*+a)-\log m)(n-t_n^-)}\big)^{b^{t_n^-}}.$$

Note that $\log(1-x)\ge -2x$ for any $x\in[0,1/2]$. Let $\delta(a):=I(x^*+a)-\log m$. Then for all sufficiently large $n\ge 1$,

$$\mathbb{P}\big(M_{n-t_n^-}\le xn+L't_n^-\big)^{b^{t_n^-}} \ge e^{-2e^{-\delta(a)(n-t_n^-)}b^{t_n^-}}.$$

Plugging this into (2.5) implies

$$\mathbb{P}(M_n\le xn)\ge e^{-c_1 b^{t_n^-}}e^{-2e^{-\delta(a)(n-t_n^-)}b^{t_n^-}}. \tag{2.8}$$

Thus we have

$$\limsup_{n\to\infty}\frac{1}{n}\log\big[-\log\mathbb{P}(M_n\le xn)\big]\le t\log b. \tag{2.9}$$

Since $I(x^*)=\log m$, letting $a\downarrow 0$ (hence $t\downarrow\frac{x^*-x}{x^*+L'}$ and $\delta(a)\downarrow 0$) gives

$$\limsup_{n\to\infty}\frac{1}{n}\log\big[-\log\mathbb{P}(M_n\le xn)\big]\le\frac{x^*-x}{x^*+L-\eta}\log b,$$

which implies (2.4) because η is arbitrary small. □

2.1.2. *Upper bound of Theorem* 1.1

In this subsection, we show that

$$\liminf_{n\to\infty}\frac{1}{n}\log\big[-\log\mathbb{P}(M_n\le xn)\big]\le\frac{x^*-x}{x^*+L}\log b. \tag{2.10}$$

Note that for any $1\le t_n^+\le n$, $Z_{t_n^+}(\cdot)$ is supported by $[-Lt_n^+,\infty)$ a.s. Since $Z_{t_n^+}\ge b^{t_n^+}$, then

$$\begin{aligned}\mathbb{P}(M_n\le xn) &= \mathbb{P}\big(Z_{t_n^+}\big([-Lt_n^+,\infty)\big)\ge b^{t_n^+};\, M_n\le xn\big)\\ &= \mathbb{P}\Big(Z_{t_n^+}\big([-Lt_n^+,\infty)\big)\ge b^{t_n^+};\, \max_{|u|=t_n^+;S_u\ge -Lt_n^+}\big(S_u+M^u_{n-t_n^+}\big)\le xn\Big)\\ &\le \mathbb{P}\big(M_{n-t_n^+}\le xn+Lt_n^+\big)^{b^{t_n^+}}.\end{aligned} \tag{2.11}$$

It remains to estimate $\mathbb{P}(M_{n-t_n^+}\le xn+Lt_n^+)^{b^{t_n^+}}$. Again, as above, the proof will be divided into two subparts depending on whether $x^*=R$ or not, respectively.

Subpart 1: *the case of* $x^*=R<\infty$. By taking $t_n^+=\lfloor\frac{(R-x)n}{R+L}\rfloor-1$ so that $xn+Lt_n^+<R(n-t_n^+)$, one has

$$\begin{aligned}\mathbb{P}\big(M_{n-t_n^+}\le xn+Lt_n^+\big)^{b^{t_n^+}} &\le \mathbb{P}\big(M_{n-t_n^+}<R(n-t_n^+)\big)^{b^{t_n^+}}\\ &= \big(1-\mathbb{P}\big(M_{n-t_n^+}\ge R(n-t_n^+)\big)\big)^{b^{t_n^+}}\\ &\le \left(1-\frac{c_2}{n-t_n^+}\right)^{b^{t_n^+}}\le e^{-c_2\frac{b^{t_n^+}}{(n-t_n^+)}},\end{aligned} \tag{2.12}$$

where we use the fact that $\mathbb{P}(M_N \geq RN) \geq c_2/N$ for some $c_2 \in (0,1)$ and all $N \geq 1$. In fact, we could construct a Galton–Watson tree with offspring $\sum_{|u|=1} \mathbf{1}_{X_u=R}$. Here $\mathbb{E}[\sum_{|u|=1} \mathbf{1}_{X_u=R}] = m\mathbb{P}(X=R) \geq 1$ since $x^*=R$. Its survival probability is positive if $\mathbb{E}[\sum_{|u|=1} \mathbf{1}_{X_u=R}] > 1$. Even when $\mathbb{E}[\sum_{|u|=1} \mathbf{1}_{X_u=R}] = 1$, it is critical and the survival probability up to generation N is larger than c_2/N for some $c_2>0$ and for all $N\geq 1$. In fact, its survival up to generation N implies that some individual at time N has position RN. So, $\mathbb{P}(M_N \geq RN) \geq c_2/N$. We hence conclude from (2.11) and (2.12) that

$$\liminf_{n\to\infty} \frac{1}{n} \log\bigl[-\log \mathbb{P}(M_n \leq xn)\bigr] \geq \frac{(R-x)\log b}{R+L}.$$

Subpart 2: *the case of* $x^* < R$. First recall Theorem 3.2 in [22] which says that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(M_n > xn) = \log m - I(x), \quad \text{for } x > x^*. \tag{2.13}$$

So for any sufficiently small $a>0$ such that $\delta(a) = I(x^*+a) - \log m \in (0,\infty)$, for any $x > -L$, let $t = \frac{x^*+a-x}{L+x^*+a} \in (0,1)$ and $t_n^+ = \lfloor tn \rfloor$ so that $x^* < \frac{xn+Lt_n^+}{n-t_n^+} \leq x^*+a$. Then for all n large enough,

$$\begin{aligned}
\mathbb{P}\bigl(M_{n-t_n^+} \leq xn + Lt_n^+\bigr)^{b^{t_n^+}} &= \biggl(1 - \mathbb{P}\biggl(M_{n-t_n^+} > \frac{xn+Lt_n^+}{n-t_n^+}\bigl(n-t_n^+\bigr)\biggr)\biggr)^{b^{t_n^+}} \\
&\leq \bigl(1 - \mathbb{P}\bigl(M_{n-t_n^+} > \bigl(x^*+a\bigr)\bigl(n-t_n^+\bigr)\bigr)\bigr)^{b^{t_n^+}} \\
&\leq \bigl(1 - \exp\bigl\{-\bigl(I\bigl(x^*+a\bigr) - \log m + \delta(a)\bigr)\bigl(n-t_n^+\bigr)\bigr\}\bigr)^{b^{t_n^+}} \\
&\leq e^{-e^{-2\delta(a)(n-t_n^+)} b^{t_n^+}},
\end{aligned} \tag{2.14}$$

where the second inequality follows from (2.13). Plugging (2.14) into (2.11) yields

$$\liminf_{n\to\infty} \frac{1}{n} \log\bigl[-\log \mathbb{P}(M_n \leq xn)\bigr] \geq -2\delta(a)(1-t) + t\log b.$$

Again letting $a \downarrow 0$ (hence $\delta(a) \downarrow 0$ and $t \downarrow \frac{x^*-x}{x^*+L}$) gives the desired upper bound (2.10).

If $\mathbb{P}(X=-L) > 0$, then the arguments for lower bound work well for $x=-L$ and $L'=L$. For the upper bound, it is easy to see that all displacements are $-L$ up to the nth generation. We thus could also obtain (1.7) for $x=-L$. □

2.2. *Proof of Theorem* 1.2

From now on, we further assume (1.3), (1.4) and (1.5) so that (1.6) holds. In fact, (1.4) is slightly stronger than the conditions given in [2]. Because of this convergence in law in Böttcher case, we can find some $y^* \in \mathbb{R}_+$ so that

$$\mathbb{P}\bigl(M_n \leq m_n - y^*\bigr) \leq 1/2 \leq \mathbb{P}\bigl(M_n \leq m_n + y^*\bigr). \tag{2.15}$$

.

Now we are ready to prove that for any increasing sequence $\ell_n = O(n)$ such that $\ell_n \uparrow \infty$ and $\limsup_{n\to\infty} \frac{\ell_n}{n} < x^*+L$,

$$\mathbb{P}(M_n \leq m_n - \ell_n) = e^{-e^{\ell_n \beta(1+o_n(1))}}. \tag{2.16}$$

Recall that $m_n = x^*n - \frac{3}{2\theta^*}\log n$ and $\beta = \frac{\log b}{x^*+L}$. Notice that $\beta \in (0,\theta^*)$ because of (1.3).

2.2.1. *Lower bound of Theorem* 1.2

Similarly to the arguments in the Section 2.1.1, for some intermediate time $t_n^- \in [1, n-1]$ and $L' = L - \eta$ with $\eta > 0$, we have

$$\begin{aligned}
\mathbb{P}(M_n \leq m_n - \ell_n) &\geq \mathbb{P}\bigl(Z_{t_n^-} = b^{t_n^-}; \forall |u| \leq t_n^-, X_u \leq -L'; M_n \leq m_n - \ell_n\bigr) \\
&\geq \mathbb{P}\Bigl(Z_{t_n^-} = b^{t_n^-}; \forall |u| \leq t_n^-, X_u \leq -L'; \max_{|v|=t_n^-} M^v_{n-t_n^-} \leq m_n - \ell_n + L't_n^-\Bigr),
\end{aligned}$$

which by branching property is larger than

$$\mathbb{P}\big(Z_{t_n^-} = b^{t_n^-}; \forall |u| \le t_n^-, X_u \le -L'\big)\mathbb{P}\big(M_{n-t_n^-} \le m_n - \ell_n + L' t_n^-\big)^{b^{t_n^-}}.$$

Here we choose $t_n^- = \lceil \frac{\ell_n + K_0}{L' + x^*} \rceil$ with $K_0 \ge 1$ a fixed large constant so that $m_n - \ell_n + L' t_n^- \ge m_{n-t_n^-} + y^*$. Consequently,

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\ge \mathbb{P}\big(Z_{t_n^-} = b^{t_n^-}; \forall |u| \le t_n^-, X_u \le -L'\big)\mathbb{P}\big(M_{n-t_n^-} \le m_{n-t_n^-} + y^*\big)^{b^{t_n^-}} \\ &\ge p_b^{\sum_{k=0}^{t_n^- -1} b^k} \mathbb{P}\big(X \le -L'\big)^{\sum_{k=1}^{t_n^-} b^k} \mathbb{P}\big(M_{n-t_n^-} \le m_{n-t_n^-} + y^*\big)^{b^{t_n^-}},\end{aligned}$$

where the last inequality holds because of the independence between offsprings and motions. Now note that $-L =$ ess inf X means $q_L := \mathbb{P}(X \le -L') \in (0, 1)$. By (2.15),

$$\mathbb{P}(M_n \le m_n - \ell_n) \ge p_b^{\sum_{k=0}^{t_n^- -1} b^k} q_L^{\sum_{k=1}^{t_n^-} b^k} (1/2)^{b^{t_n^-}} \ge e^{-c_3 b^{t_n^-}},$$

with $t_n^- = \lceil \frac{\ell_n + K_0}{L + x^* - \eta} \rceil$. Letting $n \to \infty$ then $\eta \to 0$ yields that

$$\limsup_{n\to\infty} \frac{1}{\ell_n} \log\big[-\log \mathbb{P}(M_n \le m_n - \ell_n)\big] \le \frac{\log b}{L + x^*}.$$

2.2.2. *Upper bound of Theorem 1.2*

Similarly as in the Section 2.10, for some intermediate time t_n^+ to be determined later, we have

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &= \mathbb{P}\big(Z_{t_n^+}\big([-Lt_n^+, \infty)\big) \ge b^{t_n^+}; M_n \le m_n - \ell_n\big) \\ &= \mathbb{P}\Big(Z_{t_n^+}\big([-Lt_n^+, \infty)\big) \ge b^{t_n^+}; \max_{|u| = t_n^+, S_u \ge -Lt_n^+} \big(S_u + M^u_{n-t_n^+}\big) \le m_n - \ell_n\Big) \\ &\le \mathbb{P}\big(M_{n-t_n^+} \le m_n - \ell_n + Lt_n^+\big)^{b^{t_n^+}}.\end{aligned}$$

Let $t_n^+ := \lfloor \frac{\ell_n - y^*}{L + x^*} \rfloor$ so that $m_n - \ell_n + Lt_n^+ \le m_{n-t_n^+} - y^*$. By (2.15),

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\le \mathbb{P}\big(M_{n-t_n^+} \le m_{n-t_n^+} - y^*\big)^{b^{t_n^+}} \\ &\le (1/2)^{b^{t_n^+}}. \end{aligned} \tag{2.17}$$

We hence obtain

$$\mathbb{P}(M_n \le m_n - \ell_n) \le e^{-c_4 b^{t_n^+}},$$

with $b^{t_n^+} = \Theta(e^{\beta \ell_n})$. This is sufficient to conclude Theorem 1.2.

3. Böttcher case with step size of sub/super-exponential left tail

We prove Theorem 1.3 and 1.5 in this section. We first state some preliminary lemmas in the Section 3.1. There are mainly two cases in the proofs of the theorems. When the left tail of X is sub-exponential, i.e., X has Weibull tail of $\alpha \in (0, 1]$, the proof is relatively simple and is stated in the Section 3.2. When the left tail of X is super-exponential, i.e., X has Weibull tail of $\alpha > 1$ or Gumbel tail, the proofs are given in the Sections 3.3 and 3.4 where we need a tree transformation introduced in the Section 3.3.1 to get the upper bound.

3.1. *Some preliminary results*

In this subsection, we state some technical lemmas which will be useful later.

3.1.1. *Some large deviation probability estimates*
We first state some results on the random walk $(S_n)_{n\geq 0}$ with i.i.d. increments distributed as X and started from $S_0 := 0$. Recall that $\psi(t) = \log \mathbb{E}[e^{tX}]$.

Lemma 3.1. *For positive integer sequence (a_n) such that $a_n = o(n)$,*

1. *if $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha \in (0, 1]$, $\lambda > 0$ and $a_n = o(n^\alpha)$, then for any $\varepsilon \in (0, 1)$ and for $n \gg 1$ sufficiently large,*

$$\mathbb{P}(S_{a_n} \leq -n) < e^{-\lambda(1-\varepsilon)n^\alpha}; \tag{3.1}$$

2. *if $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha > 1$ and $\lambda > 0$, then for any $\theta > 0$ and for $n \gg 1$ sufficiently large,*

$$\mathbb{P}(S_{a_n} \leq -n) < e^{-\theta n}; \tag{3.2}$$

3. *if $\mathbb{P}(X < -z) = \Theta(1)e^{-e^{z^\alpha}}$ with $\alpha > 0$, then for any $\varepsilon \in (0, 1)$ and for $n \gg 1$ sufficiently large,*

$$\mathbb{P}(S_{a_n} \leq -n) < c_5 a_n e^{-e^{(\frac{n}{a_n})^\alpha}}. \tag{3.3}$$

Proof. *Proof of* (3.1) Note that

$$\mathbb{P}(S_{a_n} < -n) \leq \mathbb{P}\left(\sum_{k=1}^{a_n}(X_k \wedge 0) < -n\right) \leq \mathbb{P}\left(\sum_{k=1}^{n^\alpha}(X_k \wedge 0) < -n\right).$$

Then an application of Theorem 3 in [21] and the remark after its proof by taking $L(n) = \lambda$, $\psi(n) = n^{1/\alpha}$ yields (3.1).

