Branching Random Walks with Selection

Patrik Gerber

Supervised by Prof. Julien Berestycki

Dissertation on a topic in Statistics
presented for MMath in Mathematics and Statistics

February 10, 2019

UNIVERSITY OF OXFORD

DEPARTMENT OF STATISTICS

Contents

1 Speed

Placeholder text.

In this essay we study Branching Random Walks (BRWs) with selection (also called N-branching random walks), which we can think of as a dynamic cloud of particles on $\mathbb{R}$ indexed by discrete time. Branching random walks with selection evolve according to two mechanisms

1 **branching** Each particle gives birth to its offspring around itself, according to some point process.

2 **selection** Out of all children of the current generation, the rightmost N are selected to form the next generation.

It will be convenient to think of BRWs and N-BRWs as stochastic processes taking values in the set $\mathfrak{M}$ of counting measures μ on $\mathbb{R}$ which put non-negative integer mass on every atom and further satisfy $\mu([x,\infty)) < \infty$ for all $x \in \mathbb{R}$. The latter condition is needed for the phrase 'rightmost particles' to be meaningful. We will write $\mathfrak{M}_N \subset \mathfrak{M}$ for measures which have total mass N and $\delta_{x_0} \in \mathfrak{M}_1$ for the unit mass at x_0. The interpretation is that if μ is the value of the (N-)BRW at some time n, then there are exactly $\mu(\{x\})$ particles at position x at time n. There is a natural partial order on $\mathfrak{M}$: we say that $\mu \preceq \nu$ if $\mu([x,\infty)) \leq \nu([x,\infty))$ for all $x \in \mathbb{R}$. Naturally, for random elements $\mathscr{L}, \mathscr{G}$ of $\mathfrak{M}$ (such as BRWs) we say that $\mathscr{L} \preceq \mathscr{G}$ if there exists a coupling $(\mathscr{L}, \mathscr{G})$ such that $\mathscr{L} \preceq \mathscr{G}$ almost surely. For $\mathscr{L} \in \mathfrak{M}$ we also define $\{l \mid l \in \mathscr{L}\} := \{I \subset \mathbb{R} \mid \sum_{i \in I} \delta_i = \mathscr{L}\}$.

Using the notation introduced above, we can construct N-branching random walks in great generality. Suppose that $\mathscr{L}$ is a random element of $\mathfrak{M}$ and that $X := (X_n)_{n \geq 0}$ is an N-branching random walk evolving according to the law of $\mathscr{L}$. Then X is inductively constructed as follows: given $X_n \in \mathfrak{M}_N$ for $n \geq 0$, take N i.i.d. copies $(\mathscr{L}_i)_{i=1}^N$ of $\mathscr{L}$ independently of X_n. Writing $X_k(1) \leq \cdots \leq X_k(N)$ for the particles of X_k for all $k \in \mathbb{N}$, we let

$$\tilde{X}_{n+1} = \sum_{i=1}^{N} \sum_{l \in \mathscr{L}_i} \delta_{X_n(i)+l}, \tag{1.1}$$

and define X_{n+1} to be the rightmost N particles in $\tilde{X}_{n+1}$. This construction allows for a natural and important coupling between (N-)BRWs. This coupling was first described in [**?**], the way we present it here is more general and similar to [3] Lemma 4.1.

Lemma 1.1 — *Let $1 \leq N_1 \leq N_2$ and $\mu_i \in \mathfrak{M}_{N_i}$ for $i = 1, 2$. Consider random elements $\mathscr{L}_i \in \mathfrak{M}_{N_i}$ for $i = 1, 2$. Then if $(X_n^{(i)})_{n \geq 0}$ is a(n) (N-)BRW which evolves according to the law of $\mathscr{L}$ and starts from μ_i respectively, then there exists a coupling such that $X_n^{(1)} \preceq X_n^{(2)}$ almost surely for all $n \geq 0$.*

Sketch of proof. We construct the coupling inductively. Given $X_n^{(1)} \preceq X_n^{(2)}$, independently take N_2 i.i.d. copies $\{(\mathscr{L}_i^{(1)}, \mathscr{L}_i^{(2)})\}_{i=1}^{N_2}$ of the coupling of $\mathscr{L}_1$ and $\mathscr{L}_2$ that witnesses the partial order. Using these, construct $\tilde{X}_{n+1}^{(1)}$ and $\tilde{X}_{n+1}^{(2)}$ as in (1.1). If the $X^{(i)}$ are regular BRWs just set $X_{n+1}^{(i)} = \tilde{X}_{n+1}^{(i)}$, if they are N-BRWs take the rightmost N-particles like before. Either way, we have $X_{n+1}^{(1)} \preceq X_{n+1}^{(2)}$ as desired. □

1.1 Exponentially decaying tails

1.1.1 Construction

The first variation of the N-branching random walk that we consider is very similar to the one studied in [1] by Bérard and Gouéré. However, we treat a slightly more general case where the number of offspring of each particle is random as opposed to being fixed at two. In this version of the N-branching random walk each particle dies and gives birth to a random number of offspring whose number is distributed like q. Given the position of the parent, say x, each child's position follows the law $p(\,\cdot - x)$ independently of the number and position of the other children.

