University of Bath

PHD

Branching processes with selection

Talyigas, Zsofia

Award date:
2022

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 02. Dec. 2022

Branching processes with selection

submitted by

Zsófia Talyigás

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

August 2022

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with the author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

Signature of Author ....

Zsófia Talyigás

Summary

The main focus of this thesis is a branching particle system with selection, called the N-particle branching random walk (N-BRW), which was first proposed by Brunet and Derrida. The N-BRW is a discrete time stochastic process, which can be viewed as a toy model of an evolving population affected by natural selection. In the N-BRW we have N particles located on the real line at all times. At each time step, each of the N particles has two offspring, which have a random displacement from the location of their parent according to some fixed jump distribution. Then among the $2N$ offspring particles, only the N rightmost particles survive to form the next generation.

The most interesting questions about the N-BRW concern the following properties. First, the speed at which the particle cloud is moving to the right on the real line; second, the shape of the particle cloud; and finally the genealogy or family tree structure of the population.

The study of the N-BRW and related branching particle systems with selection has been of great interest in recent years. Existing results and conjectures show that the long-term behaviour of the N-BRW heavily depends on the jump distribution.

For the N-BRW with 'light-tailed' (roughly means exponentially decaying tails) jump distribution, Brunet and Derrida made conjectures about the behaviours of the speed and shape of the particle cloud, and about the genealogies of the population of particles. These conjectures inspired several mathematical results in this area; for example, Bérard and Gouéré proved the conjecture concerning the speed of the particle cloud.

For the N-BRW with 'heavy-tailed' (meaning polynomially decaying tails) jump distribution, Bérard and Maillard described the behaviour of the speed and made predictions about the genealogies and spatial distribution of the population. These results and conjectures all showed substantially different behaviour from those in the case when the jump distribution is 'light-tailed'.

The first main result of this thesis proves the conjectures of Bérard and Maillard about the 'heavy-tailed' case of the N-BRW. We prove that at a typical large time the genealogy of the population is given by a star-shaped coalescent, and that almost the whole population is near the leftmost particle on the relevant space scale.

Furthermore, motivated by the fact that in the 'light-tailed' and 'heavy-tailed' cases the N-BRW shows very different behaviour, we studied an intermediate case, where the jump distribution has stretched exponential tails. The second main result of this thesis describes the behaviour of the speed of the particle cloud in the stretched exponential case, filling a gap between the 'light-tailed' and 'heavy-tailed' regimes.

Our third result is on the genealogy of the N-BRW when the jump distribution has stretched exponential tails. We give a summary on the proof of this result rather than a full proof. We also mention some of the remaining open questions about the genealogies in this case, which we intend to study in the future.

Acknowledgements

First and foremost I would like to thank my supervisors Sarah Penington and Matt Roberts for the years of collaboration throughout my PhD. I am grateful for their dedication, several ideas, guidance on research and writing, and for many chats about life in academia and anything else.

The fact that I started on a PhD at all, and that I found a studentship at Bath, is in large part thanks to my previous supervisor in Budapest, Bálint Vető.

I also wish to thank the staff and fellow PhD students of SAMBa and the Probability Laboratory at Bath for the friendly and stimulating environment with various opportunities to attend seminars, lectures, reading courses, Integrative Think Tanks, and social events.

Many thanks go to all my friends in Bath, who have been up for boardgames, tennis matches, hiking tours, lunches, and summer schools. Special thanks to everyone who has been my housemate and/or office mate over the last four years.

I am grateful to my family and friends in Budapest and elsewhere for staying in touch and for being absolutely amazing hosts whenever I was visiting; especially to my parents who have been supporting my countless years of education and everything else I have ever set my mind to.

Finally, I would like to thank Andris for all the fun we have had in the last four (well, seven) years, for his constant patience and encouragement, and for following wherever academic life takes me.

Contents

Chapter 1

Introduction

Branching processes are classical models in probability theory. The study of this area dates back to the 1840's when Bienaymé investigated the extinction of noble family lines [7]. Since then, an extensive literature has developed in this field, and applications range from population modelling and population genetics, through epidemiology to nuclear chain reactions (see several examples in [30]).

In classical branching processes, particles move around in space and branch into two or more new particles, which continue moving independently of each other. In recent years, it has been of great interest to include interaction between the particles to model phenomena seen in real-world systems. In this thesis we focus on branching processes where the type of interaction is modelling natural selection.

Questions investigated in the literature include the velocity and spatial distribution of the cloud of particles, as well as the ancestral properties of a sample of particles. This thesis contains new results concerning each of these problems.

In this introduction we describe the branching-selection particle system that we will investigate, the N-particle branching random walk; we also outline the main results and cover background on branching processes with selection.