Proof of (3.2) In fact, for any $\theta > 0$, by Markov inequality,

$$\mathbb{P}(S_{a_n} < -n) \leq \mathbb{P}\left(e^{-\theta S_{a_n}} < e^{\theta n}\right) \leq e^{-\theta n + \psi(\theta)a_n},$$

where $\psi(\theta) = \log \mathbb{E}[e^{\theta X}] < \infty$ for any $\theta > 0$ as $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha > 1$. Moreover, as $a_n = o(n)$, for $n \gg 1$ sufficiently large, one gets (3.2).

Proof of (3.3) Note that $S_{a_n} \leq -n$ implies that $\min_{1\leq k\leq a_n} X_k \leq -n/a_n$. Therefore, for n large enough,

$$\begin{aligned}\mathbb{P}(S_{a_n} < -n) &\leq \mathbb{P}\Big(\min_{1\leq k\leq a_n} X_k < -n/a_n\Big) \\ &\leq a_n \mathbb{P}(X < -n/a_n) \leq c_5 a_n e^{-e^{(n/a_n)^\alpha}},\end{aligned}$$

as $\mathbb{P}(X < -z) = \Theta(1)e^{-e^{z^\alpha}}$. □

3.1.2. *Rough upper bounds*
Now we are ready to get a rough upper bound for $\mathbb{P}(M_n \leq m_n - \ell_n)$. The idea behind this rough upper bound is that: either a particle is very low at some intermediate time t_n, or at least b^{t_n} particles at time t_n have to have a very small maixmum in their subtree.

Lemma 3.2. *Assume* (1.1), (1.2), (1.3), (1.4), (1.5) *and $b \geq 2$. For positive sequence (ℓ_n) such that $\ell_n \uparrow \infty$ and $\ell_n = O(n)$, we have the following inequalities.*

1. *If $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha \in (0, 1]$ and $\lambda > 0$, then for any $\varepsilon \in (0, 1)$ and for all n sufficiently large,*

$$\mathbb{P}(M_n \leq m_n - \ell_n) \leq e^{-\lambda(1-\varepsilon)\ell_n^\alpha}. \tag{3.4}$$

2. *If $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha > 1$ and $\lambda > 0$, then for any $\theta > 0$ and for all n sufficiently large,*

$$\mathbb{P}(M_n \leq m_n - \ell_n) \leq e^{-\theta \ell_n}. \tag{3.5}$$

3. *If $\mathbb{P}(X < -z) = \Theta(1)e^{-e^{z^\alpha}}$ with $\alpha > 0$ as $z \to \infty$, then there exist $c_6, c_7 > 0$ such that for all n large enough,*

$$\mathbb{P}(M_n \leq m_n - \ell_n) \leq c_6 \exp\left(-e^{c_7 \ell_n^{\frac{\alpha}{\alpha+1}}}\right). \tag{3.6}$$

Proof. Generally, to bound $\mathbb{P}(M_n \le m_n - \ell_n)$, we take some intermediate time $t_n = o(\ell_n)$ which will be chosen later and let $B_n := [-(1-\varepsilon)\ell_n, \infty)$ with $\varepsilon \in (0,1)$. As $Z_{t_n} \ge b^{t_n}$,

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\le \mathbb{P}\big(Z_{t_n}(B_n) \ge b^{t_n}; M_n \le m_n - \ell_n\big) + \mathbb{P}\big(Z_{t_n}(B_n) < b^{t_n}\big) \\ &= \mathbb{P}\big(Z_{t_n}(B_n) \ge b^{t_n}; M_n \le m_n - \ell_n\big) + \mathbb{P}\big(Z_{t_n}\big(B_n^c\big) \ge 1\big). \end{aligned} \tag{3.7}$$

For $t_n = o(\ell_n)$ and y^* chosen in (2.15), we have $m_n - \varepsilon\ell_n \le m_{n-t_n} - y^*$ for all n large enough. Therefore,

$$\begin{aligned}\mathbb{P}\big(Z_{t_n}(B_n) \ge b^{t_n}; M_n \le m_n - \ell_n\big) &\le \mathbb{P}\Big(Z_{t_n}(B_n) \ge b^{t_n}; \max_{|u|=t_n, S_u \in B_n} \big(S_u + M^u_{n-t_n}\big) \le m_n - \ell_n\Big) \\ &\le \mathbb{P}\Big(Z_{t_n}(B_n) \ge b^{t_n}; \max_{|u|=t_n, S_u \in B_n} M^u_{n-t_n} \le m_n - \varepsilon\ell_n\Big) \\ &\le \mathbb{P}\Big(Z_{t_n}(B_n) \ge b^{t_n}; \max_{|u|=t_n, S_u \in B_n} M^u_{n-t_n} \le m_{n-t_n} - y^*\Big). \end{aligned}$$

By Markov property at time t_n, all $M^u_{n-t_n}$ are i.i.d. copies of M_{n-t_n} for $|u| = t_n$, and independent of $(S_u, |u| = t_n)$. This implies that

$$\mathbb{P}\big(Z_{t_n}(B_n) \ge b^{t_n}; M_n \le m_n - \ell_n\big) \le \mathbb{P}\big(M_{n-t_n} \le m_{n-t_n} - y^*\big)^{b^{t_n}} \le (1/2)^{b^{t_n}}. \tag{3.8}$$

Plugging it into (3.7) yields that

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\le (1/2)^{b^{t_n}} + \mathbb{P}\big(Z_{t_n}\big(B_n^c\big) \ge 1\big) \\ &\le (1/2)^{b^{t_n}} + \mathbb{E}\big[Z_{t_n}\big(B_n^c\big)\big] \\ &= (1/2)^{b^{t_n}} + m^{t_n}\mathbb{P}\big(S_{t_n} \le -(1-\varepsilon)\ell_n\big). \end{aligned} \tag{3.9}$$

We shall apply Lemma 3.1 to bound $\mathbb{P}(S_{t_n} \le -(1-\varepsilon)\ell_n)$.

Proof of (3.4) If $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha \in (0,1]$, we choose $t_n = \lfloor \frac{2\log \ell_n}{\log b} \rfloor$ and use (3.1) to get that

$$\mathbb{P}(M_n \le m_n - \ell_n) \le (1/2)^{b^{t_n}} + m^{t_n} e^{-\lambda(1-\varepsilon)^{1+\alpha}\ell_n^\alpha},$$

which suffices to conclude (3.4) since $b^{t_n} \ge \ell_n^2$ and $t_n \ll \ell_n$.

Proof of (3.5) Similarly as above, we choose $t_n = \lfloor \frac{2\log \ell_n}{\log b} \rfloor$. Then by (3.2) for $\alpha > 1$ we obtain that for any $\theta > 0$ and for n large enough,

$$\mathbb{P}(M_n \le m_n - \ell_n) \le (1/2)^{b^{t_n}} + m^{t_n} e^{-\theta \ell_n (1-\varepsilon)},$$

which suffices to conclude (3.5).

Proof of (3.6) If $\mathbb{P}(X < -z) = \Theta(1)e^{-e^{z^\alpha}}$ with $\alpha > 0$ as $z \to \infty$, we choose $t_n = \lfloor \frac{\ell_n^{\frac{\alpha}{\alpha+1}}}{\log b} \rfloor$ and $0 < \varepsilon < 1/2$. Then by (3.3) we get

$$\mathbb{P}(M_n \le m_n - \ell_n) \le (1/2)^{b^{t_n}} + c_5 t_n m^{t_n} \exp\big\{-e^{(1-\varepsilon)^\alpha(\ell_n/t_n)^\alpha}\big\} \le c_8 \exp\big\{-e^{c_9 \ell_n^{\alpha/(\alpha+1)}}\big\},$$

for some constants $c_8, c_9 > 0$. □

3.2. *Proof for Theorem* 1.3: *Step size of Weibull tail with* $\alpha \le 1$

Let us explain the ideas before stating the proof. For the lower bound, the optimal strategy is to produce exactly b children at the first generation and to make each of them move to some position below $-\ell_n$. For the upper bound, we consider the individuals at time $t_n = \Theta(\log \ell_n)$, either there are at least b^{t_n-1} individuals located above $-\ell_n$ whose descendants at time n can hardly go below $m_n - \ell_n$, or there are at most $b^{t_n-1} - 1$ individuals located above $-\ell_n$. In the latter case, there exists no individual at the first generation whose descendants are all located above $-\ell_n$ at time t_n, i.e., we could find b i.i.d. random walks moving below $-\ell_n$ in time t_n. And the probability of this behaviour is comparable with the lower bound.

Lower bound
We shall show that if $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $0 < \alpha \le 1$, as $z \to +\infty$, then for all n sufficiently large,

$$\mathbb{P}(M_n \le m_n - \ell_n) \ge c_{10} e^{-\lambda \ell_n^\alpha b}.$$

Recall y^* from (2.15). In fact, at the first generation, we suppose that there are exactly b individuals and that all of them are located below $-(\ell_n + x^* + y^*)$. So, as $m_n - \ell_n + (\ell_n + x^* + y^*) \ge m_{n-1} + y^*$,

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\ge \mathbb{P}\big(Z_1 = b; \forall |u| = 1, X_u \le -\big(\ell_n + x^* + y^*\big); M_n \le m_n - \ell_n\big) \\ &= \mathbb{P}\Big(Z_1 = b; \forall |u| = 1, X_u \le -\big(\ell_n + x^* + y^*\big); \max_{|u|=1}\big(X_u + M_{n-1}^u\big) \le m_n - \ell_n\Big) \\ &\ge \mathbb{P}\Big(Z_1 = b; \forall |u| = 1, X_u \le -\big(\ell_n + x^* + y^*\big); \max_{|u|=1} M_{n-1}^u \le m_{n-1} + y^*\Big).\end{aligned}$$

By branching property, this implies

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\ge \mathbb{P}\big(Z_1 = b; \forall |u| = 1, X_u \le -\big(\ell_n + x^* + y^*\big)\big)\mathbb{P}\big(M_{n-1} \le m_{n-1} + y^*\big)^b \\ &= p_b \mathbb{P}\big(X \le -\big(\ell_n + x^* + y^*\big)\big)^b \mathbb{P}\big(M_{n-1} \le m_{n-1} + y^*\big)^b.\end{aligned}$$

Recall that $\mathbb{P}(X \le -(\ell_n + x^* + y^*)) \ge c_{11} e^{-\lambda \ell_n^\alpha}$ for $\ell_n \gg 1$ and $\mathbb{P}(M_{n-1} \le m_n + y^*) \ge 1/2$ by (2.15). So,

$$\mathbb{P}(M_n \le m_n - \ell_n) \ge c_{10} e^{-\lambda \ell_n^\alpha b},$$

with $c_{10} = p_b (c_{11}/2)^b > 0$.

Upper bound
Take some intermediate time $t_n = \lfloor \frac{\log \ell_n}{\log b} \rfloor$. For $B_n = [-(1-\varepsilon)\ell_n, \infty)$ with arbitrary small $\varepsilon \in (0,1)$ and for any $\delta \in (0, 1/b)$, one sees that

$$\begin{aligned}\mathbb{P}(M_n \le m_n - \ell_n) &\le \mathbb{P}\big(Z_{t_n}(B_n) \ge \delta b^{t_n}; M_n \le m_n - \ell_n\big) + \mathbb{P}\big(Z_{t_n}(B_n) < \delta b^{t_n}\big) \\ &\le \mathbb{P}\Big(Z_{t_n}(B_n) \ge \delta b^{t_n}; \max_{|u|=t_n, S_u \in B_n}\big(S_u + M_{n-t_n}^u\big) \le m_n - \ell_n\Big) + \mathbb{P}\big(Z_{t_n}(B_n) < \delta b^{t_n}\big).\end{aligned} \tag{3.10}$$

For n large enough so that $m_n - \ell_n + (1-\varepsilon)\ell_n \le m_{n-t_n} - \varepsilon \ell_n/2$, by branching property,

$$\begin{aligned}&\mathbb{P}\Big(Z_{t_n}(B_n) \ge \delta b^{t_n}; \max_{|u|=t_n, S_u \in B_n}\big(S_u + M_{n-t_n}^u\big) \le m_n - \ell_n\Big) \\ &\quad \le \mathbb{P}\Big(Z_{t_n}(B_n) \ge \delta b^{t_n}; \max_{|u|=t_n, S_u \in B_n} M_{n-t_n}^u \le m_{n-t_n} - \varepsilon \ell_n/2\Big) \\ &\quad \le \mathbb{P}(M_{n-t_n} \le m_{n-t_n} - \varepsilon \ell_n/2)^{\delta b^{t_n}}.\end{aligned}$$

It then follows from (3.4) that

$$\mathbb{P}\Big(Z_{t_n}(B_n) \ge \delta b^{t_n}; \max_{|u|=t_n, S_u \in B_n}\big(S_u + M_{n-t_n}^u\big) \le m_n - \ell_n\Big) \le e^{-\lambda \varepsilon \ell_n/8 \times \delta b^{t_n}}. \tag{3.11}$$

On the other hand, since $\delta < 1/b$, the event $Z_{t_n}(B_n) < \delta b^{t_n}$ implies that for any $|v| = 1$, $\{|u| = t_n : u \succ v\} \not\subset \{|u| = t_n, S_u \in B_n\}$. This means that

$$\begin{aligned}\mathbb{P}\big(Z_{t_n}(B_n) < \delta b^{t_n}\big) &\le \mathbb{P}\Big(\bigcap_{|v|=1} \cup_{|u|=t_n, u \succ v}\big\{S_u \in B_n^c\big\}\Big) \\ &\le \mathbb{E}\Big[\mathbb{P}\Big(\bigcup_{|u|=t_n, u \succ v}\big\{S_u \in B_n^c\big\}\Big)^{Z_1}\Big] \\ &\le \mathbb{E}\Big[\Big(\mathbb{E}\Big(\sum_{|u|=t_n, u \succ v} \mathbf{1}_{\{S_u \in B_n^c\}} \big| |v| = 1\Big)\Big)^b\Big],\end{aligned}$$

where the last inequality follows from the fact that $Z_1 \geq b$ and Markov inequality. By independence between offsprings and motions, this leads to

$$\begin{aligned}\mathbb{P}\big(Z_{t_n}(B_n) < \delta b^{t_n}\big) &\leq \big(\mathbb{E}[Z_{t_n-1}]\mathbb{P}\{S_{t_n} \in B_n^c\}\big)^b \\ &= \big(m^{t_n-1}\mathbb{P}\{S_{t_n} < -(1-\varepsilon)\ell_n\}\big)^b.\end{aligned}$$

By (3.1), one gets that for $n \gg 1$,

$$\mathbb{P}\big(Z_{t_n}(B_n) < \delta b^{t_n}\big) \leq c_{12} m^{b(t_n-1)} e^{-\lambda(1-\varepsilon)^{1+\alpha} b \ell_n^\alpha}. \tag{3.12}$$

Plugging it and (3.11) into (3.10) yields that for all $n \gg 1$ large enough,

$$\mathbb{P}(M_n \leq m_n - \ell_n) \leq e^{-\lambda\varepsilon\ell_n/8\times\delta b^{t_n}} + c_{12} m^{b(t_n-1)} e^{-\lambda(1-\varepsilon)^{1+\alpha} b \ell_n^\alpha}.$$

According to the choice of $t_n = \lfloor \frac{\log \ell_n}{\log b} \rfloor$, we could conclude that for arbitrary small $\varepsilon > 0$,

$$\limsup_{n\to\infty} \frac{1}{\ell_n^\alpha} \log \mathbb{P}(M_n \leq m_n - \ell_n) \leq -\lambda(1-\varepsilon)^{1+\alpha} b,$$

which gives the upper bound for the case of $\alpha \in (0,1]$ by letting $\varepsilon \downarrow 0+$.