Construction. Let $X = (X_n)_{n\geq 0} = (\sum_{i=1}^N \delta_{X_n(i)})_{n\geq 0}$ denote the $\mathfrak{M}_N$-valued discrete time Markov process defined by the branching-selection procedure detailed above. Note that we suppress the dependence on N in our notation for simplicity. We can construct X easily: Let $\mathcal{E}_N := (\epsilon_{l,i,j})_{l\geq 0,\, i\in[N],\, j\geq 1}$ and $\mathcal{M}_N := (\tau_{l,i})_{l\geq 0, i\in[N]}$ be i.i.d. collections of random variables distributed like p and q respectively, with the collections also independent from each other. Now, given the process up to time $n \geq 0$, we construct X_{n+1} as follows: define $Y_{n+1} := \sum_{i=1}^N \sum_{j=1}^{\tau_{n,i}} \delta_{X_n(i)+\epsilon_{n,i,j}}$ and take X_{n+1} to given by the N rightmost particles of Y_{n+1}.

Let $\nu \in \mathfrak{M}$ be a random, finite counting measure with the same distribution as the offspring of a single particle at the origin in our branching-selection mechanism (the fact that $\nu \in \mathfrak{M}$ follows from Assumption 3). In other words, the number of atoms of ν has distribution q and each atom is placed independently at position drawn from p. Let us now define the logarithmic moment generation function of ν:

$$\psi(t) := \log \mathbb{E} \int_{\mathbb{R}} e^{tx} d\nu(x).$$

Note that in their analysis Bérard and Gouéré define a slightly different function $\Lambda(t) = \psi(t) - \log 2$, however the branching random walk literature usually uses our definition. We impose the following assumptions to gain access to the results of [2]:

Assumption 1. ψ is finite in some neighbourhood of 0.

Assumption 2. There exists $t^* > 0$ in the interior of the domain of ψ such that $t^*\psi'(t^*) = \psi(t^*)$.

Assumption 1 is in fact equivalent to the requirement that p have exponentially decaying tails, furthermore it implies that p has finite moments of all orders. The third assumption concerns the distribution q:

Assumption 3. q satistfies $q(0) = 0$ and $1 < \sum_{i=1}^\infty i^2 q(i) < \infty$.

The results that follow in this section are conditional upon Assumptions 1, 2 and 3 being satisfied. We now record a technical lemma that will help us later.

Lemma 1.2 — *Let $\tau \in L^1$ be an $\mathbb{N}$-valued random variable and let $(\epsilon_n)_{n\geq 1}$ be an i.i.d. sequence of random variables with exponentially decaying tails, independent of τ. Then $M := \max_{1\leq n\leq \tau} \epsilon_n$ has exponentially decaying tails.*

Proof. Let $C, \gamma, t_0 > 0$ be such that $\mathbb{P}(|\epsilon_1| \leq t) \geq 1 - Ce^{-\gamma t}$ for all $t > t_0$. Then for $t > t_0$ large enough, Bernoulli's inequality gives

$$\begin{aligned}\mathbb{P}(M > t) &\leq 1 - \mathbb{E}\left[\mathbb{P}(|\epsilon_1| \leq t)^\tau\right] \leq 1 - \mathbb{E}\left[(1 - Ce^{-\gamma t})^\tau\right] \\ &\leq 1 - \mathbb{E}\left[1 - Ce^{-\gamma t}\tau\right] = \underbrace{C\,\mathbb{E}[\tau]}_{<\infty} e^{-\gamma t}.\end{aligned}$$

Similarly, looking at the lower tail we get

$$\mathbb{P}(M < -t) \leq 1 - \mathbb{E}\left[\mathbb{P}(|\epsilon_1| \leq t)^\tau\right] \leq C\,\mathbb{E}[\tau]\, e^{-\gamma t}.$$

□

1.1.2 Properties of the model

Denote by $\max X_n$ and $\min X_n$ the position of the right- and leftmost particle of X_n respectively. It is worth noting that $\min X_n$ and $\max X_n$ are integrable and hence finite by Assumptions 1 and 3 when started from any fixed $X_0 \in \mathfrak{M}_N$. Indeed, by independence we have

$$\mathbb{E}|\max X_n| \leq \mathbb{E}\left|\max X_0 + \sum_{l=0}^{n-1}\sum_{i=1}^{N}\sum_{j=1}^{\tau_{l,i}} \epsilon_{l,i,j}\right| \leq |\max X_0| + Nn\mathbb{E}[\tau_{0,1}]\,\mathbb{E}|\epsilon_{0,1,1}|. \tag{1.2}$$

Denote by $d(X_n) := \max X_n - \min X_n$ the diameter of X_n. We have the following result, analogous to Corollary 1 of [1]:

Proposition 1.1 — *For any $N \geq 1$ and initial population $X_0 \in \mathfrak{M}_N$, we have*

$$\frac{d(X_n)}{n} \xrightarrow[n\to\infty]{a.s.,\, L^1} 0.$$

Proof. Let $u \in \mathbb{N}_+$ and for $n \geq u$ consider the process X in the timeframe $[\![n-u, n]\!]$. Define $\mathcal{E} := \{\epsilon_{l,i,j} \mid l \in [\![n-u, n-1]\!],\, i \in [N],\, j \in [\tau_{l,i}]\}$ and let $M := \max \mathcal{E}$, $m := \min \mathcal{E}$ noting that both have exponentially decaying tails by Lemma 1.2. Write $y := \max X_{n-u}$ for the rightmost particle's position at time $n-u$. Suppose that for each $k \in [u]$ we have $\min X_{n-u+k} < y + km$. As all steps during branching are $\geq m$, this implies in particular that the descendants of the particle 'y' survive all selection steps until time n. Therefore, on the event $A_u :=$ $\{$number of descendants of y at time n is $> N\}$ almost surely $\min X_{n-u+k} \geq y + k_0 m$ for some k_0. By the definition of m this must also hold for all $k \in [\![k_0, u]\!]$, in particular for $k = u$. Noting that $\max X_n \leq y + uM$, it follows that

$$d(X_n)\mathbb{1}_{A_u} \leq u(M-m), \tag{1.3}$$

with probability one. A simple argument shows that $\mathbb{1}_{A_u} \to 1$ almost surely as $u \uparrow \infty$: take any path of length u started from 'y'. On A_u^c, along any such path the number of times that the corresponding particle has more than one child is less than N.