1.1 Branching processes

The basic setting of branching processes is called the Bienaymé (or more commonly Galton-Watson) process. This is a discrete time stochastic process on $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, which describes the number of individuals in a population throughout a large number of generations. The number of individuals in the $(n+1)$th generation is given by the sum of the number of children of every individual in the nth generation; each individual in the nth generation has an independent and identically distributed (i.i.d.) number of children.

The important book of Athreya and Ney [1] explains several properties of Bienaymé processes, starting from moments and survival probabilities to limit theorems in critical and supercritical cases. The book also deals with other variants of branching processes, such as continuous time and multitype branching processes.

Furthermore, the book introduces branching processes with spatial components, for example branching random walks, cascades and branching diffusions. In this thesis we will focus on the spatial aspect, and the number of offspring of each individual will be fixed to be 2. We now introduce the process that we will call branching random walk (BRW) throughout thesis.

The BRW model

To introduce the process we let X be a non-negative real-valued random variable. We will refer to the distribution of X as the *jump distribution*. Then we define the BRW as follows.

- Initially (at time 0) we have one particle at a position in $\mathbb{R}$.

Then at each integer time step:

- Each particle is replaced by two offspring.
- Each offspring particle performs a jump from its parents' location, independently from the other jumps, and with the same law as X.

Note that for each n, the nth generation consists of 2^n particles. One could give a more general definition, where the number of offspring is random and follows a Bienaymé process, and where the random variable X is not necessarily non-negative. However, in our case the above definition will be more relevant.

In the area of branching random walks, questions about the limiting behaviour of the rightmost position and of the rightmost trajectory have been studied in the literature. Zhan Shi's lecture notes [43] contain several results in this area and discuss the role of certain important martingales related to branching random walks.

In Section 1.4.1 we will discuss some of the existing results by Kingman [31], Hammersley [29], Biggins [8], Durrett [21], and Gantert [27] on the limiting behaviour of the rightmost particle in cases of jump distributions with different tails. These results will be interesting to compare with the behaviours we observe in the next model we introduce now, the N-particle branching random walk (N-BRW).

The N-BRW is the main focus of this thesis. In this model we introduce the selection rule, which keeps the population size constant, which is unlike the exponentially growing population of the BRW model.

1.1.1 The N-BRW model

The N-BRW is a branching process with selection, with N particles located on the real line at all times. Let us introduce the notation $[N] := \{1, 2, \ldots, N\}$. Let $\mathcal{X}_1(n) \leq \cdots \leq \mathcal{X}_N(n)$ denote the ordered positions of the N particles at time $n \in \mathbb{N}_0$. The positions $\mathcal{X}_1(n)$ will be referred to as leftmost and $\mathcal{X}_N(n)$ as rightmost particles at time n. Similarly to the BRW, we let X be a non-negative real-valued random variable, whose probability distribution is called the jump distribution.

In the process, the N particles start from some initial configuration $\mathcal{X}_1(0), \ldots, \mathcal{X}_N(0)$. Then, at each time step a branching and a selection step is performed, which are defined as follows:

Branching:

- Each particle is replaced by two offspring.
- Each of the $2N$ offspring particles performs a jump from its parents' location, independently from the other jumps, and with the same law as X.

Selection:

- From the $2N$ offspring particles only the N rightmost particles survive to form the next generation. (We say that the N leftmost particles are 'killed'.)

We give a formal definition of the N-BRW in Chapter 2.

The N-BRW model, together with other similar branching processes with selection, was proposed by Brunet and Derrida in the physics literature [14, 15]. Based on their further work with Mueller and Munier [11, 12] they conjecture that there is a class of models which shows similar scaling properties in velocity, spatial distribution and genealogy, independently of the details of the models.

Here, by genealogy we refer to the family tree structure of a uniform sample of particles from the population. One of the reasons why Brunet and Derrida proposed this model is that it is a toy model for an evolving population under natural selection. One can think of the locations of the particles as representing the fitness levels of individuals in a population, and the selection rule says that only the fittest half of the offspring individuals survive and have descendants in future generations. Considering the toy model aspect of the N-BRW it is particularly interesting to study the genealogies in this process.

The N-BRW with '*light-tailed*' jump distribution is conjectured to be in the class of models described by Brunet and Derrida, where by light-tailed we mean that the jump distribution has some exponential moments. Bérard and Gouéré investigated the limiting behaviour of the velocity of the particle cloud in the N-BRW, and proved one of the conjectures of Brunet and Derrida rigorously in the light-tailed case [2].

