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 4, 2019

UNIVERSITY OF
OXFORD

DEPARTMENT OF
STATISTICS

Contents

0.1 Exponentially decaying tails

0.1.1 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, opposed to being fixed at two. Denote by $X = (X_n)_{n\geq 0} = (\sum_{i=1}^{N} \delta_{X_n(i)})_{n\geq 0}$ the N-branching random walk with particles at positions $X_n(1) \leq ... \leq X_n(N)$. At each timestep the system undergoes the following two steps:

(i) Branching: Each particle dies and gives birth to a random number of offspring. The number of children is distributed with law q and is independent of the past of the process. Given the position of the parent, say x, each child's position follows the law $p(\,\cdot - x)$ independently.

(ii) Selection: Out of all children, the N rightmost are selected to form the population at the next timestep.

Construction. As before, let $X = (X_n)_{n\geq 0}$ denote the $\mathcal{C}_N$-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 $\mathcal{E}_N := (\epsilon_{l,i,j})_{l\geq 0,\, i\in [\![1,N]\!],\, j\geq 1}$ and $\mathcal{M}_N := (\tau_{l,i})_{l\geq 0, i\in [\![1,N]\!]}$ 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}. This construction gives rise to an important monotonicity property that we record in the following Lemma:

Lemma 0.1 ([1, Corollary 2]) — *For any $1 \leq N_1 \leq N_2$ and $\tau_i \in \mathcal{C}_{N_i}$ with $i = 1, 2$ such that $\tau_1 \preceq \tau_2$, there exists a coupling $(X_n^{(1)}, X_n^{(2)})_{n\geq 0}$ between two versions of the branching-selection particle system started from τ_1 and τ_2 respectively satisfying $X_n^{(1)} \preceq X_n^{(2)}$ almost surely for all $n \geq 0$.*

Proof. The proof is a straightworward extension of our construction. The idea is to take the i.i.d. families $\mathcal{E}_{N_2}$ and $\mathcal{M}_{N_2}$ defined as above and use them to define both processes. Note that the argument hinges on the fact that in our notation $X_n(1) \leq ... \leq X_n(N)$. □

0.1.2 Assumptions

Let $\nu \in \mathcal{C}$ be a random, finite counting measure with the same distribution as the offspring of a single particle at origin in our branching-selection mechanism (the fact that $\nu \in \mathcal{C}$ 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:

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

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 now state the assumptions necessary 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 has exponentially decaying tails, $q(0) = 0$ and $\sum_{i=1}^{\infty} iq(i) > 1$.

We will use the notation $\mu := \int_{[1,\infty)} x dq(x) > 1$. 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 0.2 — *Let $\tau \in \mathbb{N}$ be a random variable with exponentially decaying tails and let $(\epsilon_i)_{i\geq 1}$ be an i.i.d. sequence of random variables with exponentially decaying tails, independent of τ. Then $M := \max_{1\leq i\leq \tau} \epsilon_i$ 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 $t > t_0$ Bernoulli's inequality gives

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

Similarly, looking at the lower tail we get

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

□

0.1.3 PROPERTIES OF THE MODEL

Denote by $\max X_n$ and $\min X_n$ the right- and leftmost atom 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 $X_0 \in \mathcal{C}_N$. Indeed, by independence we have

$$\mathbb{E}|\max X_n| \leq \mathbb{E}\left|\max X_0 + \sum_{l=0}^{n-1}\sum_{i=1}^{N}\sum_{j=1}^{\tau_{l,i}} \epsilon_{l,i,j}\right| \leq |\max X_0| + Nn\mathbb{E}\left[\tau_{0,1}\right]\mathbb{E}|\epsilon_{0,1,1}|. \tag{0.1}$$

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 0.1 — *For any $N \geq 1$ and initial population $X_0 \in \mathcal{C}_N$, we have*