$$\mathbb{P}\left(A_u^c\right) \leq \sum_{k=0}^{N-1} \binom{u}{k} q(1)^{u-k}(1-q(1))^k \leq N u^{N-1} q(1)^{u-(N-1)} \to 0 \tag{1.4}$$

as $u \uparrow \infty$ since $q(1) < 1$. Fix $\epsilon > 0$ and take u large enough so that $\mathbb{P}\left(A_u^c\right) < \epsilon^2$. Consider the decomposition

$$\frac{d(X_n)}{n} = \frac{d(X_n)}{n}\mathbb{1}_{A_u} + \frac{d(X_n)}{n}\mathbb{1}_{A_u^c}. \tag{1.5}$$

Taking expectations and then taking n to infinity, the first term vanishes by (1.3). The second term is upper bounded by $(\mathbb{P}\left(A_u^c\right)\mathbb{E}\left[d(X_n)^2/n^2\right])^{1/2}$ using Hölder's inequality. A rough bound on $d(X_n)$ suffices now: at each branching step $l \geq 0$ take the maximum and the minimum of the $\sum_{j=1}^{N} \tau_{l,j}$ random walk steps. The diameter certainly grows by no more than the difference between these two at each step. By Lemma 1.2 this yields $\mathbb{E}\left[d(X_n)^2\right] = \mathcal{O}(n^2)$ which implies that the second term in 1.5 is $\mathcal{O}(\epsilon)$. Taking ϵ to zero concludes the proof of L^1 convergence. Almost sure convergence is a consequence of the proof of the next Proposition. □

Proposition 1.2 ([1, Proposition 2]) — *There exists $v_N = v_N(p) \in \mathbb{R}$ such that for any initial population $X_0 \in \mathfrak{M}_N$ the following holds almost surely and in L^1:*

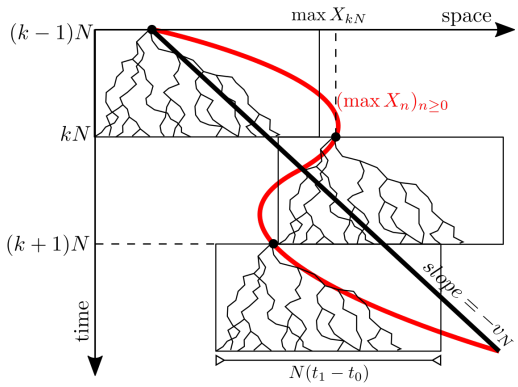
$$\lim_{n\to\infty} \frac{\min X_n}{n} = \lim_{n\to\infty} \frac{\max X_n}{n} = v_N. \tag{1.6}$$

Proof. First we treat the case $X_0 = N\delta_0$. Recall the definition of $\mathcal{E}_N$ and $\mathcal{M}_N$ from the construction of X. For each $l \geq 0$ we define the process $(X_n^l)_{n\geq 0}$ by shifting the origin of time by l. More precisely, given the process up to time $n \geq 0$, define X_{n+1}^l to be given by the N rightmost particles of $\sum_{i=1}^{N}\sum_{j=1}^{\tau_{n+l,i}} \delta_{X_n^l(i)+\epsilon_{n+l,i,j}}$. It is clear that each $(X_n^l)_{n\geq 0}$ is distributed as the N-branching random walk with offspring law p. Start $(X_n^l)_{n\geq 0}$ from $N\delta_0$ for each $l \geq 0$ so that $(X_n^0)_{n\geq 0} = (X_n)_{n\geq 0}$ almost surely. From Lemma 1.1 it follows easily that

$$\max X_{n+m}^0 \leq \max X_n^0 + \max X_m^n \qquad \forall\, n, m \geq 0. \tag{1.7}$$