3.3. *Proof for Theorem 1.3: Step size of Weibull tail with $\alpha > 1$*

Here, we are going to use some ideas different from that in the Section 3.2. Observe that we need to control every branch to get an atypical lower maximum at time n. The optimal choice is to control the motions at the beginning when there are not so many individuals. So we need to have a b-regular tree up to some intermediate time $t_n = \Theta(\log \ell_n)$ and to make sure that each individual in this generation moves to some position below $-\ell_n$. As just modifying a few displacements near the root can have a significant effect on the maximum, we force the individuals of earlier generations to make larger displacements. This helps us to get the lower bound.

For the upper bound, one sees that if there are more than $b^{t_n-\delta_n}$ individuals positioned above $-\ell_n$ at some time $t_n = \Theta(\log \ell_n)$, it will be extremely difficult to make maximal position at time n less than $m_n - \ell_n$. So we should have more than $b^{t_n} - b^{t_n-\delta_n}$ individuals positioned below $-\ell_n$. By the tree transformation introduced below, we could get a b-regular tree up to generation t_n for which one individual in the δ_nth generation with all its descendants is removed. For this almost b-regular tree, the event that all individuals at time t_n stay below $-\ell_n$ is reminiscent of that in the lower bound.

3.3.1. *A technical lemma and tree transformation*

Here we introduce a technical lemma and the associated tree transformation in the deterministic setting which will be used later for the upper bound estimates with super-exponential tails.

Let $\mathbf{t}$ be a fixed tree of H generations such that every individual $u \in \mathbf{t}$ with $|u| \leq H-1$ has at least b offsprings. Mark each $u \in \mathbf{t} \setminus \{\rho\}$ with $x_u \in \mathbb{R}$ and mark ρ with $x_\rho = 0$. Set $s_u := \sum_{\rho \prec v \preceq u} x_v$ with $s_\rho := 0$. Then we regard s_u as the spatial position of u. Assume that in the Hth generation, there are at most b^{H-K} individuals located below a given level $\ell > 0$. Here $H, K \in \mathbb{N}$ and $\ell \in \mathbb{R}_+$.

Lemma 3.3. *Let $\alpha > 0$. For any $C_2 \geq 0$ and $\varepsilon \in (0,1)$, let $M \geq 1$ be a large fixed integer such that*

$$\sum_{x=M}^{\infty} e^{-\lambda\varepsilon x^\alpha + C_2} < 1.$$

Then for $\alpha > 1$, $\lambda > 0$ and all ℓ sufficiently large,

$$\begin{aligned}\Sigma_{\text{Weibull}}(\mathbf{t}) :&= \sum_{(x_u)_{u\in\mathbf{t}\setminus\{\rho\}} \in (\mathbb{N}\cap[M,\infty))^{\#\mathbf{t}-1}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\} \mathbf{1}_{\{\sum_{|u|=H} \mathbf{1}_{\{s_u\leq\ell\}} \leq b^{H-K}\}} \\ &\leq \exp\big\{-\lambda(1-\varepsilon)C(b,\alpha)\big(1-b^{-K}\big)^{\alpha+1}\ell^\alpha\big\},\end{aligned} \tag{3.13}$$

where $C(b,\alpha)$ is defined in Theorem 1.3.

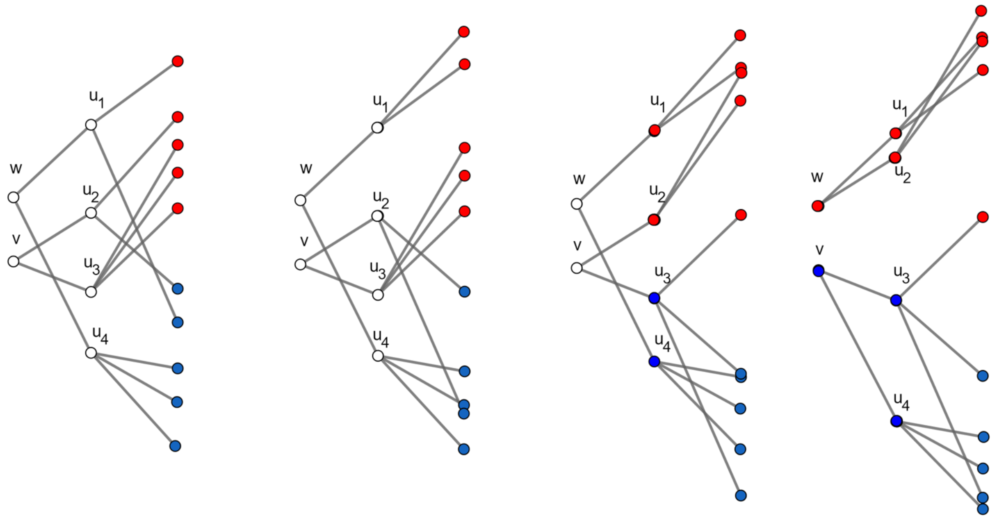

Fig. 2. We first exchange u_1's blue child with u_2's red child; then we exchange u_2's blue children with two of u_3's red children. So we color u_1 and u_2 red and color u_3 and u_4 blue (notice that one of u_3's children is red). Next, we exchange u_2 and its subtree with u_4 and its subtree. Then w is colored red and v is colored blue.

Similarly, for any $C_2 \geq 0$ and $\varepsilon \in (0,1)$, take $M \geq 1$ a large fixed integer such that

$$\sum_{x=M}^{\infty} e^{-\varepsilon e^{x^\alpha}+C_2} < 1 \text{ and } x \mapsto e^{-e^{x^\alpha}} \text{ is convex on } [M,\infty).$$

Then for $\alpha > 0$ and all ℓ sufficiently large,

$$\Sigma_{\text{Gumbel}}(\mathbf{t}) := \sum_{(x_u)_{u\in\mathbf{t}\setminus\rho}\in(\mathbb{N}\cap[M,\infty))^{\#\mathbf{t}-1}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} e^{x_u^\alpha} + C_2(\#\mathbf{t}-1)\Big\}\mathbf{1}_{\{\sum_{|u|=H}\mathbf{1}_{\{s_u\leq\ell\}}\leq b^{H-K}\}}$$
$$\leq \exp\big\{-(1-\varepsilon)\big(1-b^{-K}\big)e^{(\frac{(\alpha+1)\log b}{\alpha}(1-b^{-K})\ell)^{\frac{\alpha}{\alpha+1}}}\big\}. \tag{3.14}$$

To prove Lemma 3.3, we will introduce a tree-transformation. Recall that $\mathbf{t}$ is a fixed tree of H generations. We mark each $u \in \mathbf{t}\setminus\{\rho\}$ with $x_u \in \mathbb{R}$ and mark ρ with 0. For each $\mathbf{x} \in \{(x_u)_{u\in\mathbf{t}\setminus\{\rho\}} : x_u \in \mathbb{R}\}$, we regard $\mathbf{t}(\mathbf{x}) := \{(u, x_u), u \in \mathbf{t}\}$ as a marked tree. Then $\mathbf{t}$ is just the genealogical tree of $\mathbf{t}(\mathbf{x})$.

Given $\mathbf{x} \in \{(x_u)_{u\in\mathbf{t}\setminus\{\rho\}} : x_u \in \mathbb{R}\}$, we shall show that by manipulating the order of $u \in \mathbf{t}$ according to $\mathbf{x}$, one could construct a new marked tree $\mathbf{t}_*(\mathbf{x})$, where the lexicographical orders of individuals are totally rearranged so that the most recent common ancestor u^* of individuals located below a given level ℓ at the Hth generation is of the generation J with $H \geq J \geq K$. However, $\mathbf{t}_*(\mathbf{x})$ and $\mathbf{t}(\mathbf{x})$, viewed as sets of individuals, contain exactly the same individuals. And the mark of each individual remains the same. The detailed construction will be explained in the following.

We first colour the individuals in Hth generation. At the Hth generation, there are at most b^{H-K} individuals positioned below ℓ, which are all coloured blue. The other individuals above ℓ are coloured red.

At the $(H-1)$th generation, the individuals are called $u_{(1)}, u_{(2)}, \ldots, u_{(|\mathbf{t}|_{H-1})}$ according to their positions such that $s_{u_{(1)}} \geq s_{u_{(2)}} \geq \cdots \geq s_{u_{(|\mathbf{t}|_{H-1})}}$, where $|\mathbf{t}|_{H-1} =: \#\{u \in \mathbf{t} : |u| = H-1\}$. Let us start with $u_{(1)}$ and its children. If all children of $u_{(1)}$ are red, then we turn to $u_{(2)}$. Otherwise, we keep its red children and replace its blue children by the red children of other individuals of the $(H-1)$th generation. More precisely, saying that there are k blue children of $u_{(1)}$, we collect the red children of $u_{(2)}$ and then the red children of $u_{(3)}, \ldots$, until we find exactly k red ones to be exchanged with the original k blue children of $u_{(1)}$.

Note that in this way the number of children $u_{(1)}$ is unchanged and that all of them are positioned above ℓ and red. Now, we put $u_{(1)}$ aside and restart from $u_{(2)}$ by doing the same exchanges with $u_{(3)}, u_{(4)}, \ldots$. We would stop at some $u_{(k)}$ such that there is no red child left for $u_{(k+1)}, \ldots$. At this stage, there are at most 3 types of individuals at the $(H-1)$th generation: the ones with only red children; the ones with only blue children and the one with red children and blue children (note that there is at most one individual who has both red and blue children). Then the individuals with only red children are all coloured red. The others of the $(H-1)$th generation are coloured blue. An example is given in Fig. 2. Notice that the number of blue individuals of the $(H-1)$th generation are at most b^{H-K-1}.

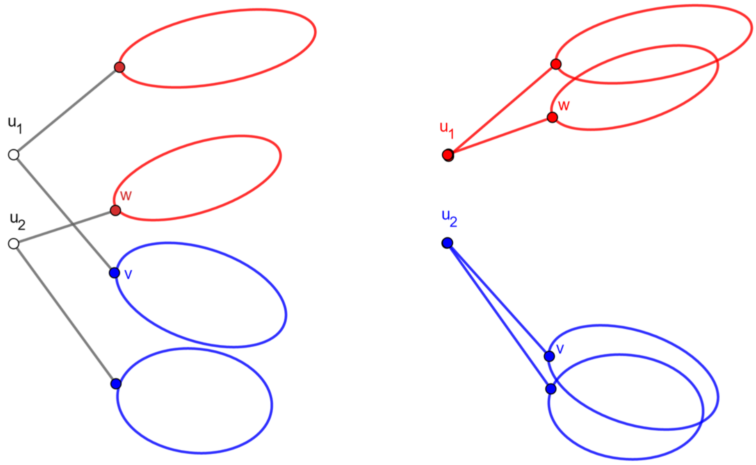

Fig. 3. Both u_1 and u_2 have two offsprings. After exchanging subtrees rooted at w and v, u_1 is colored red and u_2 is colored blue.

By iteration, we exchange individuals and colour the tree from one generation to the previous generation. When we exchange two individuals w and v, we exchange two subtrees rooted at w and v, as well as their displacements; see Fig. 3. Therefore, the positions of red individuals get higher, and obviously stay above ℓ. Finally, we stop at some generation J where only one individual is coloured blue for the first time. We hence obtain the new marked tree $\mathbf{t}_*(\mathbf{x})$, which depends on $\mathbf{t}$ and $\mathbf{x}$. Note that the ancestor u^* of blue ones is in the Jth generation with $J \geq K$. Observe that, for all red individuals, their descendants at Hth generation are positioned above ℓ. Note also that J depends on H, K, ℓ and $\mathbf{t}$.

In particular, if there is no blue individuals at all, we take $J = H$ and take one of individuals in the Hth generation to be u^*.

Now we cut this u^* and remove all its descendants from $\mathbf{t}_*(\mathbf{x})$ to get a pruned tree $\mathbf{t}_*^{\setminus u^*}(\mathbf{x})$. Note that all individuals of this tree $\mathbf{t}_*^{\setminus u^*}(\mathbf{x})$ up to the generation $H-1$ have at least b children, except the parent of u^*. And the parent of u^* has at least $b-1$ children. So we can extract from $\mathbf{t}_*^{\setminus u^*}(\mathbf{x})$ an "almost" b-ary regular tree $\mathbf{t}_b(\mathbf{x})$ so that its all descendants in the Hth generation are located above ℓ. Here in $\mathbf{t}_b(\mathbf{x})$, the parent of u^* has $b-1$ children (as u^* is removed), and all others except the leaves have exactly b children. Notice that in this construction, we do not change the mark x_u for any individual u. Denote by $\mathbf{t}_b^s(\mathbf{x})$ the genealogical tree of the marked tree $\mathbf{t}_b(\mathbf{x})$. Let $\mathbf{T_t}$ be the collection of all possible genealogical trees of $\mathbf{t}_b(\mathbf{x})$ from $\mathbf{t}$. That is

$$\mathbf{T_t} := \{\mathbf{t}_b^s(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^{\mathbf{t}}\},$$

where $\mathbb{R}^{\mathbf{t}} := \{(x_u)_{u\in\mathbf{t}\setminus\{\rho\}} : x_u \in \mathbb{R}\}$. With a little abuse of notation, we still use $\{x_u\}_{u\in\mathbf{t}_b^s(\mathbf{x})}$ and $\{s_u\}_{u\in\mathbf{t}_b^s(\mathbf{x})}$ to represent the displacements and positions of individuals in $\mathbf{t}_b^s(\mathbf{x})$.

Remark 3.1. A similar idea of tree transformation is used in [13]. However, in [13], a B-regular tree is extracted from the underlying tree, when $b < B := \sup\{k \geq 1 : p_k > 0\} < \infty$.

Now we are ready to show this technical lemma by using the above tree transformation.