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

Proof. Let $u_\kappa := \kappa\lceil \log N\rceil$ for some $\kappa \in \mathbb{N}_+$ and take $n \geq u_N$. Define $\mathcal{E} := \{\epsilon_{l,i,j} \mid l \in [\![n - u_N, n-1]\!],\, i \in [\![1, N]\!],\, j \in [\![1, \tau_{l,i}]\!]\}$ and let $M := \max \mathcal{E}$, $m := \min \mathcal{E}$ noting that both have exponentially decaying tails by Lemma 0.2. Now consider the process X in the time-frame $[\![n - u_N, n]\!]$ where the evolution of X is governed by $\mathcal{E}$. Write $y := \max X_{n-u_N}$ for the rightmost particle's position at time $n - u_N$. Suppose that for each $k \in [\![1, u_N]\!]$ we have $\min X_{n-u_N+k} < y + km$. As all steps during branching are $\geq m$, this implies in particular that the descendents of the particle 'y' survive all selection steps until time n. Therefore, on the event $A_\kappa := \{\text{number of descendants of } y \text{ at time } n \text{ is } > N\}$ almost surely $\min X_{n-u_N+k} \geq y + k_0 m$ for some k_0. By the definition of m this must also hold for all $k \in [\![k_0, u_N]\!]$, in particular for $k = u_N$. Noting that $\max X_n \leq y + u_N M$, it follows that

$$d(X_n)\mathbb{1}_{A_\kappa} \leq u_N(M - m), \tag{0.2}$$

with probability one. A simple argument shows that $\mathbb{1}_{A_\kappa} \to 1$ almost surely as $\kappa \uparrow \infty$:

$$\mathbb{P}(A_\kappa^c) \leq \sum_{k=0}^{N-1}\binom{u_\kappa}{k} q(1)^{u_\kappa - k}(1 - q(1))^k \leq u_\kappa^{N-1} q(1)^{u_\kappa} \to 0 \text{ as } \kappa \uparrow \infty \text{ since } q(1) < 1. \tag{0.3}$$

Fix $\epsilon > 0$ and take κ large enough so that $\mathbb{P}(A_\kappa^c) < \epsilon^2$. Consider the decomposition

$$\frac{d(X_n)}{n} = \frac{d(X_n)}{n}\mathbb{1}_{A_\kappa} + \frac{d(X_n)}{n}\mathbb{1}_{A_\kappa^c}. \tag{0.4}$$

Taking expectations and then taking n to infinity, the first term vanishes by 0.2. The second term is upper by $(\mathbb{P}(A_\kappa^c)\,\mathbb{E}\left[d(X_n)^2/n^2\right])^{1/2}$ 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 0.2 this yields $\mathbb{E}\left[d(X_n)^2\right] = \mathcal{O}(n^2)$ which implies that the second term in 0.4 is $\mathcal{O}(\epsilon)$. Taking ϵ to zero concludes the proof of L^1 convergence. Almost sure convergence is a consequence of the next Proposition. □

Proposition 0.2 ([1, Proposition 2]) — *There exists $v_N = v_N(p) \in \mathbb{R}$ such that for any initial population $X_0 \in \mathcal{C}_N$ the following holds almost surely and in L^1:*

$$\lim_{n\to\infty} \frac{\min X_n}{n} = \lim_{n\to\infty} \frac{\max X_n}{n} = v_N. \tag{0.5}$$

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. For this proof we 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 0.1 it follows easily that

$$\max X_{n+m}^0 \leq \max X_n^0 + \max X_m^n \qquad \forall\, n, m \geq 0. \tag{0.6}$$

For notational simplicity define $Y_{i,j} = \max X_{j-i}^i$ for $0 \leq i \leq j$. Then 0.6 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 0.3. 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 0.1 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 $X_0 \in \mathcal{C}_N$ note that the result is a consequence of Lemma 0.1 and a sandwiching argument between the initial configurations $N\delta_{\min X_0}$ and $N\delta_{\max X_0}$. $\square$

Lemma 0.3 — *The random variables $Y_{i,j}$ as defined in the proof of Proposition 0.2 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 0.1. Finally, $\mathbb{E}Y_{0,n} = \mathbb{E}\max X_n \geq n\,\mathbb{E}\min\{\epsilon_{0,i,j} \mid i \in [\![1, N]\!],\ j \in [\![1, \tau_{0,i}]\!]\}$ where the expectation is finite by Lemma 0.2. $\square$

Proposition 0.3 ([1, analogue of Proposition 3]) — *The sequence $(v_N)_{N\geq 1}$ is non-decreasing.*

Proof. This is again a consequence of Lemma 0.1. $\square$

Remark 0.1. From Proposition 0.3 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 $v := \phi'(t^*)$ and $\chi := \frac{\pi^2}{2} t^* \phi''(t^*)$ are finite. In Section 0.1.5 we will see that v_∞ is in fact equal to v.