For notational simplicity define $Y_{i,j} = \max X_{j-i}^i$ for $0 \leq i \leq j$. Then (1.7) reads $Y_{0,j} \leq Y_{0,i} + Y_{i,j}$ for all $0 \leq i \leq j$, which is familiar territory for Kingman's Subadditive Ergodic Theorem. We postpone showing that the conditions of the theorem hold to Lemma 1.3. Applying the theorem yields $\lim_{n\to\infty} n^{-1}\max X_n = \lim_{n\to\infty}\mathbb{E}\left[n^{-1}\max X_n\right] = \inf_n \mathbb{E}\left[n^{-1}\max X_n\right] = v_N \in \mathbb{R}$ where the first limit is almost sure. Noting that the process $(-X_n)_{n\geq 0}$ satisfies all the same assumptions as X, we can deduce from the identity $\min X_n = -\max(-X_n)$ that $\lim_{n\to\infty} n^{-1}\min X_n = \lim_{n\to\infty}\mathbb{E}\left[n^{-1}\min X_n\right] = \inf_n \mathbb{E}\left[n^{-1}\min X_n\right] = \tilde{v}_N \in \mathbb{R}$ exists too, where the first limit is almost sure. From the proof of Proposition 1.1 we immediately get $\tilde{v}_N = v_N$ by uniqueness of L^1 limits, which gives $\lim_{n\to\infty} n^{-1} d(X_n) = v_N - \tilde{v}_N = 0$ almost surely as claimed. The proof is complete in the case $X_0 = N\delta_0$. By translation invariance of the dynamics of the system the result also follows for initial conditions of the form $N\delta_{x_0}$ for any $x_0 \in \mathbb{R}$. Finally, for arbitrary $X_0 \in \mathfrak{M}_N$ note that the result is a consequence of Lemma 1.1 and a sandwiching argument between the initial configurations $N\delta_{\min X_0}$ and $N\delta_{\max X_0}$. □

If we look at the previous proof, we see that the existence of v_N and $\tilde{v}_N$ (the almost sure and L^1 limits of the left- and rightmost particles) when started from $X_0 = N\delta_0$ was shown without relying on Proposition 1.1. We can in fact deduce Proposition 1.1 by an argument inspired by one of Prof. Berestycki's suggestions:



Alternative proof of Proposition 1.1. Let $Y = (Y_n)_{n\geq 0}$ be a branching random walk (without selection) with offspring distribution q and step distribution p. Start Y from δ_0 noting that initially there is only one particle. It is easy to see that the probability ρ_1 that the number of particles in Y_n has reached N by time N is strictly positive. Similar to the proof of Proposition 1.1, define $\mathcal{E} := \{\epsilon_j \mid 1 \leq i \leq \sum_{i=1}^{N^2} \tau_i,\ (\epsilon_j)_{j\geq 0} \overset{iid}{\sim} p \perp \tau_i \overset{iid}{\sim} q\}$ and $M := \min \mathcal{E}, m := \max \mathcal{E}$, where we should think of $\mathcal{E}$ as the set of possible random walk steps that Y can take up to time N. By Lemma 1.2 we can choose $t_0 < t_1$ such that $\rho_2 := \mathbb{P}(t_0 \leq m \leq M \leq t_1) > 0$. We can now write

$$\begin{aligned}
&\mathbb{P}(\{Y_N \text{ has } N \text{ particles}\} \cap \{Nt_1 \geq \max Y_N \geq \min Y_N \geq Nt_0\}) = \\
&\qquad = \mathbb{P}(Y_N \text{ has } N \text{ particles})\, \mathbb{P}(Nt_1 \geq \max Y_N \geq \min Y_N \geq Nt_0 | \, Y_N \text{ has } N \text{ particles}) \\
&\qquad \geq \rho_1\rho_2 > 0. \qquad (1.8)
\end{aligned}$$

Suppose that we couple X with $((Y_n^{(k)})_{0\leq n\leq N})_{k\geq 0}$ which are independent copies of Y placed at the space-time points $(\max X_{kN}, kN)_{k\geq 0}$. By the second Borel-Cantelli lemma and (1.8) it follows that almost surely infinitely many of the $(Y_n^{(k)})_{0\leq n\leq N}$ must have N particles by time N and have $Nt_0 \leq \min Y_N^{(k)} \leq \max Y_N^{(k)} \leq Nt_1$. This in turn implies that for infinitely many $k \geq 0$ the diameter $d(X_{kN})$ is less than $N(t_1 - t_0)$, which immediately yields $\tilde{v}_N = v_N$. □

Proposition 1.3 ([1, analogue of Proposition 3]) — *The sequence $(v_N)_{N\geq 1}$ is non-decreasing.*

Proof. This is again a consequence of Lemma 1.1. □

Remark 1.1. From Proposition 1.3 we can deduce that v_N increases to a possibly infinite limit v_∞ as N goes to infinity. Assumption 1 implies that Λ is smooth on the interior of $\mathcal{D}(\Lambda)$ so that both quantities $v := \psi'(t^*)$ and $\chi := \frac{\pi^2}{2} t^* \psi''(t^*)$ are finite. In Section 1.1.4 we will see that v_∞ is in fact equal to v.

Lemma 1.3 — *The random variables $Y_{i,j}$ as defined in the proof of Proposition 1.2 satisfy the hypothesis of Kingman's Subadditive Theorem.*

Proof. For each $k \geq 1$ the sequence $\{Y_{k,2k}, Y_{2k,3k}, ...\} = \{\max X_k^k, \max X_k^{2k}, ...\}$ is i.i.d. so stationary and ergodic. Clearly the distribution of $(Y_{i,i+k})_{k\geq 0} = (\max X_k^i)_{k\geq 0}$ is independent of i. $\mathbb{E}Y_{0,1}^+ = \mathbb{E}(\max X_1)^+ < \infty$ because $\max X_1 \in L^1$ by (1.2). Finally, $\mathbb{E}Y_{0,n} = \mathbb{E}\max X_n \geq n\,\mathbb{E}\min\{\epsilon_{0,i,j} \mid i \in [N],\ j \in [\tau_{0,i}]\}$ where the expectation is finite by Lemma 1.2. □