Proof of Lemma 3.3. First, one sees that the tree transformation introduced above ensures

$$\begin{aligned}\Big\{\mathbf{x} \in \mathbb{R}^{\mathbf{t}} : \sum_{|u|=H; u\in\mathbf{t}} \mathbf{1}_{\{s_u \leq \ell\}} \leq b^{H-K}\Big\} &\subset \Big\{\mathbf{x}\in\mathbb{R}^{\mathbf{t}} : \min_{u\in\mathbf{t}_b^s(\mathbf{x}), |u|=H} s_u > \ell\Big\} \\ &= \bigcup_{\mathbf{t_0}\in\mathbf{T_t}} \Big\{\mathbf{x}\in\mathbb{R}^{\mathbf{t}} : \mathbf{t}_b^s(\mathbf{x}) = \mathbf{t_0}, \min_{u\in\mathbf{t_0}, |u|=H} s_u > \ell\Big\}.\end{aligned}$$

Now, let us turn to bound $\Sigma_{\text{Weibull}}(\mathbf{t})$ and $\Sigma_{\text{Gumbel}}(\mathbf{t})$. Recall that the choice of M ensures

$$\sum_{\mathbf{x}\in M_N^{\mathbf{t}}} e^{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda\varepsilon x_u^\alpha + C_2(\#\mathbf{t}-1)} = \prod_{u\in\mathbf{t}\setminus\{\rho\}} \sum_{x_u\in\mathbb{N}\cap[M,\infty)} e^{-\lambda\varepsilon x_u^\alpha + C_2} < 1, \tag{3.15}$$

where

$$M_N^{\mathbf{t}} := \{(x_u)_{u\in\mathbf{t}\setminus\rho} : x_u \in \mathbb{N}\cap[M,\infty)\} = \bigcup_{\mathbf{t_0}\in\mathbf{T_t}} \{\mathbf{x}\in M_N^{\mathbf{t}} : \mathbf{t}_b^s(\mathbf{x}) = \mathbf{t_0}\}.$$

Therefore, if we consider all possible choices of $\mathbf{t}_b$, then

$$\begin{aligned}\Sigma_{\mathrm{Weibull}}(\mathbf{t}) &\le \sum_{\mathbf{x}\in M_N^{\mathbf{t}}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_b^s(\mathbf{x}),|u|=H} s_u>\ell\}} \\ &= \sum_{\mathbf{t}_0\in\mathbf{T_t}} \sum_{\substack{\mathbf{x}\in M_N^{\mathbf{t}}\\ \mathbf{t}_b^s(\mathbf{x})=\mathbf{t_0}}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_0,|u|=H} s_u>\ell\}} \\ &\le \sum_{\mathbf{t}_0\in\mathbf{T_t}} \sum_{\substack{\mathbf{x}\in M_N^{\mathbf{t}}\\ \mathbf{t}_b^s(\mathbf{x})=\mathbf{t_0}}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda\varepsilon x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\}\exp\Big\{-\lambda(1-\varepsilon)\sum_{u\in\mathbf{t}_0\setminus\{\rho\}} x_u^\alpha\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_0,|u|=H} s_u>\ell\}}.\end{aligned}$$

Let

$$\begin{aligned}\mathcal{M}_{\mathrm{Weibull}}(\mathbf{t}) &:= \max_{\mathbf{t}_0\in\mathbf{T_t}} \max_{\substack{\mathbf{x}\in M_N^{\mathbf{t}}\\ \mathbf{t}_b^s(\mathbf{x})=\mathbf{t_0}}} \exp\Big\{-\lambda(1-\varepsilon)\sum_{u\in\mathbf{t}_0\setminus\{\rho\}} x_u^\alpha\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_0,|u|=H} s_u>\ell\}} \\ &= \max_{\mathbf{x}\in M_N^{\mathbf{t}}} \exp\Big\{-\lambda(1-\varepsilon)\sum_{u\in\mathbf{t}_b^s(\mathbf{x})\setminus\{\rho\}} x_u^\alpha\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_b^s(\mathbf{x}),|u|=H} s_u>\ell\}}.\end{aligned}$$

We hence get from (3.15) that

$$\begin{aligned}\Sigma_{\mathrm{Weibull}}(\mathbf{t}) &\le \mathcal{M}_{\mathrm{Weibull}}(\mathbf{t}) \sum_{\mathbf{t}_0\in\mathbf{T_t}} \sum_{\substack{\mathbf{x}\in M_N^{\mathbf{t}}\\ \mathbf{t}_b^s(\mathbf{x})=\mathbf{t_0}}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda\varepsilon x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\} \\ &= \mathcal{M}_{\mathrm{Weibull}}(\mathbf{t}) \sum_{\mathbf{x}\in M_N^{\mathbf{t}}} \exp\Big\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda\varepsilon x_u^\alpha + C_2(\#\mathbf{t}-1)\Big\} \\ &\le \mathcal{M}_{\mathrm{Weibull}}(\mathbf{t}).\end{aligned} \tag{3.16}$$

Similarly for $\Sigma_{\mathrm{Gumbel}}(\mathbf{t})$, we have

$$\Sigma_{\mathrm{Gumbel}}(\mathbf{t}) \le \mathcal{M}_{\mathrm{Gumbel}}(\mathbf{t}) := \max_{\mathbf{x}\in M_N^{\mathbf{t}}} \exp\Big\{-(1-\varepsilon)\sum_{u\in\mathbf{t}_b^s(\mathbf{x})\setminus\{\rho\}} e^{x_u^\alpha}\Big\}\mathbf{1}_{\{\min_{u\in\mathbf{t}_0,|u|=H} s_u>\ell\}}. \tag{3.17}$$

It remains to bound $\mathcal{M}_{\mathrm{Weibull}}(\mathbf{t})$ and $\mathcal{M}_{\mathrm{Gumbel}}(\mathbf{t})$.

For any $\mathbf{t}_0 \in \mathbf{T_t}$, let $|\mathbf{t}_0|_k := \#\{u \in \mathbf{t}_0 : |u| = k\}$ denote the population size of the kth generation of $\mathbf{t}_0$ and $\overline{x}_k := \frac{\sum_{|v|=k,v\in\mathbf{t}_0} x_v}{|\mathbf{t}_0|_k}$ denote the averaged displacement at the kth generation. Here we shall bound $\sum_{k=1}^H \overline{x}_k$ from below on $\{\min_{u\in\mathbf{t}_0,|u|=H} s_u > \ell\}$. According to the construction of $\mathbf{t}_b(\mathbf{x})$ as above, there exists an integer $J \in [K, H]$ such that

$$|\mathbf{t}_0|_k = b^k, \quad \forall 1 \le k \le J-1; \quad \text{and} \quad |\mathbf{t}_0|_k = b^k - b^{k-J}, \quad \forall J \le k \le H. \tag{3.18}$$

Observe that on $\{\min_{u\in\mathbf{t}_0,|u|=H} s_u > \ell\}$,

$$\sum_{|u|=H,u\in\mathbf{t}_0} s_u \ge \ell \times |\mathbf{t}_0|_H. \tag{3.19}$$

On the other hand,

$$\sum_{|u|=H,u\in\mathbf{t}_0} s_u = \sum_{|u|=H,u\in\mathbf{t}_0} \sum_{\rho\prec v\preceq u} x_v = \sum_{k=1}^{H} \sum_{|v|=k,v\in\mathbf{t}_0} \Big(x_v \sum_{|u|=H} \mathbf{1}_{\{v\preceq u\}} \Big),$$

where $\sum_{|u|=H} \mathbf{1}_{\{v\preceq u\}} \le b^{H-|v|}$ as $\mathbf{t}_0$ is a pruned b-ary tree. So, it follows from (3.19) and (3.18) that

$$\sum_{k=1}^{H} \sum_{|v|=k,v\in\mathbf{t}_0} x_v b^{H-k} \ge \ell\big(b^H - b^{H-J}\big). \tag{3.20}$$

Again by (3.18),

$$\sum_{k=1}^{H} \sum_{|v|=k} x_v b^{H-k} = b^H \left(\sum_{k=1}^{J-1} \overline{x}_k + \frac{b^k - b^{k-J}}{b^k} \sum_{k=J}^{H} \overline{x}_k \right) \le b^H \sum_{k=1}^{H} \overline{x}_k,$$

which, together with (3.20), gives

$$\sum_{k=1}^{H} \overline{x}_k \ge \big(1 - b^{-J}\big)\ell. \tag{3.21}$$

We shall treat $\mathcal{M}_{\text{Weibull}}(\mathbf{t})$ and $\mathcal{M}_{\text{Gumbel}}(\mathbf{t})$ separately.

Proof of (3.13) For any $\mathbf{t}_0 \in \mathbf{T}_\mathbf{t}$, let us find a suitable lower bound for

$$\min_{(x_u)_{u\in\mathbf{t}_0\setminus\{\rho\}}\in[M,+\infty)^{\#\mathbf{t}_0-1}} \sum_{u\in\mathbf{t}_0\setminus\{\rho\}} \lambda x_u^\alpha,$$

on $\{\min_{u\in\mathbf{t}_0,|u|=H} s_u > \ell\}$. For $\alpha > 1$, by convexity on $\mathbb{R}_+$ of $x \mapsto x^\alpha$, one has

$$\sum_{u\in\mathbf{t}_0} x_u^\alpha = \sum_{k=1}^{H} \sum_{u|=k,u\in\mathbf{t}_0} x_u^\alpha \ge \sum_{k=1}^{H} |\mathbf{t}_0|_k (\overline{x}_k)^\alpha.$$

So,

$$\sum_{u\in\mathbf{t}_0} x_u^\alpha \ge \big(1 - b^{-J}\big) \sum_{k=1}^{H} b^k (\overline{x}_k)^\alpha. \tag{3.22}$$

Let us take a positive sequence $(\mu_k)_{k\ge1}$, which will be determined later, with $\mu_\alpha := \sum_{k=1}^H \mu_k^\alpha$ and write

$$\sum_{k=1}^{H} b^k (\overline{x}_k)^\alpha = \mu_\alpha \sum_{k=1}^{H} \frac{\mu_k^\alpha}{\mu_\alpha} \big(\mu_k^{-1} b^{k/\alpha} \overline{x}_k\big)^\alpha,$$

which again by convexity implies

$$\sum_{k=1}^{H} b^k (\overline{x}_k)^\alpha \ge \mu_\alpha \left(\sum_{k=1}^{H} \frac{\mu_k^\alpha}{\mu_\alpha} \mu_k^{-1} b^{k/\alpha} \overline{x}_k \right)^\alpha = \mu_\alpha^{1-\alpha} \left(\sum_{k=1}^{H} \mu_k^{\alpha-1} b^{k/\alpha} \overline{x}_k \right)^\alpha. \tag{3.23}$$

We choose $\mu_k = b^{-\frac{k}{\alpha(\alpha-1)}}$ so that $\mu_k^{\alpha-1} b^{k/\alpha} \overline{x}_k = \overline{x}_k$ for any $k \ge 1$. Moreover,

$$\mu_\alpha = \sum_{k=1}^{H} b^{-\frac{k}{\alpha-1}} \le \frac{1}{b^{\frac{1}{\alpha-1}} - 1}.$$

So, (3.23) yields

$$\sum_{k=1}^{H} b^k (\overline{x}_k)^\alpha \geq \left(b^{\frac{1}{\alpha-1}} - 1\right)^{\alpha-1} \left(\sum_{k=1}^{H} \overline{x}_k\right)^\alpha. \tag{3.24}$$

Plugging it into (3.24) yields that

$$\sum_{k=1}^{H} b^k (\overline{x}_k)^\alpha \geq \left(b^{\frac{1}{\alpha-1}} - 1\right)^{\alpha-1} \left(\left(1 - b^{-J}\right)\ell\right)^\alpha.$$

Going back to (3.22), as $J \geq K$, we end up with

$$\sum_{u\in\mathbf{t}_0} x_u^\alpha \geq \left(b^{\frac{1}{\alpha-1}} - 1\right)^{\alpha-1} \left(1 - b^{-K}\right)^{\alpha+1} \ell^\alpha.$$

This shows that

$$\mathcal{M}_{\text{Weibull}}(\mathbf{t}) \leq \exp\left\{-\lambda\left(b^{\frac{1}{\alpha-1}} - 1\right)^{\alpha-1}\left(1 - b^{-K}\right)^{\alpha+1}\ell^\alpha\right\}.$$

Plugging it into (3.16) proves (3.13).

Proof of (3.14) Similarly, to bound $\mathcal{M}_{\text{Gumbel}}(\mathbf{t})$, we shall find a suitable lower bound for $\sum_{u\in\mathbf{t}_0\setminus\{\rho\}} e^{x_u^\alpha}$ on $\{\min_{u\in\mathbf{t}_0,|u|=H} s_u > \ell\}$ for any $\mathbf{t}_0 \in \mathbf{T}_\mathbf{t}$. Note that for any $\alpha > 0$, there exists $M \geq 1$ such that $x \mapsto e^{x^\alpha}$ is convex on $[M, \infty)$. For such M, one sees that

$$\sum_{u\in\mathbf{t}_0} e^{x_u^\alpha} = \sum_{k=1}^{H} |\mathbf{t}_0|_k \sum_{|u|=k} \frac{1}{|\mathbf{t}_0|_k} e^{x_u^\alpha} \geq \sum_{k=1}^{H} |\mathbf{t}_0|_k e^{\overline{x}_k^\alpha}.$$

As $|\mathbf{t}_0|_k \geq (1 - b^{-J})b^k$ for any $1 \leq k \leq H$, one gets that

$$\sum_{u\in\mathbf{t}_0} e^{x_u^\alpha} \geq \left(1 - b^{-J}\right) \sum_{k=1}^{H} b^k e^{\overline{x}_k^\alpha} \geq \left(1 - b^{-J}\right) e^{\Xi_H}, \tag{3.25}$$

where

$$\Xi_H := \max_{1\leq k\leq H} \left\{\overline{x}_k^\alpha + k\log b\right\} \geq H\log b + M^\alpha.$$

By this definition, one sees

$$\overline{x}_k \leq (\Xi_H - k\log b)^{1/\alpha}, \quad \forall k \in \{1, \dots, H\}.$$

So, combining it with (3.21) yields

$$\sum_{k=1}^{H} (\Xi_H - k\log b)^{1/\alpha} \geq \left(1 - b^{-J}\right)\ell.$$

On the other hand, by monotonicity of $x \mapsto (\Xi_H - x\log b)^{1/\alpha}$ on $[0, \frac{\Xi_H}{\log b}]$, one has

$$\sum_{k=1}^{H} (\Xi_H - k\log b)^{1/\alpha} \leq \int_0^H (\Xi_H - x\log b)^{1/\alpha}\, dx \leq \frac{\alpha}{(1+\alpha)\log b} \Xi_H^{1+\frac{1}{\alpha}}.$$

We then deduce that

$$\Xi_H \geq \left(\frac{(\alpha+1)\log b}{\alpha}\left(1 - b^{-J}\right)\ell\right)^{\frac{\alpha}{\alpha+1}}.$$

Going back to (3.25), one gets

$$\sum_{u\in\mathbf{t}_0} e^{x_u^\alpha} \geq \left(1-b^{-K}\right)\exp\left\{\left(\frac{(\alpha+1)\log b}{\alpha}\left(1-b^{-K}\right)\ell\right)^{\frac{\alpha}{\alpha+1}}\right\}.$$

since $J \geq K$. This shows

$$\mathcal{M}_{\text{Gumbel}}(\mathbf{t}) \leq \exp\left\{-(1-\varepsilon)\left(1-b^{-K}\right)e^{\left(\frac{(\alpha+1)\log b}{\alpha}(1-b^{-K})\ell\right)^{\frac{\alpha}{\alpha+1}}}\right\}.$$

Plugging it into (3.17) proves (3.14). □

3.3.2. *Lower bound*

We shall prove here that if $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$ with $\alpha > 1$, as $z \to +\infty$, then

$$\liminf_{n\to\infty} \frac{1}{\ell_n^\alpha}\log\mathbb{P}(M_n \leq m_n - \ell_n) \geq -\lambda\left(b^{\frac{1}{\alpha-1}}-1\right)^{\alpha-1}. \tag{3.26}$$

By the assumption of Theorem 1.3, there exist two constants $0 < c_{13} < c_{14} < \infty$ such that for any $x > 0$,

$$c_{13}e^{-\lambda x^\alpha} \leq \mathbb{P}(X \leq -x) \leq c_{14}e^{-\lambda x^\alpha}. \tag{3.27}$$