0.1.4 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 $\mathcal{T}$. Φ must be such that $\Phi(\text{root}) = 0$ and $\{\Phi(v) - \Phi(u) \mid u \text{ is the parent of } v\}$ is i.i.d. 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 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 0.6. Recall the definitions of v and χ from Remark 0.1.

Theorem 0.4 ([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,*

$$\rho(\infty,\epsilon) \le \exp\left(-\left(\frac{\chi+o(1)}{\epsilon}\right)^{1/2}\right). \tag{0.7}$$

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

Theorem 0.5 ([2, Consequence of proof of Theorem 1.2]) — *Let $\rho(m,\epsilon)$ denote the probability that the root of the BRW with offspring distribution p is $(m, v-\epsilon)$-good. For any $0<\beta<\chi$, there exists $\theta > 0$ such that for all large m,*

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

0.1.5 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 0.6 — *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 0.4 and 0.8 to the N-branching random walk. Let $(\mathrm{BRW}_i)_{i\in[\![1,N]\!]} = ((\mathcal{T}_i,\Phi_i))_{i\in[\![1,N]\!]}$ 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 $(G_n)_{n\ge 0}$ 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 $(\mathfrak{X}_n)_{n\ge0}$ 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 mentioning explicitly 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 0.6.

Lemma 0.7 ([3, Adapted by Bérard and Gouéré from Lemma 5.2]) — *Let $v_1 < v_2 \in \mathbb{R}$ and $1 \le m \le n \in \mathbb{N}$. Suppose $0 =: x_0, ..., x_n$ is a sequence of real numbers such that $\max_{i\in[\![0,n-1]\!]}(x_{i+1} - x_i) \le K$ for some $K>0$, and define $I := \{i \in [\![0, n-m]\!] \mid x_{i+j} - x_i \ge jv_1, \quad \forall j \in [\![0,m]\!]\}$. If $x_n \ge v_2 n$, then $|I| \ge \frac{v_2-v_1}{K-v_1}\frac{m}{n} - \frac{K}{K-v_1}$.*

Proof of lower bound in Theorem 0.6. As before, set $X_0 = N\delta_0$. Our aim is to show $v_N := \lim_{n\to\infty} \mathbb{E}\left[n^{-1}\max X_n\right] \le v_\infty - \chi/(\log N)^2 + o((\log N)^{-2})$. However, we shall show this with v_∞ replaced by v, which combined with the upper bound also proves that $v_\infty = v$. Let $\beta \in (0,\chi)$ and let $\theta > 0$ be as in Theorem 0.8. Let $\lambda > 0$, and define

$$m := \left\lceil \theta^{3/2}\left(\frac{(1+\lambda)\log N}{(\chi-\beta)^{1/2}}\right)^3\right\rceil, \tag{0.8}$$

and $\epsilon := \theta\, m^{-2/3}$. The scale of ϵ and m is carefully chosen so that by Theorem 0.8,

$$\rho(m,\epsilon) \le N^{-(1+\lambda)} \qquad \text{for all large } N. \tag{0.9}$$

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 $\delta > 0$:

$$\begin{aligned}\mathbb{E}\left[n^{-1}\max X_n\right] &= \mathbb{E}\left[n^{-1}\max X_n\left[\mathbb{1}_{\{\max X_n < nv_2\}} + \mathbb{1}_{\{nv_2 \le \max X_n < n(v+\delta)n\}} + \mathbb{1}_{\{(v+\delta)n \le \max X_n\}}\right]\right] \\ &\le v_2 + (v+\delta)\underbrace{\mathbb{P}\left(\max X_n \le v_2 n\right)}_{(I)} + \underbrace{\mathbb{E}\left[n^{-1}\max X_n \mathbb{1}_{\{(v+\delta)n \le \max X_n\}}\right]}_{(II)}.\end{aligned} \tag{0.10}$$