1.1.3 Killed branching random walks

Adapting the notation used in [1], we formally define a Branching Random Walk (BRW) to be a pair $(\mathcal{T}, \Phi)$, where $\mathcal{T}$ is a Galton-Watson tree with offspring distribution q and Φ is a map assigning a random variable $\Phi(u)$ to each vertex $u \in \mathcal{T}$, independently of the structure of $\mathcal{T}$. Φ must be such that $\Phi(\text{root}) = 0$ and $\{\Phi(v) - \Phi(u) \mid u \text{ is the parent of } v\}$ is an i.i.d. collection with common distribution p. We call $\Phi(u)$ the value of the BRW at vertex u and write $\mathcal{T}(n)$ for the set of vertices in $\mathcal{T}$ at depth n. We say a (possibly finite) sequence of vertices $u_1, u_2, ...$ is a path if u_{i+1} is the parent of u_i for each $i \geq 1$.
Suppose that we have a BRW $(\mathcal{T}, \Phi)$ and take $v \in \mathbb{R}$ and $m \geq 1$. We say that vertex u is (m, v)-good if there exists a path $u = u_0, u_1, ..., u_m$ such that $\Phi(u_i) - \Phi(u) \geq vi$ for all $i \in [\![0, m]\!]$. This is essentially saying that there exists a path started from u that stays to the right of the space-time line through $(u, \Phi(u))$ with slope v, for at least m steps. The definition of an (∞, v)-good vertex is analogous. We now state two results from [2] that we will need to prove Theorem 1.6. Recall the definitions of v and χ from Remark 1.1.

Theorem 1.4 ([2, Theorem 1.2]) — *Let $\rho(\infty, \epsilon)$ denote the probability that the root of the BRW with offspring distribution q and step distribution p is $(\infty, v - \epsilon) -$ good. Then, as $\epsilon > 0$ goes to zero,*

$$\rho(\infty, \epsilon) \leq \exp\left(-\left(\frac{\chi + o(1)}{\epsilon}\right)^{1/2}\right). \tag{1.9}$$

A similar result can be stated for the probability of observing a $(m, v - \epsilon)$-good root with m finite:

Theorem 1.5 ([2, Consequence of proof of Theorem 1.2]) — *Let $\rho(m, \epsilon)$ denote the probability that the root of the BRW with offspring distribution q and step distribution p is $(m, v - \epsilon)$-good. For any $0 < \beta < \chi$, there exists $\theta > 0$ such that for all large m,*

$$\rho(m, \epsilon) \leq \exp\left(-\left(\frac{\chi - \beta}{\epsilon}\right)^{1/2}\right), \qquad \text{with } \epsilon := \theta m^{-2/3}.$$

1.1.4 Brunet-Derrida behaviour

We are now ready to present and prove our main result in this section, the analogue of Bérard and Gouéré's Theorem 1:

Theorem 1.6 — *As N goes to infinity,*

$$v_\infty - v_N = \frac{\chi}{(\log N)^2} + o((\log N)^{-2}).$$

First let us describe the coupling between the N-branching random walk and N independent branching random walks which allows us to relate Theorems 1.4 and 1.5 to the N-branching random walk. Let $(\text{BRW}_i)_{i \in [N]} = ((\mathcal{T}_i, \Phi_i))_{i \in [N]}$ be a set of N independent copies of the BRW with offspring distribution q and step distribution p. Define $\mathbb{T}_n := \bigsqcup_{i=1}^N \mathcal{T}_i(n)$ to be the disjoint union of vertices at depth n in the N BRWs, and fix an arbitrary (nonrandom) total order on $\mathbb{T}_n$ for each n. We now inductively define a sequence $(G_n)_{n \geq 0}$ of random subsets of $\mathbb{T}_n$, each with exactly N elements. These random subsets will correspond to the particles alive in the coupled N-braching random walk at time n. Define $G_0 = \mathbb{T}_0$ and given G_n, define H_n to be the vertices in $\mathbb{T}_{n+1}$ that descend from vertices in G_n. Finally, set G_{n+1} to be the set of N vertices in H_n with the gratest value, resolving ties via the fixed total order on $\mathbb{T}_{n+1}$. If we now define (with some abuse of notation) $\mathfrak{X}_n = \sum_{u, i: u \in G_n \cap \mathcal{T}_i} \delta_{\Phi_i(u)}$ then $(\mathfrak{X}_n)_{n \geq 0}$ has the same distribution as X started from $N\delta_0$. Going forward we will alternate between the notation of the two constructions of the N-branching random walk that we have given. Concretely, we will refer to $\mathcal{T}$, Φ, $\epsilon_{n,i,j}$ and $\tau_{n,i}$ without explicitly explaining the obvious relationships between these objects. Let us now record a technical lemma that will be used in the proof of the lower bound in Theorem 1.6.

LEMMA 1.7 ([4, Adapted by Bérard and Gouéré from Lemma 5.2]) — *Let* $v_1 < v_2 \in \mathbb{R}$ *and* $1 \leq m \leq n \in \mathbb{N}$. *Suppose* $0 =: x_0, ..., x_n$ *is a sequence of real numbers such that* $\max_{i\in[\![0,n-1]\!]}(x_{i+1} - x_i) \leq K$ *for some* $K > 0$, *and define* $I := \{i \in [\![0, n-m]\!] \mid x_{i+j} - x_i \geq jv_1, \quad \forall j \in [\![0, m]\!]\}$. *If* $x_n \geq v_2 n$, *then* $|I| \geq \frac{v_2-v_1}{K-v_1}\frac{n}{m} - \frac{K}{K-v_1}$.