We choose $t_n^- = o(\ell_n)$ such that $t_n^- \uparrow \infty$ and suppose that up to the t_n^-th generation, the genealogical tree is a b-regular tree. For any individual $|u| = k$ with $1 \leq k \leq t_n^-$, we suppose that its displacement X_u is less than $-a_k$ with some $a_k > 0$. We will determine t_n^- and the sequence $(a_k)_{k\geq 1}$ later. Therefore,

$$\begin{aligned}\mathbb{P}(M_n \leq m_n - \ell_n) &\geq \mathbb{P}\left(Z_{t_n^-} = b^{t_n^-}; \forall|u| = k \in \{1,\dots,t_n^-\}, X_u < -a_k; M_n \leq m_n - \ell_n\right)\\ &\geq \mathbb{P}\left(Z_{t_n^-} = b^{t_n^-}; \forall|u| = k \in \{1,\dots,t_n^-\}, X_u < -a_k; \max_{|z|=t_n^-}\left(S_z + M^z_{n-t_n^-}\right) \leq m_n - \ell_n\right)\\ &\geq \mathbb{P}\left(Z_{t_n^-} = b^{t_n^-}; \forall|u| = k \in \{1,\dots,t_n^-\}, X_u < -a_k; \max_{|z|=t_n^-} M^z_{n-t_n^-} \leq m_n - \ell_n + \sum_{k=1}^{t_n^-} a_k\right).\end{aligned}$$

Once again by branching property, one has

$$\begin{aligned}&\mathbb{P}(M_n \leq m_n - \ell_n)\\ &\quad\geq \mathbb{P}\left(Z_{t_n^-} = b^{t_n^-}; \forall|u| = k \in \{1,\dots,t_n^-\}, X_u < -a_k\right)\mathbb{P}\left(M_{n-t_n^-} \leq m_n - \ell_n + \sum_{k=1}^{t_n^-} a_k\right)^{b^{t_n^-}}.\end{aligned} \tag{3.28}$$

For the first term on the right hand side, by independence of branching structure and displacements,

$$\begin{aligned}&\mathbb{P}\left(Z_{t_n^-} = b^{t_n^-}; \forall|u| = k \in \{1,\dots,t_n^-\}, X_u < -a_k\right)\\ &\quad= p_b^{\sum_{k=0}^{t_n^- -1} b^k}\prod_{k=1}^{t_n^-}\mathbb{P}(X < -a_k)^{b^k},\end{aligned}$$

which by (3.27), is larger than

$$p_b^{\frac{b^{t_n^-}-1}{b-1}}\prod_{k=1}^{t_n^-} c_{13}^{b^k} e^{-\lambda(a_k)^\alpha b^k} = p_b^{\frac{b^{t_n^-}-1}{b-1}} c_{13}^{\frac{b^{t_n^-+1}-b}{b-1}}\exp\left\{-\lambda\sum_{k=1}^{t_n^-} a_k^\alpha b^k\right\}. \tag{3.29}$$

Now, we take the values of a_k. Let $b_\alpha := b^{\frac{1}{\alpha-1}}$ and $a_k = \frac{(b_\alpha-1)}{b_\alpha^k}\ell_n$. Note that $\sum_{k=1}^{t_n^-} a_k = (1-b_\alpha^{-t_n^-})\ell_n$. Take $t_n^- = \lceil(\alpha-1)\frac{\log\ell_n}{\log b}\rceil$ so that for n large enough,

$$m_n - \ell_n + \sum_{k=1}^{t_n^-} a_k = m_n - \ell_n + \left(1-b_\alpha^{-t_n^-}\right)\ell_n \geq m_{n-t_n^-} + y^*, \tag{3.30}$$

with y^* chosen in (2.15). Meanwhile, one obtains that

$$b^{t_n^-}=\Theta\big(\ell_n^{\alpha-1}\big)\quad\text{and}\quad\sum_{k=1}^{t_n^-}a_k^\alpha b^k=\ell_n^\alpha(b_\alpha-1)^{\alpha-1}\big(1-b_\alpha^{-t_n^-}\big)=\ell_n^\alpha(b_\alpha-1)^{\alpha-1}-\Theta\big(\ell_n^{\alpha-1}\big).$$

Plugging them into (3.29) yields

$$\mathbb{P}\big(Z_{t_n^-}=b^{t_n^-};\forall|u|=k\in\{1,\dots,t_n^-\},X_u<-a_k\big)\geq\exp\big\{-\lambda\ell_n^\alpha(b_\alpha-1)^{\alpha-1}-\Theta\big(\ell_n^{\alpha-1}\big)\big\}. \tag{3.31}$$

Applying it and (3.30) to (3.28), together with (2.15), gives that

$$\begin{aligned}\mathbb{P}(M_n\leq m_n-\ell_n)&\geq\exp\big\{-\lambda\ell_n^\alpha(b_\alpha-1)^{\alpha-1}-\Theta\big(\ell_n^{\alpha-1}\big)\big\}\mathbb{P}\big(M_{n-t_n^-}\leq m_{n-t_n^-}+y^*\big)^{b^{t_n^-}}\\&\geq\exp\big\{-\lambda\ell_n^\alpha(b_\alpha-1)^{\alpha-1}-\Theta\big(\ell_n^{\alpha-1}\big)\big\}(1/2)^{\ell_n^{\alpha-1}}.\end{aligned}$$

This suffices to conclude (3.26).

3.3.3. *Upper bound*

We are going to use the rough upper bound (3.5) in Lemma 3.2 and get a better estimate. We still use some intermediate time $t_n^+=\lfloor t^+\log\ell_n\rfloor$ where $t^+>0$ will be determined later. The rough idea is similar to what we used above to prove (3.5). Take $B_n=[-(1-\varepsilon)\ell_n,\infty)$ with $\varepsilon\in(0,1)$. Observe that for $\delta_n:=\delta\log\ell_n$ with some $\delta\in(0,t^+)$, we have

$$\mathbb{P}(M_n\leq m_n-\ell_n)\leq\mathbb{P}\big(Z_{t_n^+}(B_n)\geq b^{t_n^+-\delta_n};M_n\leq m_n-\ell_n\big)+\mathbb{P}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big). \tag{3.32}$$

Similarly to (3.8), by branching property at time t_n^+, one has

$$\begin{aligned}\mathbb{P}\big(Z_{t_n^+}(B_n)\geq b^{t_n^+-\delta_n};M_n\leq m_n-\ell_n\big)&\leq\mathbb{P}(M_{n-t_n^+}\leq m_n-\varepsilon\ell_n)^{b^{t_n^+-\delta_n}}\\&\leq\mathbb{P}(M_{n-t_n^+}\leq m_{n-t_n^+}-\varepsilon\ell_n/2)^{b^{t_n^+-\delta_n}}.\end{aligned}$$

By use of the rough upper bound (3.5) with $\theta=4$, we get

$$\mathbb{P}\big(Z_{t_n^+}(B_n)\geq b^{t_n^+-\delta_n};M_n\leq m_n-\ell_n\big)\leq e^{-\varepsilon\ell_n b^{t_n^+-\delta_n}}. \tag{3.33}$$

It remains to bound $\mathbb{P}(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n})$. Let $\mathbf{t}$ denote a fixed tree of t_n^+ generations and $\mathbb{P}^{\mathbf{t}}(\cdot)$ denote the conditional probability $\mathbb{P}(\cdot|\mathcal{T}_{t_n^+}=\mathbf{t})$ where $\mathcal{T}_{t_n^+}$ denotes the genealogical tree $\mathcal{T}$ up to the t_n^+th generation. Observe that

$$\mathbb{P}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big)=\sum_{\mathbf{t}}\mathbb{P}(\mathcal{T}_{t_n^+}=\mathbf{t})\mathbb{P}^{\mathbf{t}}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big). \tag{3.34}$$

Here for convenience, for $C_2=(\log c_{14})_+$ and $\varepsilon\in(0,1)$, we can replace each displacement X_u by $X_u^+:=(-X_u)\vee M$ for some large and fixed constant M chosen in (3.13) of Lemma 3.3. Now denote the new positions achieved by these new displacements by

$$S_u^+:=\sum_{\rho\prec v\preceq u}X_v^+,\quad\forall|u|\leq t_n^+.$$

Obviously, $S_u^+\geq\sum_{\rho\prec v\preceq u}(-X_v)=-S_u$. So, if $Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}$, then

$$\sum_{|u|=t_n^+}\mathbf{1}_{\{S_u^+\leq(1-\varepsilon)\ell_n\}}\leq\sum_{|u|=t_n^+}\mathbf{1}_{\{S_u\in B_n\}}=Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}.$$

Therefore, for $\varepsilon\in(0,1/2)$ and for n sufficiently large such that $t_n^+=\lfloor t^+\log\ell_n\rfloor\leq\varepsilon\ell_n$,

$$\begin{aligned}&\mathbb{P}^{\mathbf{t}}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big)\\&\quad\leq\mathbb{P}^{\mathbf{t}}\Big(\sum_{|u|=t_n^+}1_{\{S_u^+\leq(1-\varepsilon)\ell_n\}}\leq b^{t_n^+-\delta_n}\Big)\end{aligned}$$

$$\leq \sum_{(x_u)_{u\in\mathbf{t}}\in(\mathbb{N}\cap[M,\infty))^{\#\mathbf{t}}} \prod_{u\in\mathbf{t}\setminus\{\rho\}} \mathbb{P}\big(X_u^+ \in [x_u, x_u+1)\big)\mathbf{1}_{\{\sum_{|u|=t_n^+} 1_{\{s_u\leq(1-2\varepsilon)\ell_n\}}\leq b^{t_n^+-\delta_n}\}}$$

$$\leq \sum_{(x_u)_{u\in\mathbf{t}}\in(\mathbb{N}\cap[M,\infty))^{\#\mathbf{t}}} \exp\bigg\{-\sum_{u\in\mathbf{t}\setminus\{\rho\}} \lambda x_u^\alpha + C_2(\#\mathbf{t}-1)\bigg\}\mathbf{1}_{\{\sum_{|u|=t_n^+} 1_{\{s_u\leq(1-2\varepsilon)\ell_n\}}\leq b^{t_n^+-\delta_n}\}}, \tag{3.35}$$

where $s_u := \sum_{\rho\prec v\preceq u} x_v$ and $C_2 = (\log c_{14})_+$ and for the second inequality, we use the fact that as $x_u \leq X_u^+ < x_u + 1$, one has $s_u \leq S_u^+ < s_u + t_n^+$ and

$$\mathbf{1}_{\{S_u^+\leq(1-\varepsilon)\ell_n\}} \geq \mathbf{1}_{\{s_u\leq(1-\varepsilon)\ell_n - t_n^+\}} \geq \mathbf{1}_{\{s_u\leq(1-2\varepsilon)\ell_n\}}.$$

So $\{\sum_{|u|=t_n^+} 1_{\{S_u^+\leq(1-\varepsilon)\ell_n\}} \leq b^{t_n^+-\delta_n}\} \subset \{\sum_{|u|=t_n^+} 1_{\{s_u\leq(1-2\varepsilon)\ell_n\}} \leq b^{t_n^+-\delta_n}\}$. Now by use of (3.13) with $H = t_n^+$, $K = \delta_n$, $\ell = (1-2\varepsilon)\ell_n$, we obtain

$$\mathbb{P}^{\mathbf{t}}\big(Z_{t_n^+}(B_n) < b^{t_n^+-\delta_n}\big) \leq \exp\big\{-\lambda(1-\varepsilon)C(b,\alpha)\big(1-b^{-\delta_n}\big)^{\alpha+1}(1-2\varepsilon)^\alpha \ell_n^\alpha\big\}.$$

Plugging it into (3.34) brings out

$$\mathbb{P}\big(Z_{t_n^+}(B_n) < b^{t_n^+-\delta_n}\big) \leq \exp\big\{-\lambda(1-\varepsilon)C(b,\alpha)\big(1-b^{-\delta_n}\big)^{\alpha+1}(1-2\varepsilon)^\alpha \ell_n^\alpha\big\}, \tag{3.36}$$

which, combined with (3.32) and (3.33), implies

$$\mathbb{P}(M_n \leq m_n - \ell_n) \leq e^{-\varepsilon\ell_n b^{t_n^+-\delta_n}} + e^{-\lambda(1-\varepsilon)C(b,\alpha)(1-b^{-\delta_n})^{\alpha+1}(1-2\varepsilon)^\alpha \ell_n^\alpha},$$

with $t_n^+ = \lfloor t^+ \log \ell_n \rfloor$, $\delta_n = \delta \log \ell_n$. We choose here $t^+ = \frac{3\alpha-1}{3\log b}$ and $\delta = \frac{1}{3\log b}$ so that

$$\ell_n b^{t_n^+-\delta_n} \sim \ell_n^{\alpha+1/3}.$$

Consequently, letting $n \uparrow \infty$ and then $\varepsilon \downarrow \infty$ shows

$$\limsup_{n\to\infty} \frac{1}{\ell_n^\alpha} \log \mathbb{P}(M_n \leq m_n - \ell_n) \leq -\lambda C(b,\alpha),$$

which is what we need.

3.4. *Proof of Theorem 1.5: Step size of Gumbel tail*

The arguments for Gumbel tail are similar to that for Weibull tail of $\alpha > 1$. For upper bound, we will employ the tree-transformation and Lemma 3.3 in Section 3.3.1 again.

3.4.1. *Lower bound of Theorem 1.5*

We are going to demonstrate that if $\mathbb{P}(X \leq -z) = \Theta(1)\exp(-e^{z^\alpha})$ as $z \to +\infty$ with $\alpha > 0$, then

$$\mathbb{P}(M_n \leq m_n - \ell_n) \geq \exp\big\{-e^{\beta(\alpha,b)\ell_n^{\frac{\alpha}{\alpha+1}} + o(\ell_n^{\frac{\alpha}{\alpha+1}})}\big\},$$

where $\beta(\alpha,b) := (\frac{1+\alpha}{\alpha}\log b)^{\frac{\alpha}{\alpha+1}}$.