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 $\sqcup_{i=1}^{n} G_i$ that are (m, v_1)-good with respect to their respective BRWs. Define $K = \kappa \log(2Nn)$ for some $\kappa > 0$ and notice that the quantity $\frac{v_2 - v_1}{K - v_1}\frac{m}{n} - \frac{K}{K - v_1} = \Theta(N^{\xi}(\log N)^{-4})$ so that for large enough N it is positive. Let $u_0, u_1, ..., u_n$ be a path in $\mathcal{T}_{i_0}$ for some $i_0 \in [\![1,N]\!]$ 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 $u_0, ..., u_n$ be the path from the root to the rightmost particle at time n of the coupled N-branching random walk. On the event $E := \{\max X_n \ge v_2 n\}$, we apply Lemma 0.7 to the sequence of real numbers $(\Phi_{i_0}(u_i))_{i \in [\![1,n]\!]}$ 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 $\ge K$. These events are respectively included in the events that $B_n \ge 1$ and that $M := \max\{\epsilon_{l,i,j} \mid l \in [\![0, n-1]\!],\ i \in [\![1,N]\!],\ j = 1,2\} \ge K$. We can use this to bound the probability of E:

$$\mathbb{P}(E) \le \mathbb{P}(M \ge K) + \mathbb{P}(B_n \ge 1). \tag{0.11}$$

Consider a vertex $u \in \mathcal{T}_{i_0}(d)$ for some $i_0 \in [\![1,N]\!]$ 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 $\{\Phi_j(v) \mid j \in [\![1,N]\!],\ \mathcal{T}_j \ni v\text{'s depth } \le d\}$. 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

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

Taking expectations gives

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

where we used that G_n has N elements for all n. Recall that the distribution p has exponentially decaying tails, so that there exist $C, \gamma > 0$ such that $\mathbb{P}_{X \sim p}(X > t) \le C\exp(-\gamma t)$ for all large t. This gives $\mathbb{P}(M \ge K) \le 1 - (1 - \exp(-\gamma\kappa\log(2Nn)))^{2Nn} \le (2Nn)^{1-\gamma\kappa} = o((\log N)^{-2})$, for $\kappa > \gamma^{-1}$. Together with **??** this shows that $(I) = o((\log)^{-2})$.
To show that $(II) = o((\log N)^{-2})$ first consider the obvious inequality $\exp(t\max X_n) \le \sum_{i \in [\![1,N]\!],\ u \in \mathcal{T}_i(n)} \exp(t\Phi_i(u))$. Taking expectations gives $\mathbb{E}[\exp(t\max X_n)] \le N2^n\exp(n\Lambda(t))$, where we used a telescoping sum along the path connecting the root and u and the fact that $\#\mathcal{T}_i(n) = 2^n$ for each i. Recalling from Assumption 2 and Remark 0.1 that $\Lambda(t^*) = vt^* - \log 2$, we obtain

$$\mathbb{E}[\exp(t^*\max X_n)] \le N\exp(vnt^*). \tag{0.12}$$

Lemma 0.8 — *Let $b > 0$. Then for all large enough a,*

$$x\mathbb{1}_{\{x \ge a\}} \le \exp\left(b\left(x - \frac{a}{2}\right)\right), \qquad \forall x \in \mathbb{R}. \tag{0.13}$$

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 $f(a) \ge 0$ for all large a concludes the proof. ☐

Apply Lemma 0.8 with $X = \max X_n - vn$, $a = \delta n, b = t^*$ and take expectations to get

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

which combined with 0.14 and a Chernoff bound gives

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

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

$$\mathbb{E}\left[\lceil N^{\xi}\rceil^{-1} \max X_{\lceil N^{\xi}\rceil}\right] \leq v - (1-\gamma)\frac{\chi-\beta}{(1+\lambda)^2(\log N)^2} + o((\log N)^{-2}). \tag{0.14}$$

Recall from the proof of Proposition 0.2 that $v_N = \inf_n n^{-1}\mathbb{E}[\max X_n]$, so the left hand side in 0.16 can be replaced by v_N. Taking γ, β, λ and ξ to zero gives the desired result. □

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