Proof of lower bound in Theorem 1.6. As before, we first treat the case $X_0 = N\delta_0$. Our aim is to show $v_N := \lim_{n\to\infty} \mathbb{E}\left[n^{-1} \max X_n\right] \leq v_\infty - \chi/(\log N)^2 + o((\log N)^{-2})$. However, we shall show this with v_∞ replaced by v, which combined with the upper bound also proves that $v_\infty = v$. Set $\beta \in (0, \chi)$ and let $\theta > 0$ be as in Theorem 1.5. Let $\lambda > 0$, and define

$$m := \left\lceil \theta^{3/2}\left(\frac{(1+\lambda)\log N}{(\chi-\beta)^{1/2}}\right)^3\right\rceil, \tag{1.10}$$

and $\epsilon := \theta\, m^{-2/3}$. The scale of ϵ and m is carefully chosen so that by Theorem 1.5,

$$\rho(m, \epsilon) \leq N^{-(1+\lambda)} \qquad \text{for all large } N. \tag{1.11}$$

Take $\gamma \in (0, 1)$ and define $v_1 = v - \epsilon$ and $v_2 = v - (1-\gamma)\epsilon$ noting that $v_1 < v_2 < v$. Finally, let $n = \lceil N^\xi \rceil$ for some $0 < \xi < \lambda$ and consider the following inequality with $\delta > 0$:

$$\begin{aligned}\mathbb{E}\left[n^{-1}\max X_n\right] &= \mathbb{E}\left[n^{-1}\max X_n\left[\mathbb{1}_{\{\max X_n < nv_2\}} + \mathbb{1}_{\{nv_2 \leq \max X_n < n(v+\delta)n\}} + \mathbb{1}_{\{(v+\delta)n \leq \max X_n\}}\right]\right] \\ &\leq v_2 + (v+\delta)\underbrace{\mathbb{P}\left(\max X_n \leq v_2 n\right)}_{(I)} + \underbrace{\mathbb{E}\left[n^{-1}\max X_n \mathbb{1}_{\{(v+\delta)n\leq \max X_n\}}\right]}_{(II)}. \end{aligned} \tag{1.12}$$

The strategy for the proof is to show that both (I) and (II) are $o((logN)^{-2})$. The result then follows, as $v_2 = v - (1-\gamma)(\chi-\beta)(1+\lambda)^{-2}(\log N)^{-2}$ where γ, β, λ can be taken arbitrarily small.

Let B_n be the number of vertices in $\sqcup_{i=0}^n G_i$ that are (m, v_1)-good with respect to their respective BRWs. Define $K = \kappa \log(N)$ for some $\kappa > 0$ and notice that the quantity $\frac{v_2-v_1}{K-v_1}\frac{n}{m} - \frac{K}{K-v_1}$ is positive for large enough N. Let $u_0, u_1, ..., u_n$ be a path in $\mathcal{T}_{i_0}$ for some $i_0 \in [N]$ such that $u_0 = root_{i_0}$ and $u_n \in G_n$ with $\Phi_{i_0}(u_n) = \max X_n$. In other words, let BRW_{i_0} be the random walk that the rightmost particle of the coupled N-branching random walk lives in at time n, and let $u_0, ..., u_n$ be the path connecting it to the root. On the event $E := \{\max X_n \geq v_2 n\}$, we apply Lemma 1.7 to the sequence of real numbers $(\Phi_{i_0}(u_i))_{i\in[n]}$ to see that either there is an (m, v_1)-good vertex among the u_i or one of the random walk steps along the path is $\geq K$. These events are respectively included in the events $\{B_n \geq 1\}$ and $\{M := \max\{\epsilon_{l,i,j} \mid l \in [\![0, n-1]\!], i \in [N], j \in [\tau_{l,i}]\} \geq K\}$. We can use this to bound the probability of E:

$$(I) = \mathbb{P}(E) \leq \mathbb{P}(M \geq K) + \mathbb{P}(B_n \geq 1). \tag{1.13}$$

Consider a vertex $u \in \mathcal{T}_{i_0}(d)$ for some $i_0 \in [N]$ at depth $d \in [\![0, n]\!]$. The event $\{u \in G_d\}$ is measurable with respect to the sigma algebra generated by the random variables $\{\Phi_j(v) \mid j \in [N],\, \mathcal{T}_j \ni v\text{'s depth } \leq d\}$. On the other hand, the event $\{u \text{ is } (m, v_1)\text{-good}\}$ is determined by the variables $\{\Phi_{i_0}(v) - \Phi_{i_0}(u) \mid \mathcal{T}_{i_0} \ni v\text{'s depth } > d\}$, so that the two events are independent. We can write B_n as

$$B_n = \sum_{i=1}^N \sum_{u\in\mathcal{T}_i} \mathbb{1}_{\{u \text{ is } (n,v_1)\text{-good}\}} \mathbb{1}_{\{u\in G_d \text{ for some } d\in[\![0,n]\!]\}}.$$