By the assumption of Theorem 1.5, there exist two constants $0 < c_{15} < c_{16} < \infty$ such that for any $x \geq 0$,

$$c_{15} e^{-e^{x^\alpha}} \leq \mathbb{P}(X < -x) \leq c_{16} e^{-e^{x^\alpha}}. \tag{3.37}$$

Note that here $\alpha > 0$. Using the similar arguments as in Section 3.3.2, we take some intermediate time $t_n^- = o(\ell_n)$ and a positive sequence $(a_k)_{1\le k\le t_n^-}$. Then, observe that

$$\begin{aligned}\mathbb{P}(M\le m_n-\ell_n) &\ge \mathbb{P}\big(Z_{t_n^-}=b^{t_n^-};\forall|u|=k\in\{1,\dots,t_n^-\},X_u\le -a_k;M_n\le m_n-\ell_n\big)\\ &\ge \mathbb{P}\big(Z_{t_n^-}=b^{t_n^-};\forall|u|=k\in\{1,\dots,t_n^-\},X_u\le -a_k\big)\mathbb{P}\Bigg(M_{n-t_n^-}\le m_n-\ell_n+\sum_{k=1}^{t_n^-}a_k\Bigg)^{b^{t_n^-}}\\ &= p_b^{\sum_{k=0}^{t_n^- -1}b^k}\prod_{k=1}^{t_n^-}\mathbb{P}(X<-a_k)^{b^k}\,\mathbb{P}\Bigg(M_{n-t_n^-}\le m_n-\ell_n+\sum_{k=1}^{t_n^-}a_k\Bigg)^{b^{t_n^-}}.\end{aligned}$$

By (3.37), one has

$$\mathbb{P}(M\le m_n-\ell_n)\ge p_b^{\frac{b^{t_n^-}-1}{b-1}}c_{15}^{\frac{b^{t_n^-+1}-b}{b-1}}\exp\Bigg\{-\sum_{k=1}^{t_n^-}e^{a_k^\alpha}b^k\Bigg\}\mathbb{P}\Bigg(M_{n-t_n^-}\le m_n-\ell_n+\sum_{k=1}^{t_n^-}a_k\Bigg)^{b^{t_n^-}}. \tag{3.38}$$

Here we take $t_n^- := t^-\ell_n^{\frac{\alpha}{\alpha+1}}$ and $a_k := (\log b)^{1/\alpha}(t_n^- + 1 - k)^{1/\alpha}$ with $t^- := (\frac{1+\alpha}{\alpha})^{\frac{\alpha}{\alpha+1}}(\log b)^{-\frac{1}{\alpha+1}}$. Now observe that for arbitrary small $\varepsilon > 0$ and n large enough,

$$\begin{aligned}\sum_{k=1}^{t_n^-}a_k &= (\log b)^{1/\alpha}\sum_{k=1}^{t_n^-}\big(t_n^-+1-k\big)^{1/\alpha}\ge(\log b)^{1/\alpha}\int_1^{t_n^-}\big(t_n^-+1-s\big)^{1/\alpha}\,ds\\ &= \ell_n-\Theta(1)\ge \ell_n-\big(m_n-m_{n-t_n^-}-y^*\big).\end{aligned}$$

This leads to the fact that

$$\begin{aligned}\mathbb{P}\Bigg(M_{n-t_n^-}\le m_n-\ell_n+\sum_{k=1}^{t_n^-}a_k\Bigg)^{b^{t_n^-}} &\ge \mathbb{P}\big(M_{n-t_n^-}\le m_{n-t_n^-}+y^*\big)^{b^{t_n^-}}\\ &\ge e^{-\Theta(b^{t_n^-})}.\end{aligned}$$

On the other hand, note that

$$\sum_{k=1}^{t_n^-}e^{a_k^\alpha}b^k=\sum_{k=1}^{t_n^-}b^{t_n^-+1}=bt_n^-e^{(t^-\log b)\ell_n^{\frac{\alpha}{\alpha+1}}}.$$

Going back to (3.38), as $b^{t_n^-}\ll t_n^-e^{(t^-\log b)\ell_n^{\frac{\alpha}{\alpha+1}}}$ and $t_n^-=e^{o(\ell_n^{\frac{\alpha}{\alpha+1}})}$, one obtains that

$$\mathbb{P}(M\le m_n-\ell_n)\ge\exp\big\{-t_n^-e^{(t^-\log b)\ell_n^{\frac{\alpha}{\alpha+1}}}-\Theta\big(b^{t_n^-}\big)\big\}=\exp\big\{-e^{\beta(\alpha,b)\ell_n^{\frac{\alpha}{\alpha+1}}+o(\ell_n^{\frac{\alpha}{\alpha+1}})}\big\},$$

where $\beta(\alpha,b)=t^-\log b=(\frac{1+\alpha}{\alpha}\log b)^{\frac{\alpha}{\alpha+1}}$.

3.4.2. *Upper bound of Theorem 1.5*

It remains to prove that if $\mathbb{P}(X\le -z)=\Theta(1)\exp(-e^{z^\alpha})$ as $z\to+\infty$ with $\alpha>0$, then

$$\liminf_{n\to\infty}\frac{1}{\ell_n^{\frac{\alpha}{\alpha+1}}}\log\big[-\log\mathbb{P}(M_n\le m_n-\ell_n)\big]\ge\left(\frac{\alpha+1}{\alpha}\log b\right)^{\frac{\alpha}{\alpha+1}}.$$

Let $t_n^+ := t^+\ell_n^{\frac{\alpha}{\alpha+1}} = o(\ell_n)$ and $\delta_n := \delta\ell_n^{\frac{\alpha}{\alpha+1}}$ with some $0<\delta<t^+<\infty$. Using the similar arguments as in the Section 3.3.3, in view of (3.32) and to (3.33), one sees that for any $\varepsilon\in(0,1/2)$,

$$\begin{aligned}\mathbb{P}(M_n\le m_n-\ell_n) &\le \mathbb{P}\big(Z_{t_n^+}(B_n)\ge b^{t_n^+-\delta_n}; M_n\le m_n-\ell_n\big)+\mathbb{P}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big)\\ &\le \mathbb{P}(M_{n-t_n^+}\le m_{n-t_n^+}-\varepsilon\ell_n/2)^{b^{t_n^+-\delta_n}}+\mathbb{P}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big),\end{aligned}$$

which by (3.6) is bounded by

$$\exp\big(-e^{c_{31}(\varepsilon\ell_n/2)^{\frac{\alpha}{\alpha+1}}}b^{t_n^+-\delta_n}\big)+\mathbb{P}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big).$$

Similarly to (3.34) and (3.35), one also sees that

$$\begin{aligned}&\mathbb{P}(M_n\le m_n-\ell_n)\\ &\quad\le \exp\big(-e^{c_{31}(\varepsilon\ell_n/2)^{\frac{\alpha}{\alpha+1}}}b^{t_n^+-\delta_n}\big)+\sum_{\mathbf{t}}\mathbb{P}(\mathcal{T}_{t_n^+}=\mathbf{t})\mathbb{P}^{\mathbf{t}}\big(Z_{t_n^+}(B_n)<b^{t_n^+-\delta_n}\big)\\ &\quad\le \exp\big(-e^{c_{17}(\varepsilon\ell_n/2)^{\frac{\alpha}{\alpha+1}}}b^{t_n^+-\delta_n}\big)\\ &\qquad+\sum_{\mathbf{t}}\mathbb{P}(\mathcal{T}_{t_n^+}=\mathbf{t})\sum_{(x_u)_{u\in\mathbf{t}}\in(\mathbb{N}\cap[M,\infty))^{\#\mathbf{t}}}\exp\Big\{-\sum_{u\in\mathbf{t}}e^{x_u^\alpha}+C_2\#\mathbf{t}\Big\}\mathbf{1}_{\{\sum_{|u|=t_n^+}1_{\{s_u\le(1-2\varepsilon)\ell_n\}}\le b^{t_n^+-\delta_n}\}},\end{aligned}\tag{3.39}$$

with M chosen in accordance with $C_2=(\log c_{16})_+$ and $\varepsilon\in(0,1)$. Applying (3.14) yields

$$\begin{aligned}\mathbb{P}(M_n\le m_n-\ell_n)&\le \exp\big(-e^{c_{17}(\varepsilon\ell_n/2)^{\frac{\alpha}{\alpha+1}}}b^{t_n^+-\delta_n}\big)\\ &\quad+\exp\big\{-(1-\varepsilon)\big(1-b^{-\delta_n}\big)e^{(\frac{(\alpha+1)\log b}{\alpha}(1-b^{-\delta_n})(1-2\varepsilon)\ell_n)^{\frac{\alpha}{\alpha+1}}}\big\}.\end{aligned}\tag{3.40}$$

Here we choose $t^+=[(\frac{\alpha+1}{\alpha}(1-2\varepsilon)\log b)^{\frac{\alpha}{\alpha+1}}-\eta_\varepsilon/6]/\log b$ and $\delta=\frac{\eta_\varepsilon}{6\log b}$ where $\eta_\varepsilon=c_{17}(\frac{\varepsilon}{2})^{\frac{\alpha}{\alpha+1}}$ so that

$$e^{c_{17}(\varepsilon\ell_n/2)^{\frac{\alpha}{\alpha+1}}}b^{t_n^+-\delta_n}\gg e^{(\frac{\alpha+1}{\alpha}(1-2\varepsilon)\log b)^{\frac{\alpha}{\alpha+1}}\ell_n^{\frac{\alpha}{\alpha+1}}}.$$

This suffices to conclude

$$\liminf_{n\to\infty}\frac{1}{\ell_n^{\frac{\alpha}{\alpha+1}}}\log\big[-\log\mathbb{P}(M_n\le m_n-\ell_n)\big]\ge(1-\varepsilon)\Big(\frac{\alpha+1}{\alpha}(1-2\varepsilon)\log b\Big)^{\frac{\alpha}{\alpha+1}},$$

for arbitrary small $\varepsilon>0$. This is exactly what we need.

4. Small ball probability of D_∞ in Böttcher case

This section is devoted to proving Propositions 1.4 and 1.6. In fact, we only prove Proposition 1.4 where $\mathbb{P}(X<-x)=\Theta(1)e^{-\lambda x^\alpha}$ as $x\to\infty$. And we feel free to omit the proof of Proposition 1.6 as it follows from similar arguments.

Write D for D_∞ in (1.6) for simplicity. It is easy to see that for any time $n\ge1$,

$$D\overset{a.s.}{=}\sum_{|u|=n}e^{\theta^*(S_u-nx^*)}D^{(u)},\tag{4.1}$$

where given $(S_u;|u|=n)$, $(D^{(u)})_{\{|u|=n\}}$ are i.i.d. copies of D. It is known from [34] that under (1.1), (1.2), (1.3), (1.4) and (1.5), there exists a constant $C_D>0$ such that as $x\to+\infty$,

$$\mathbb{P}(D>x)\sim\frac{C_D}{x}.\tag{4.2}$$

However, on the survival set, the left tail of D has different regimes. In fact, from (4.1), one sees that to get $\{D<\varepsilon\}$, we could force all individuals of some generation t_ε to move to some much lower position than $t_\varepsilon x^*$ so that $e^{\theta^*(S_u-t_\varepsilon x^*)}\ll\varepsilon$ and $D^{(u)}$ are typical. This idea is very similar to that used in the previous sections. That is how we get Propositions 1.4 and 1.6 parallel to Theorems 1.3 and 1.5, respectively. The detailed arguments are in the following.

4.1. *Lower bound of Proposition 1.4*

First observe from (4.1) that for any $n \geq 1$ and $\delta > 0$,

$$\mathbb{P}(D < \varepsilon) = \mathbb{P}\Bigg(\sum_{|u|=n} e^{\theta^*(S_u - nx^*)} D^{(u)} < \varepsilon\Bigg)$$
$$\geq \mathbb{P}\Bigg(\forall |u| = n, e^{\theta^*(S_u - nx^*)} \leq \varepsilon^{1+\delta}; \sum_{|u|=n} D^{(u)} < \varepsilon^{-\delta}\Bigg),$$

where $\sum_{|u|=n} D^{(u)} = \Theta_{\mathbb{P}}(Z_n \log Z_n)$ because of (4.2) and weak law for triangular arrays (Theorem 2.2.6 in [18]). Therefore, by independence,

$$\mathbb{P}(D < \varepsilon) \geq \mathbb{P}\Bigg(\forall |u| = n, e^{\theta^*(S_u - nx^*)} \leq \varepsilon^{1+\delta}; Z_n = b^n; \sum_{|u|=n} D^{(u)} < \varepsilon^{-\delta}\Bigg)$$
$$= \mathbb{P}\Bigg(\forall |u| = n, S_u \leq (1+\delta)\frac{\log \varepsilon}{\theta^*} + nx^*; Z_n = b^n\Bigg)\mathbb{P}\Bigg(\sum_{k=1}^{b^n} D_k < \varepsilon^{-\delta}\Bigg),$$

where $D_k; k \geq 1$ are i.i.d. copies of D. Again by weak law for triangular arrays (Theorem 2.2.6 in [18]), $\sum_{k=1}^{b^n} D_k = (C_D + o_{\mathbb{P}}(1))b^n \log(b^n)$. As long as we take $n = t_\varepsilon \ll \frac{-\delta \log \varepsilon}{\log b}$ so that $nb^n \ll \varepsilon^{-\delta}$, $\mathbb{P}(\sum_{k=1}^{b^n} D_k < \varepsilon^{-\delta}) = 1 + o(1)$. So for $\varepsilon > 0$ small enough,

$$\mathbb{P}(D < \varepsilon) \geq \frac{1}{2}\mathbb{P}\Bigg(\forall |u| = t_\varepsilon, S_u \leq (1+\delta)\frac{\log \varepsilon}{\theta^*} + t_\varepsilon x^*; Z_{t_\varepsilon} = b^{t_\varepsilon}\Bigg). \tag{4.3}$$

The sequel of this proof will be divided into two parts for Weibull tail with $\alpha > 1$ and $\alpha \in (0, 1]$. Write $a_\varepsilon := -\log \varepsilon$ for convenience.

Subpart 1: *the case of* $\alpha > 1$

Recall that $b_\alpha = b^{1/(\alpha-1)}$. Choose $t_\varepsilon = \lfloor \frac{\log((1+\delta)a_\varepsilon/\theta^* x^*)}{\log b_\alpha} \rfloor$ and $x_k = \frac{(b_\alpha - 1)(1+\delta)a_\varepsilon}{\theta^* b_\alpha^k}$. Then $t_\varepsilon \gg 1$ and $\sum_{k=1}^{t_\varepsilon}(-x_k) = (1 - b_\alpha^{-t_\varepsilon})\frac{-(1+\delta)a_\varepsilon}{\theta^*} \leq (1+\delta)\frac{\log \varepsilon}{\theta^*} + t_\varepsilon x^*$. As a consequence,

$$\mathbb{P}(D < \varepsilon) \geq \frac{1}{2}\mathbb{P}\Bigg(\forall |u| = t_\varepsilon, S_u \leq (1+\delta)\frac{\log \varepsilon}{\theta^*} + t_\varepsilon x^*; Z_{t_\varepsilon} = b^{t_\varepsilon}\Bigg)$$
$$\geq \frac{1}{2}\mathbb{P}\big(Z_{t_\varepsilon} = b^{t_\varepsilon}; \forall |u| = k \in \{1, \ldots, t_\varepsilon\}, X_u < -x_k\big)$$
$$\geq \exp\Bigg\{-\lambda\big(1 - b_\alpha^{-t_\varepsilon}\big)\bigg(\frac{(1+\delta)a_\varepsilon}{\theta^*}\bigg)^\alpha (b_\alpha - 1)^{\alpha-1} - \Theta\bigg(\bigg(\frac{(1+\delta)a_\varepsilon}{\theta^* x^*}\bigg)^{\alpha-1}\bigg)\Bigg\}, \tag{4.4}$$

where the inequality follows from the same reasonings as (3.31). Letting $\varepsilon \downarrow 0$ then $\delta \downarrow 0$ implies that

$$\liminf_{\varepsilon \to 0+} \frac{1}{(-\log \varepsilon)^\alpha} \log \mathbb{P}(D_\infty < \varepsilon) \geq -\frac{\lambda}{(\theta^*)^\alpha}\big(b^{\frac{1}{\alpha-1}} - 1\big)^{\alpha-1}.$$

Subpart 2: *the case of* $\alpha \in (0, 1]$

Choose $t_\varepsilon = 1$. Then it follows from (4.3) that for $\varepsilon > 0$ sufficiently small,

$$\mathbb{P}(D < \varepsilon) \geq \frac{1}{2}\mathbb{P}\Bigg(Z_1 = b; X_u \leq (1+\delta)\frac{\log \varepsilon}{\theta^*} + x^*, \text{ for all } |u| = 1\Bigg)$$
$$\geq c_{18} e^{-\lambda b((1+\delta)\frac{-\log \varepsilon}{\theta^*} - x^*)^\alpha}, \tag{4.5}$$

which implies

$$\liminf_{\varepsilon \to 0+} \frac{1}{(-\log \varepsilon)^\alpha} \log \mathbb{P}(D_\infty < \varepsilon) \geq -\frac{\lambda(1+\delta)^\alpha}{(\theta^*)^\alpha} b.$$

Then we obtain the lower bound by letting $\delta \to 0$.