Later on, Bérard and Maillard described the behaviour of the speed and the limiting process of the N-BRW when the jump distribution has regularly varying tails (we call this case '*heavy-tailed*') [3]. Furthermore, they predicted the long-term behaviour of the spatial distribution and the genealogy of the population in this regularly varying tails regime.

We give a detailed summary of the works [2] and [3] in Sections 1.4.2 and 1.4.3. The results of these two articles show that the N-BRW exhibits substantially different behaviours under light-tailed and heavy-tailed conditions. Inspired by earlier works, in this thesis we investigate the N-BRW with different jump distributions. We prove the conjectures that appeared in [3], and we study an intermediate case between light-tailed and heavy-tailed, in which the jump distribution has stretched exponential tails.

1.2 Genealogy and spatial distribution for polynomial-tailed jump distributions

In Chapter 2 we investigate the N-BRW with a jump distribution with regularly varying tails. The precise statement of our result involves heavy notation which we are not going to introduce in this section. Instead, we state the result informally, and for now, we also assume a specific jump distribution instead of a general one with regularly varying tails. We will introduce regularly varying functions in Chapter 2.

Assume that we have an N-BRW with jump distribution given by

$$\mathbb{P}(X > x) = \min(1, x^{-\alpha}) \quad \text{for all } x \geq 0, \tag{1.2.1}$$

and for some $\alpha > 0$.

1.2.1 Scaling

Before stating our main result, we mention a simple consequence of the selection rule, which motivates the scaling in our result. As is illustrated in Figure 1-1, if a particle performs an extremely large jump and ends up far ahead of every other particle, then its descendants will have a large advantage compared to the other particles' descendants, and they will have a good chance of surviving in future generations. It is then possible that the number of descendants of the particle which made the extremely large jump doubles at each time step until $\lceil \log_2 N \rceil$ time after the jump, when these descendants take over the whole population.

This simple idea is the reason for the choice of our time scale, which will be denoted by

$$\ell_N := \lceil \log_2 N \rceil . \tag{1.2.2}$$

Then our relevant space scale will be

$$a_N := (2N\ell_N)^{1/\alpha}, \tag{1.2.3}$$

which is shown as the order of magnitude of the size of an extremely large jump in Figure 1-1. With this choice of a_N, for any positive constant c, the expected number of jumps which are larger than ca_N during a time interval of length ℓ_N is of constant order, as N goes to infinity. The reason for this is that at each time step there are $2N$ jumps in the N-BRW; thus, by (1.2.1), in ℓ_N time we expect