Taking expectations gives

$$\mathbb{E}[B_n] \leq N(n+1)\rho(m, \epsilon) = \mathcal{O}(N^{\xi-\lambda}) = o((\log N)^{-2}) \qquad \text{as } N \text{ goes to infinity}, \tag{1.14}$$

where we used that G_n has N elements for all n. We now want to bound $\mathbb{P}(M \geq K)$, the probability of $S := \sum_{l=0}^{n-1}\sum_{i=1}^N \tau_{l,i} \in L^1$ i.i.d. variables with distribution p to be larger than K. Since p has exponentially decaying tails, we can take $C, \gamma, t_0 > 0$ so that $p([t, \infty)) \leq C\exp(-\gamma t)$ for all $t > t_0$. Then for $\kappa > t_0$ large enough for Bernoulli's inequality to apply, we have

$$\mathbb{P}(M \geq K) = 1 - \mathbb{E}\left[(1 - p([K, \infty)))^S\right] \leq 1 - \mathbb{E}\left[(1 - C\exp(-\gamma K))^S\right] \tag{1.15}$$

$$= 1 - \mathbb{E}\left[(1 - CN^{-\gamma\kappa})^S\right] \leq CN^{-\gamma\kappa}\mathbb{E}[S] = C\underbrace{\mathbb{E}[\tau]}_{<\infty} N^{1+\xi-\gamma\kappa}. \tag{1.16}$$

Thus, for large enough κ, $\mathbb{P}(M \geq K) = o((\log N)^{-2})$. This, combined with (1.14) and Markov's inequality gives $(I) = o((\log N)^{-2})$ as desired. We now turn to showing $(II) = o((\log N)^{-2})$. Consider the obvious inequality $\exp(t \max X_n) \leq \sum_{i\in[N],\, u\in\mathcal{T}_i(n)} \exp(t\Phi_i(u))$. If we set $\mathcal{G} := \sigma\{\tau_{l,i} \mid l \in [\![0, n-1]\!],\, i \in [N]\}$, then we have

$$\mathbb{E}\left[\exp(t \max X_n)\right] \leq \mathbb{E} \sum_{i=1}^{N} \sum_{u\in\mathcal{T}_i(n)} \mathbb{E}\left[\exp(t\Phi_i(u)) \middle| \mathcal{G}\right] = \mathbb{E} \sum_{i=1}^{N} \sum_{u\in\mathcal{T}_i(n)} \mathbb{E}_{\epsilon\sim p}\left[\exp(t\epsilon)\right]^n \tag{1.17}$$

$$= N \mathbb{E}_{\tau\sim q}\left[\tau\right]^n \mathbb{E}_{\epsilon\sim p}\left[\exp(t\epsilon)\right]^n, \tag{1.18}$$

where we used a telescoping sum along the path connecting u and the corresponding root for each vertex u in the sum. We can rewrite this in terms of $\psi(t)$:

$$N\mathbb{E}\left[\tau_{0,1}\right]^n \mathbb{E}_{\epsilon\sim p}\left[\exp(t\epsilon)\right]^n = N\mathbb{E}_{\epsilon\sim p\,\perp\,\tau\sim q}\left[\tau \exp(t\epsilon)\right]^n \tag{1.19}$$

$$= N\mathbb{E}_{\epsilon_j \overset{iid}{\sim} p\,\perp\,\tau\sim q}\Big[\sum_{j=1}^{\tau} \exp(t\epsilon_j)\Big]^n = N\exp(n\psi(t)). \tag{1.20}$$

Recalling from Assumption 2 and Remark 1.1 that $\psi(t^*) = vt^*$, we obtain

$$\mathbb{E}\left[\exp(t^*(\max X_n - vn))\right] \leq N. \tag{1.21}$$

Lemma 1.8 — *Let $b > 0$. Then for all large enough a,*

$$x\mathbb{1}_{\{x\geq a\}} \leq \exp\left(b\left(x - \frac{a}{2}\right)\right), \qquad \forall x \in \mathbb{R}. \tag{1.22}$$

Proof. Differentiate the map $f : x \mapsto \exp(b(x - a/2)) - x$ to find that for large enough a, f is increasing on $[a, \infty)$. Noting that $f(a) \geq 0$ for all large a concludes the proof. ☐

Apply Lemma 1.8 with $X = \max X_n - vn$, $a = \delta n, b = t^*$ and take expectations to get

$$\mathbb{E}\left[(\max X_n - vn)\mathbb{1}_{\{\max X_n \geq (v+\delta)n\}}\right] \leq \mathbb{E}\left[\exp(t^*(X_n - vn - \delta n/2))\right],$$

which combined with (1.21) and a Chernoff bound gives

$$(II) = \mathbb{E}\left[\max X_n \mathbb{1}_{\{\max X_n \geq (v+\delta)n\}}\right] \leq N\exp(-t^*\delta n/2)(1 + |v|n) = o((logN)^{-2}).$$

We have shown that for any choice of $\gamma \in (0,1)$, $\beta \in (0,\chi)$ and $\lambda > \xi > 0$, for all N large enough

$$\mathbb{E}\left[\lceil N^\xi \rceil^{-1} \max X_{\lceil N^\xi \rceil}\right] \leq v - \frac{(1-\gamma)(\chi-\beta)}{(1+\lambda)^2(\log N)^2} + o((\log N)^{-2}). \tag{1.23}$$

Recall from the proof of Proposition 1.2 that $v_N = \inf_n n^{-1}\mathbb{E}\left[\max X_n\right]$, so the left hand side in (1.23) can be replaced by v_N. Taking γ, β, λ and ξ to zero gives the desired result. ☐