4.2. *Upper bound of Proposition* 1.4

Again we divide the proof into two parts for Weibull tail with $\alpha > 1$ and $\alpha \in (0, 1]$. Define

$$U_0(t,\ell) := \{u \in \mathcal{T} : |u| = t \text{ and } \theta^*(S_u - tx^*) \geq \ell\}.$$

Subpart 1: *The case of* $\alpha > 1$

By (4.1), observe that

$$\mathbb{P}(D < \varepsilon) = \mathbb{P}\Big(\sum_{|u|=t} e^{\theta^*(S_u - tx^*)} D^{(u)} < \varepsilon\Big) \leq \mathbb{P}\big(e^{\theta^*(S_u - tx^*)} D^{(u)} < \varepsilon, \ \forall |u| = t\big). \tag{4.6}$$

We first obtain a rough bound. In fact,

$$\begin{aligned}\mathbb{P}(D < \varepsilon) &\leq \mathbb{P}\big(D^{(u)} < 1, \ \forall u \in U_0(t, \log\varepsilon); \#U_0(t, \log\varepsilon) \geq b^t\big) + \mathbb{P}\big(\#U_0(t, \log\varepsilon) < b^t\big) \\ &\leq \mathbb{P}(D < 1)^{b^t} + \mathbb{P}\Big(Z_t\Big(\Big[\frac{\log\varepsilon}{\theta^*} + tx^*, \infty\Big)\Big) < b^t\Big) \\ &\leq e^{-c_{18} b^t} + \mathbb{P}\Big(\sum_{|u|=t} \mathbf{1}_{\{S_u < \frac{\log\varepsilon}{\theta^*} + tx^*\}} \geq 1\Big),\end{aligned} \tag{4.7}$$

because $\mathbb{P}(D < 1) < 1$ and $Z_t \geq b^t$. By Markov inequality and Chernoff inequality,

$$\mathbb{P}\Big(\sum_{|u|=t} \mathbf{1}_{\{S_u < \frac{\log\varepsilon}{\theta^*} + tx^*\}} \geq 1\Big) \leq \mathbb{E}[Z_t]\mathbb{P}\Big(S_t < \frac{\log\varepsilon}{\theta^*} + tx^*\Big) \leq e^{\theta \frac{\log\varepsilon}{\theta^*} + c_{19} t},$$

for any $\theta > 0$ and $c_{19} = c_{19}(\theta) > 0$. Recall that $a_\varepsilon = -\log\varepsilon$. We take $t = \lfloor 2\log a_\varepsilon / \log b \rfloor$ so that $b^t \gg a_\varepsilon$ and $t \ll a_\varepsilon$ for $\varepsilon > 0$ small enough. As a consequence, for $\varepsilon > 0$ sufficiently small,

$$\mathbb{P}(D < \varepsilon) \leq e^{-2a_\varepsilon}, \tag{4.8}$$

which is a rough upper bound. Now again by (4.6), for any $\delta \in (0, 1)$, $t_\varepsilon \in \mathbb{N}_+$ and $\delta_\varepsilon \in (0, t_\varepsilon) \cap \mathbb{N}$,

$$\begin{aligned}&\mathbb{P}(D < \varepsilon) \\ &\quad\leq \mathbb{P}\Big(\sup_{u \in U_0(t_\varepsilon, (1-\delta)\log\varepsilon)} D^{(u)} < \varepsilon^\delta, \#U_0\big(t_\varepsilon, (1-\delta)\log\varepsilon\big) \geq b^{t_\varepsilon - \delta_\varepsilon}\Big) + \mathbb{P}\big(\#U_0\big(t_\varepsilon, (1-\delta)\log\varepsilon\big) < b^{t_\varepsilon - \delta_\varepsilon}\big) \\ &\quad\leq \mathbb{P}\big(D < \varepsilon^\delta\big)^{b^{t_\varepsilon - \delta_\varepsilon}} + \mathbb{P}\big(\#U_0\big(t_\varepsilon, (1-\delta)\log\varepsilon\big) < b^{t_\varepsilon - \delta_\varepsilon}\big).\end{aligned} \tag{4.9}$$

By (4.8), one sees that

$$\mathbb{P}\big(D < \varepsilon^\delta\big)^{b^{t_\varepsilon - \delta_\varepsilon}} \leq e^{2\delta b^{t_\varepsilon - \delta_\varepsilon} \log\varepsilon}.$$

On the other hand, for the second term on the r.h.s. of (4.9), for $t_\varepsilon = \Theta(\log a_\varepsilon) \ll a_\varepsilon$,

$$\begin{aligned}\mathbb{P}\big(\#U_0\big(t_\varepsilon, (1-\delta)\log\varepsilon\big) < b^{t_\varepsilon - \delta_\varepsilon}\big) &\leq \mathbb{P}\Big(\sum_{|u|=t_\varepsilon} 1_{S_u \geq t_\varepsilon x^* + (1-\delta)\frac{\log\varepsilon}{\theta^*}} < b^{t_\varepsilon - \delta_\varepsilon}\Big) \\ &\leq \mathbb{P}\Big(\sum_{|u|=t_\varepsilon} 1_{S_u \geq -(1-2\delta)\frac{a_\varepsilon}{\theta^*}} < b^{t_\varepsilon - \delta_\varepsilon}\Big),\end{aligned}$$

which by the same arguments as deducing (3.36), is less than

$$\exp\Big\{-(1-\delta)\lambda(b_\alpha - 1)^{\alpha-1}\Big(\frac{a_\varepsilon}{\theta^*}\Big)^\alpha (1-4\delta)^\alpha \big(1 + o_\varepsilon(1)\big)\Big\}.$$

Consequently, (4.9) becomes that

$$\mathbb{P}(D<\varepsilon)\le e^{-2\delta b^{t_\varepsilon-\delta_\varepsilon}a_\varepsilon}+\exp\left\{-\lambda(1-\delta)(b_\alpha-1)^{\alpha-1}\left(\frac{a_\varepsilon}{\theta^*}\right)^\alpha(1-4\delta)^\alpha\big(1+o_\varepsilon(1)\big)\right\}.$$

Let $t_\varepsilon=\lfloor\frac{\alpha-1/3}{\log b}\log a_\varepsilon\rfloor$, $\delta_\varepsilon=\lfloor\frac{1/3}{\log b}\log a_\varepsilon\rfloor$ so that

$$b^{t_\varepsilon-\delta_\varepsilon}a_\varepsilon\gg a_\varepsilon^\alpha\gg b^{t_\varepsilon}\log\big(e^{C_2}Aa_\varepsilon\big),\qquad A^\alpha\ge\frac{2}{\theta^*}(b_\alpha-1)^{\alpha-1}.$$

This implies that for any $\delta\in(0,1/4)$,

$$\limsup_{\varepsilon\downarrow 0}\frac{1}{(-\log\varepsilon)^\alpha}\log\mathbb{P}(D<\varepsilon)\le-(1-\delta)\frac{\lambda}{(\theta^*)^\alpha}(b_\alpha-1)^{\alpha-1}(1-4\delta)^\alpha,$$

which gives the upper bound for the case of $\alpha>1$.

Subpart 2: *The case of* $0<\alpha\le 1$
For $\delta\in(0,1/b)$, similar to (4.7), we have, for any $t_\varepsilon\in(0,a_\varepsilon)\cap\mathbb{N}$,

$$\begin{aligned}\mathbb{P}(D<\varepsilon)&\le\mathbb{P}\big(D^{(u)}<1,\ \forall u\in U_0(t_\varepsilon,\log\varepsilon);\#U_0(t_\varepsilon,\log\varepsilon)\ge\delta b^{t_\varepsilon}\big)+\mathbb{P}\big(\#U_0(t_\varepsilon,\log\varepsilon)<\delta b^{t_\varepsilon}\big)\\&\le\mathbb{P}(D<1)^{\delta b^{t_\varepsilon}}+\mathbb{P}\big(\#U_0(t_\varepsilon,\log\varepsilon)<\delta b^{t_\varepsilon}\big)\\&\le e^{-c_{32}\delta b^{t_\varepsilon}}+\mathbb{P}\left(\sum_{|u|=t_\varepsilon}\mathbf{1}_{S_u\ge t_\varepsilon x^*-\frac{a_\varepsilon}{\theta^*}}<\delta b^{t_\varepsilon}\right).\end{aligned}\tag{4.10}$$

Note that for $t_\varepsilon=\Theta(\log a_\varepsilon)\le\delta' a_\varepsilon$ with some $\delta'\in(0,1)$ and for $0<\varepsilon\ll 1$,

$$\begin{aligned}\mathbb{P}\left(\sum_{|u|=t_\varepsilon}\mathbf{1}_{S_u\ge t_\varepsilon x^*-\frac{a_\varepsilon}{\theta^*}}<\delta b^{t_\varepsilon}\right)&=\mathbb{P}\left(Z_{t_\varepsilon}\left[t_\varepsilon x^*-\frac{a_\varepsilon}{\theta^*},\infty\right)<\delta b^{t_\varepsilon}\right)\\&\le\mathbb{P}\left(Z_{t_\varepsilon}\left[-\big(1-\delta'\big)\frac{a_\varepsilon}{\theta^*},\infty\right)<\delta b^{t_\varepsilon}\right),\end{aligned}$$

which by the same reasonings as (3.12), yields that

$$\mathbb{P}\left(Z_{t_\varepsilon}\left[-\big(1-\delta'\big)\frac{a_\varepsilon}{\theta^*},\infty\right)<\delta b^{t_\varepsilon}\right)\le e^{-\lambda b(1-\delta')^{1+\alpha}(\frac{a_\varepsilon}{\theta^*})^\alpha+\Theta(t_\varepsilon)}.$$

Going back to (4.10), one sees that

$$\mathbb{P}(D<\varepsilon)\le e^{-c_{18}\delta b^{t_\varepsilon}}+e^{-\lambda b(1-\delta')^{1+\alpha}(\frac{a_\varepsilon}{\theta^*})^\alpha+\Theta(t_\varepsilon)}.$$

By taking $t_\varepsilon=\lfloor\frac{2}{\log b}\log a_\varepsilon\rfloor$, one obtains that for any $\delta'\in(0,1)$,

$$\limsup_{\varepsilon\downarrow 0}\frac{1}{(-\log\varepsilon)^\alpha}\log\mathbb{P}(D<\varepsilon)\le-\frac{\lambda b}{(\theta^*)^\alpha}\big(1-\delta'\big)^{1+\alpha}.$$

The the desired upper bound for the case of $\alpha\in(0,1]$ follows obviously.

5. Moderate deviation in Schröder case: Proof of Theorem 1.7

This section is devoted to studying the moderate deviation of $\mathbb{P}(M_n\le m_n-\ell_n)$ in the case where $p_0+p_1>0$. Inspired by the ideas in Böttcher case, we need that there is only one branch up to some generation with this single random walk moving to some lower place and then from this lower place we start a typical branching random walk. However, we can not use this idea to get the upper bound for which we will consider the first time when the population exceeds ℓ_n^3. In fact, this idea is borrowed from [22] where the Lemma 5.1 helps us to couple the branching random walk at the beginning generations with one single random walk.

We first recall some results in the literature, which will be used later. The following result is the well-known Cramér theorem; see Theorem 3.7.4 in [14].

Lemma 5.1. *Under the assumption* (1.2a), *we have for any* $a > 0$, *as* $n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \le -an) = -I(-a). \tag{5.1}$$

Remark 5.1. Note that if X has Weibull left tail with $\alpha \in (0, 1)$, i.e., $\mathbb{P}(X < -z) = \Theta(1)e^{-\lambda z^\alpha}$, Theorem 3 in [21] shows that for any $x > 0$, $\mathbb{P}(S_n \le -n^{1/\alpha}x) = e^{-(\lambda+o_n(1))x^\alpha}$.

The next two statements characterize asymptotic behaviors of lower deviation probability for Galton–Watson process; see Corollary 5 in [19] or Proposition 3 in [20]. Define $b_1 := \min\{k \ge 1 : p_k > 0\}$ and recall $\gamma = \log f'(q)$ and $\mathbb{P}^s(\cdot) = \mathbb{P}(\cdot|\mathcal{T} = \infty)$.

Lemma 5.2. *Assume* (1.1) *and* $0 < p_0 + p_1 < 1$. *Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^s(Z_n = b_1) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Z_n = b_1) = \gamma, \tag{5.2}$$

with b_1 *the minimal positive offspring number, and for every subexponential sequence* a_n *with* $a_n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^s(Z_n \le a_n) = \gamma. \tag{5.3}$$

We also have the following fact whose proof can e.g. be found in Lemma 1.2.15 in [14]. For $i \ge 1$, let $(a_n^i)_{n\ge1}$ be a sequence of positive numbers and $a^i = \limsup_{n\to\infty} \frac{1}{n} \log a_n^i$. Then, for all $k \ge 2$ it holds that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{k} \log a_n^i = \max_{i\in\{1,\dots,k\}} a^i. \tag{5.4}$$

5.1. *Lower bound*

For the lower bound, we consider the case that there are only b_1 particles at some generation t_n, and the random walk of one of those b_1 particles moves to the level $-at_n$. Furthermore, families induced by other $b_1 - 1$ particles at t_nth generation die out before time n.

Recall that $\ell^* = \lim_{n\to\infty} \frac{\ell_n}{n} \in [0, \infty)$. For any $\varepsilon > 0$ and $y \ge (x^* - \ell^*) \vee 0$ such that $a = \ell^* - x^* + 2\varepsilon + y > 0$, let $t_n = \lceil \frac{\ell_n}{\ell^*+y+\varepsilon} \rceil$. Note that $t_n < n$ for n large enough. By using Markov property at time t_n, we have for n large enough,

$$\begin{aligned}
&\mathbb{P}^s(M_n \le m_n - \ell_n) \\
&\quad \ge \mathbb{P}^s\Big(Z_{t_n} = b_1; \exists |u| = t_n, S_u \le -at_n, S_u + M^u_{n-t_n} \le m_n - \ell_n, \bigcup_{|v|=t_n, v\neq u} \{|w| = n : v \prec w\} = \varnothing\Big) \\
&\quad \ge \mathbb{P}(Z_{t_n} = b_1)\mathbb{P}(S_{t_n} \le -at_n)\mathbb{P}^s(M_{n-t_n} \le m_n + at_n - \ell_n)\mathbb{P}(Z_{n-t_n} = 0|Z_0 = b_1 - 1) \\
&\quad \ge \mathbb{P}(Z_{t_n} = b_1)\mathbb{P}(S_{t_n} \le -at_n)\mathbb{P}^s(M_{n-t_n} \le m_n + at_n - \ell_n)(q/2)^{b_1-1},
\end{aligned} \tag{5.5}$$

where in the last inequality we use the fact that $\lim_{n\to\infty} \mathbb{P}(Z_{n-t_n} = 0|Z_0 = b_1 - 1) = q^{b_1-1}$. Recall that $m_n = x^*n - \frac{3}{2\theta^*} \log n$. Then one can check for n large enough,

$$m_n + at_n - \ell_n - m_{n-t_n} \ge \varepsilon t_n + \frac{3}{2\theta^*} \log\left(\frac{n - t_n}{n}\right) \ge 0.$$

Thus

$$\liminf_{n\to\infty} \mathbb{P}^s(M_{n-t_n} \le m_n + at_n - \ell_n) > 0$$

and then for n large enough,

$$\mathbb{P}^s(M_n \le m_n - \ell_n) \ge c_{20}\mathbb{P}(Z_{t_n} = b_1)\mathbb{P}(S_{t_n} \le -at_n). \tag{5.6}$$

Combining it with (5.1) and (5.2) yields that

$$\liminf_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s(M_n \le m_n - \ell_n) \ge \frac{-I(-a) - \gamma}{\ell^* + y + \varepsilon}.$$

Letting $\varepsilon \downarrow 0$, together with the fact that l.h.s. is independent of y, gives

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s(M_n \le m_n - \ell_n) &\ge \sup_{y \ge (x^*-\ell^*)\vee 0} \frac{-I(x^* - \ell^* - y) + \gamma}{\ell^* + y} \\ &= \sup_{a \ge \ell^* \vee x^*} \frac{\gamma - I(x^* - a)}{a}.\end{aligned}$$

5.2. *Upper bound*

For simplicity, write Z_t for $Z_{\lfloor t \rfloor}$ and $t \ge 0$. Let

$$T_n = \inf\{t \ge 0 : Z_{t\ell_n} \ge \ell_n^3\}$$

and for $\delta > 0$ and $\varepsilon > 0$ small enough set

$$F(\delta) = \left\{\delta, 2\delta, \ldots, \frac{1}{(\ell^* \vee x^*)(1 + 2\varepsilon)}\right\}.$$

Then

$$\begin{aligned}&\mathbb{P}^s(M_n \le m_n - \ell_n) \\ &\quad\le \mathbb{P}^s\big(Z_{\frac{\ell_n}{(\ell^*\vee x^*)(1+2\varepsilon)}} \le \ell_n^3\big) + \sum_{t\in F(\delta)} \mathbb{P}^s\big(M_n \le m_n - \ell_n; T_n \in (t - \delta, t]\big).\end{aligned} \tag{5.7}$$

Note that by (5.3), we have

$$\lim_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s\big(Z_{\frac{\ell_n}{(\ell^*\vee x^*)(1+2\varepsilon)}} \le \ell_n^3\big) = \frac{\gamma}{(\ell^* \vee x^*)(1 + 2\varepsilon)} \tag{5.8}$$

and

$$\limsup_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s\big(T_n \in (t - \delta, t]\big) \le \lim_{n\to\infty} \frac{1}{\ell_n} \log \mathbb{P}^s\big(Z_{(t-\delta)\ell_n} \le \ell_n^3\big) = \gamma(t - \delta). \tag{5.9}$$

Meanwhile,

$$\begin{aligned}&\mathbb{P}^s\big(M_n \le m_n - \ell_n | T_n \in (t - \delta, t]\big) \\ &\quad= \mathbb{P}^s\Big(\max_{|u|=t\ell_n} \big(S_u + M^u_{n-t\ell_n}\big) \le m_n - \ell_n | T_n \in (t - \delta, t]\Big) \\ &\quad\le \mathbb{P}^s\Big(S_{t\ell_n} + \max_{|u|=t\ell_n} M^u_{n-t\ell_n} \le m_n - \ell_n | T_n \in (t - \delta, t]\Big) \\ &\quad\le \mathbb{P}\big(S_{t\ell_n} \le m_n - (1 - \varepsilon)\ell_n - m_{n-t\ell_n}\big) + \mathbb{P}^s\Big(\max_{|u|=t\ell_n} M^u_{n-t\ell_n} \le m_{n-t\ell_n} - \varepsilon\ell_n | T_n \in (t - \delta, t]\Big) \\ &\quad=: I_1 + I_2,\end{aligned}$$

where in the first inequality, we use Lemma 5.1 in [22] and the fact that (S_u) and $(M^u_{n-t\ell_n})$ are independent. We first estimate I_1. For any $t \in F(\delta)$, one can check that $tx^* - 1 + \varepsilon < 0$ and

$$m_n - (1-\varepsilon)\ell_n - m_{n-t\ell_n} = \frac{3}{2\theta^*}\log\left(\frac{n-t\ell_n}{n}\right) + \left(tx^* - 1 + \varepsilon\right)\ell_n \le \left(tx^* - 1 + \varepsilon\right)\ell_n.$$

Thus

$$\limsup_{n\to\infty} \frac{1}{\ell_n}\log \mathbb{P}\left(S_{t\ell_n} \le m_n - (1-\varepsilon)\ell_n - m_{n-t\ell_n}\right) \le -tI\left(\frac{tx^* - 1 + \varepsilon}{t}\right).$$

Next, we turn to I_2.

$$I_2 = \mathbb{E}^s\left[\mathbb{P}^s(M_{n-t\ell_n} \le m_{n-t\ell_n} - \varepsilon\ell_n)^{Z_{t\ell_n}} | T_n \in (t-\delta, t]\right] \le \mathbb{P}^s(M_{n-t\ell_n} \le m_{n-t\ell_n} - \varepsilon\ell_n)^{\ell_n^2} + \mathbb{P}^s\left(Z_{t\ell_n} \le \ell_n^2 | T_n \in (t-\delta, t]\right).$$

Notice that as $\mathbb{P}^s(T_n \in (t-\delta, t]) \ge \frac{1-q^{n^3}}{1-q}\mathbb{P}(T_n \in (t-\delta, t])$, we have

$$\mathbb{P}^s\left(Z_{tn} \le \ell_n^2 | T_n \in (t-\delta, t]\right) \le \left(1 - q + o(1)\right)\mathbb{P}\left(\exists k \le \delta n,\ Z_k \le \ell_n^2 | Z_0 = \ell_n^3\right) \le \binom{\ell_n^3}{\ell_n^2} q^{\ell_n^3 - \ell_n^2}.$$

And by (1.6), one sees that for all n sufficiently large,

$$\mathbb{P}^s(M_{n-t\ell_n} \le m_{n-t\ell_n} - \varepsilon\ell_n) \le e^{-c_{21}} < 1,$$

with some $c_{21} > 0$. Thus $I_2 \le e^{-c_{21}\ell_n^2}$ and hence

$$\limsup_{n\to\infty} \frac{1}{\ell_n}\log \mathbb{P}^s\left(M_n \le m_n - \ell_n | T_n \in (t-\delta, t]\right) \le -tI\left(\frac{tx^* - 1 + \varepsilon}{t}\right). \tag{5.10}$$

Going back to (5.7), together with (5.8), (5.9) and (5.4), one has

$$\limsup_{n\to\infty} \frac{1}{\ell_n}\log \mathbb{P}^s(M_n \le m_n - \ell_n) \le \frac{\gamma}{(\ell^* \vee x^*)(1+2\varepsilon)} \vee \sup_{t\in F(\delta)}\left((t-\delta)\gamma - tI\left(\frac{tx^* - 1 + \varepsilon}{t}\right)\right),$$

which by letting $\varepsilon \downarrow 0$ and $\delta \downarrow 0$ implies

$$\limsup_{n\to\infty} \frac{1}{\ell_n}\log \mathbb{P}^s(M_n \le m_n - \ell_n) \le \sup_{t\in(0, \frac{1}{\ell^*\vee x^*})}\left(t\gamma - tI\left(\frac{tx^* - 1}{t}\right)\right) = \sup_{y \ge (x^* - \ell^*)_+} \frac{-I(x^* - \ell^* - y) + \gamma}{\ell^* + y}. \tag{5.11}$$

We have completed the proof.

6. Discussions

In this paper, we study the branching random walk by assuming that the branching and the motions are independent. However, things will be more complicated in the general setting where the reproduction law is given by a point process representing the displacements of children of one individual. Let us give an example here. Suppose that the point process is $\sum_{|u|=1} \delta_{X_u} = \delta_X \mathbf{1}_{\{X \ge 0\}} + 2\delta_X \mathbf{1}_{\{X<0\}}$ with X a random variable such that $\mathbb{P}(X \ge 0) \in (0,1)$ and $\mathbb{E}[e^{tX}] < \infty$ for any $t \in \mathbb{R}$. Obviously, the branching random walk generated by the law of this point process is in Schröder case with $b_1 = 1$. However, for the lower bound of $\mathbb{P}(M_n \le m_n - \ell_n)$, the strategy for the Schröder case does not work: if there is one single individual at time t_n, it has to move to some positive level. One may instead apply the arguments in the Böttcher case. So, it seems to have more regimes in the general setting.

Acknowledgement

We are grateful to Elie Aïdékon, Yueyun Hu and Bastien Mallein for enlightening discussions. We also would like to thank the referees for carefully reading our paper, spotting several inaccuracies and for making valuable suggestions to improve the presentation. Xinxin Chen is supported by ANR MALIN. Hui He is supported by NSFC (No. 11671041, 11531001).

References

[1] L. Addario-Berry and B. Reed. Minima in branching random walks. *Ann. Probab.* **37** (2009) 1044–1079. MR2537549 https://doi.org/10.1214/08-AOP428
[2] E. Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.* **41** (2013) 1362–1426. MR3098680 https://doi.org/10.1214/12-AOP750
[3] E. Aïdékon, Y. Hu and Z. Shi. Large deviations for level sets of branching Brownian motion and Gaussian free fields. *Zap. Naučn. Semin. POMI* **457** (2017) 12–36. MR3723574
[4] A. Bhattacharya. Large deviation for extremes in branching random walk with regularly varying displacements. Preprint, 2018. Available at https://arxiv.org/abs/1802.05938.
[5] J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Adv. in Appl. Probab.* **8** (1976) 446–459. MR0420890 https://doi.org/10.2307/1426138
[6] J. D. Biggins. Branching out. In *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*. Cambridge University Press, Cambridge, 2010. MR2744237
[7] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Adv. in Appl. Probab.* **36** (2004) 544–581. MR2058149 https://doi.org/10.1239/aap/1086957585
[8] M. D. Bramson. Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31** (1978) 531–581. MR0494541 https://doi.org/10.1002/cpa.3160310502
[9] M. D. Bramson, J. Ding and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.* **69** (2016) 62–123. MR3433630 https://doi.org/10.1002/cpa.21621
[10] M. D. Bramson, J. Ding and O. Zeitouni. Convergence in law of the maximum of nonlattice branching random walk. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 1897–1924. MR3573300 https://doi.org/10.1214/15-AIHP703
[11] D. Buraczewski and M. Maślanka. Large deviation estimates for branching random walks. *ESAIM Probab. Stat.* **23** (2019) 823–840. MR4045541 https://doi.org/10.1051/ps/2019006
[12] 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) 299–314. MR0968823 https://doi.org/10.1007/BF00356108
[13] X. Chen and H. He. On large deviation probabilities for empirical distribution of super-critical branching random walks with unbounded displacements. *Probab. Theory Related Fields* **175** (2019) 255–307. MR4009709 https://doi.org/10.1007/s00440-018-0891-4
[14] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer-Verlag, Berlin, 1998. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
[15] B. Derrida and Z. Shi. Large deviations for the branching Brownian motion in presence of selection or coalescence. *J. Stat. Phys.* **163** (2016) 1285–1311. MR3503265 https://doi.org/10.1007/s10955-016-1522-z
[16] B. Derrida and Z. Shi. Large deviations for the rightmost position in a branching Brownian motion. In *Modern Problems of Stochastic Analysis and Statistics. MPSAS 2016*, V. Panov (Ed.) *Springer Proceedings in Mathematics & Statistics* **208**. Springer, Cham, 2017. MR3747671
[17] B. Derrida and Z. Shi. Slower deviations of the branching Brownian motion and of branching random walks. *J. Phys. A* **50** (2017), 344001. MR3681502 https://doi.org/10.1088/1751-8121/aa7f98
[18] R. Durrett. *Probability: Theory and Examples (Fouth Edition)*. Cambridge University Press, Cambridge, 2010. MR2722836 https://doi.org/10.1017/CBO9780511779398
[19] K. Fleischmann and V. Wachtel. Lower deviation probabilities for supercritical Galton–Watson processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **43** (2007) 233–255. MR2303121 https://doi.org/10.1016/j.anihpb.2006.03.001
[20] K. Fleischmann and V. Wachtel. Large deviations for sums indexed by the generations of a Galton–Watson process. *Probab. Theory Related Fields* **141** (2008) 445–470. MR2391161 https://doi.org/10.1007/s00440-007-0090-1
[21] N. Gantert. The maximum of a branching random walk with semiexponential increments. *Ann. Probab.* **28** (2000) 1219–1229. MR1797310 https://doi.org/10.1214/aop/1019160332
[22] N. Gantert and T. Höfelsauer. Large deviations for the maximum of a branching random walk. *Electron. Commun. Probab.* **23** (2018) 12 pp. MR3812066 https://doi.org/10.1214/18-ECP135
[23] D. R. Grey. Explosiveness of age-dependent branching processes. *Z. Wahrsch. Verw. Gebiete* **28** (1973/74) 129–137. MR0359040 https://doi.org/10.1007/BF00533364
[24] J. M. Hammersley. Postulates for subadditive processes. *Ann. Probab.* **2** (1974) 652–680. MR0370721 https://doi.org/10.1214/aop/1176996611
[25] Y. Hu. How big is the minimum of a branching random walk? *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 233–260. MR3449302 https://doi.org/10.1214/14-AIHP651

[26] Y. Hu and Z. Shi. A subdiffusive behaviour of recurrent random walk in random environment on a regular tree. *Probab. Theory Related Fields* **138** (2007) 521–549. MR2299718 https://doi.org/10.1007/s00440-006-0036-z
[27] 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) 742–789. MR2510023 https://doi.org/10.1214/08-AOP419
[28] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** (1975) 790–801. MR0400438 https://doi.org/10.1214/aop/1176996266
[29] S. P. Lalley and Y. Shao. On the maximal displacement of critical branching random walk. *Probab. Theory Related Fields* **162** (2015) 71–96. MR3350041 https://doi.org/10.1007/s00440-014-0566-8
[30] Q. Liu. Fixed points of a generalised smoothing transformation and applications to the branching random walk. *Adv. in Appl. Probab.* **30** (1998) 85–112. MR1618888 https://doi.org/10.1239/aap/1035227993
[31] Q. Liu. Asymptotic properties of supercritical age-dependent branching processes and homogeneous branching random walks. *Stochastic Process. Appl.* **82** (1999) 61–87. MR1695070 https://doi.org/10.1016/S0304-4149(99)00008-3
[32] Q. Liu. Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stochastic Process. Appl.* **95** (2001) 83–107. MR1847093 https://doi.org/10.1016/S0304-4149(01)00092-8
[33] Q. Liu. On generalised multiplicative cascades. *Stochastic Process. Appl.* **86** (2000) 263–286. MR1741808 https://doi.org/10.1016/S0304-4149(99)00097-6
[34] T. Madaule. The tail distribution of the Derivative martingale and the global minimum of the branching random walk. Preprint, 2016. Available at https://arxiv.org/abs/1606.03211v2.
[35] E. Neuman and X. Zheng. On the maximal displacement of subcritical branching random walks. *Probab. Theory Related Fields* **167** (2017) 1137–1164. MR3627435 https://doi.org/10.1007/s00440-016-0702-8
[36] A. Rouault. Precise estimates of presence probabilities in the branching random walk. *Stochastic Process. Appl.* **44** (1993) 27–39. MR1198661 https://doi.org/10.1016/0304-4149(93)90036-4
[37] Z. Shi. Branching random walks. In *École d'Été de Probabilités de Saint-Flour XLII-2012*. *Lecture Notes in Mathematics* **2151**. Springer, Cham, 2015. MR3444654 https://doi.org/10.1007/978-3-319-25372-5
[38] O. Zeitouni. *Branching Random Walks and Gaussian Fields. Probability and Statistical Physics in St. Petersburg*. *Proc. Sympos. Pure Math.* **91**, 437–471. Amer. Math. Soc., Providence, RI, 2016. MR3526836