Lemma 1.9 ([1, Lemma 3]) — *Let $(M_n)_{n\geq 0}$ be a supercritical Galton-Watson process with square integrable offspring distribution started from $M_0 = 1$. Then there exist constants $r > 0$ and $\phi > 1$ such that $\mathbb{P}(M_n \geq \phi^n) \geq r$ for all $n \geq 0$.*

Proof of upper bound in Theorem 1.6. Let $\tau \sim q$ and take $R < v$ such that $p([R,\infty)) > \mathbb{E}\left[\tau\right]^{-1}$. Define $M := (M_n)_{n\geq 0}$ to be a Galton-Watson process started from $M_0 = 1$ with offspring distribution $\tilde{q}$ such that $\tilde{q}|\tau \sim \text{Binomial}(\tau, p([R,\infty)))$. Then M is supercritical and has square integrable offspring distribution. Hence by Lemma 1.9 there exist $r > 0$ and $\phi > 1$ usch such that $\mathbb{P}(M_n \geq \phi^n) \geq r$ for all $n \geq 0$.

Define $\lambda \in (0,1)$ and let $\epsilon := \chi((1-\lambda)\log N)^{-2}$. Theorem 1.4 gives $\rho(\infty, \epsilon) = N^{\lambda-1+o(1)}$ as $N \to \infty$. Further define $s := \lceil \frac{\log N}{\log \phi} \rceil + 1$ and for $\eta \in (0,1)$ define $m := \lceil \frac{(c-R)s}{\eta\epsilon} \rceil$ and finally set $n = s + m$. Consider a vertex u at depth m in a $BRW = (\mathcal{T}, \Phi)$ with offspring distribution q and step distribution p. The probability that there are at least ϕ^s distinct paths $u =: u_m, ..., u_n$

with $\Phi(u_{i+1}) - \Phi(u_i) \geq R$ for all $i \in [\![m, n-1]\!]$ is greater than r by Lemma 1.9. Recall that the probability of the root being $(m, v-\epsilon)$-good is $\rho(m,\epsilon)$. In light of the previous discussion, we see that the probability of observing a path $root = w_0, ..., w_n$ in the BRW such that $\Phi(w_k) \geq k(v-\epsilon)$ for $k \in [\![0, m]\!]$ and $\Phi(w_{k+1}) - \Phi(v_k) \geq R$ for $k \in [\![m, n-1]\!]$ is at least $\rho(m,\epsilon)r$. By the choice of m and n, such a path must in fact be $(n, v-(1+\eta)\epsilon)$-good. For $i \in [N]$ define A_j to be the event that BRW_i contains no more than ϕ^s distinct $(n, v-(1+\eta)\epsilon)$-good paths starting at the root. By independence we get

$$\mathbb{P}\left(\bigcap_{i=1}^{N} A_i\right) \leq (1-\rho(m,\epsilon))^N. \tag{1.24}$$

Denote $B := \{\min X_k < (v-(1+\eta)\epsilon) \text{ for all } k \in [n]\}$. On the event $B \cap [\cap_{i=1}^N A_i]^c$ one of the BRW_is has $> \phi^s > N$ particles at time n that have stayed to the right of the space time line with slope $v-(1+\eta)\epsilon$ for all of $[n]$. By the definition of B this implies that there are $> N$ particles alive in the N-branching random walk which is a contradiction. Therefore we must have $B \subset \cap_{i=1}^N A_i$. Using the fact that $\rho(m,\epsilon) \leq \rho(\infty,\epsilon) = N^{\lambda-1+o(1)}$ and the inequality $1-x \leq \exp(-x)$ for all $x \in \mathbb{R}$, we get

$$\mathbb{P}(B) \leq \mathbb{P}\left(\bigcap_{i=1}^{N} A_i\right) \leq \exp(-N^{\lambda+o(1)}) \tag{1.25}$$

Lemma 1.10 ([1, Proposition 4]) — *With the previous notations, for all N large enough,*

$$v_N \geq (v-(1+\eta)\epsilon)(1-n\mathbb{P}(B)) - n\mathbb{E}\left[|\Theta_n|\mathbb{1}_B\right], \tag{1.26}$$

where Θ_n is distributed as the minimum of $\sum_{l=0}^{n-1}\sum_{i=1}^{N}\tau_{l,i}$ i.i.d. random variables distributed like p, independent from $(\tau_{l,i})_{l\in[\![0,n-1]\!],\, i\in[N]}$ which are i.i.d. with distribution q.

Proof. Identical to the one given in [1]. □

We immediately have $(v-(1+\eta)\epsilon)n\mathbb{P}(B) = o((\log N)^{-2})$ by (1.25). To bound the other term, note that $\mathbb{E}\left[|\Theta_n|\mathbb{1}_B\right] \leq (\mathbb{E}\left[\Theta_n^2\right]\mathbb{P}(B))^{1/2}$ by Hölder's inequality □

References

[1] Jean Bérard and Jean-Baptiste Gouéré. Brunet-derrida behavior of branching-selection particle systems on the line. *Communications in Mathematical Physics*, 298(2):323–342, 2010.

[2] Nina Gantert, Yueyun Hu, and Zhan Shi. Asymptotics for the survival probability in a supercritical branching random walk. *arXiv preprint arXiv:0811.0262*, 2008.

[3] Bastien Mallein. N-branching random walk with α-stable spine. *Theory of Probability & Its Applications*, 62(2):295–318, 2018.

[4] Robin Pemantle. Search cost for a nearly optimal path in a binary tree. *The Annals of Applied Probability*, pages 1273–1291, 2009.