$$2N\ell_N\mathbb{P}(X > ca_N) = c^{-\alpha}$$

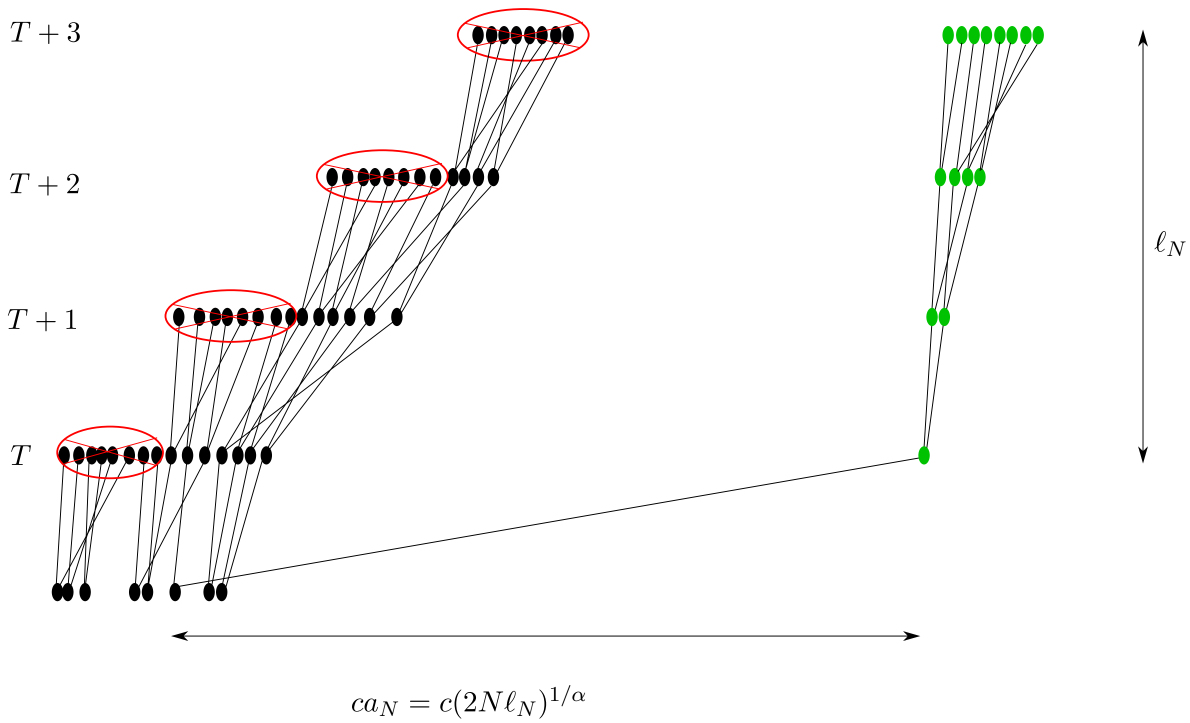

Figure 1-1: Big jump effect and scaling of the N-BRW: The black and green dots represent particles. Particles are ordered horizontally from left to right. Time increases from bottom to top. Each particle is connected with a line with two other particles from the next generation: these are its children, displayed at the horizontal locations where they jumped to. Particles circled and crossed with red are killed in the selection step. This example shows the particle configurations of the N-BRW between times $T-1$ and $T+3$ with $N = 8$. The bottommost green particle performed a big jump of order a_N. The descendants of the green particle are also displayed in green. These descendants take over the population in ℓ_N time steps.

jumps of size larger than ca_N. Jumps of order a_N will be referred to as *big jumps* in the N-BRW with polynomial tails. We note that Figure 1-1 illustrates the scaling and a possible effect of a big jump, but it does not describe the typical behaviour of the process. Below we give the informal statement of the theorem we prove in Chapter 2.

1.2.2 The main result (in words)

Consider the N-BRW with jump distribution given by (1.2.1). Recall that $\mathcal{X}_1(n)$ denotes the leftmost position at time n. When we say 'with high probability', we mean with probability converging to 1 as $N \to \infty$. Our theorem says the following.

For all $\eta > 0$, $M \in \mathbb{N}$ and $t \in \mathbb{N}_0$ with $t > 4\ell_N$, the N-BRW has the following properties with high probability:

- **Spatial distribution:** *At time t there are $N - o(N)$ particles within distance ηa_N of the leftmost particle, i.e. in the interval $[\mathcal{X}_1(t), \mathcal{X}_1(t) + \eta a_N]$.*

- **Genealogy:** *The genealogy of the population on an ℓ_N time scale is asymptotically given by a star-shaped coalescent, and the time to coalescence is between ℓ_N and $2\ell_N$.*

  That is, there exists a time $T \in [t-2\ell_N, t-\ell_N]$ such that with high probability, if we choose M particles uniformly at random at time t, then every one of these particles descends from the rightmost particle at time T. Furthermore, with high probability no two particles in the sample of size M have a common ancestor after time $T + \varepsilon_N \ell_N$, where ε_N is any sequence satisfying $\varepsilon_N \to 0$ and $\varepsilon_N \ell_N \to \infty$, as $N \to \infty$.

The genealogy result is illustrated in Figure 1-2. We call the coalescent star-shaped, because looking on an ℓ_N time scale all coalescences of the lineages of the particles appear to happen at the same time.

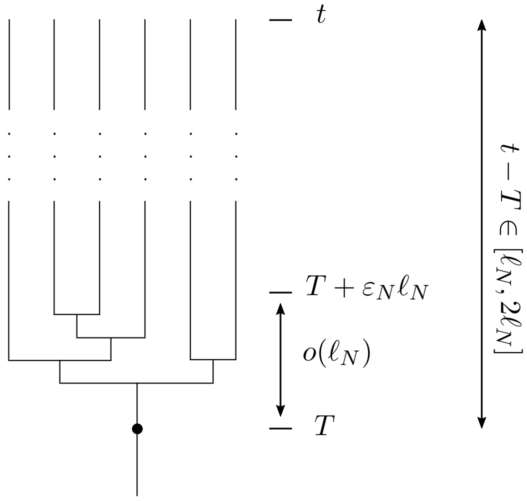

Figure 1-2: Coalescence of the ancestral lineages of $M = 6$ particles sampled at time t. To each particle in the sample we associate a vertical line, representing its ancestral line. Two lines coalesce into one when the particles they are associated with have a common ancestor for the first time going backwards from time t. The three dots in each line indicate that the picture is not proportional: the time between t and T is of order ℓ_N, whereas the time between all coalescences and T is $o(\ell_N)$.

1.2.3 Tribe heuristics

In our proof we built on the idea of 'tribes' which was described in the work of Bérard and Maillard [3]. The message of the tribe heuristics is, that at a typical large time we should think of the population as follows: On the a_N space scale, $N - o(N)$ particles are close to the leftmost particle; we say that these particles belong to the *big tribe*. The rest of the population is to the right of the big tribe in *small tribes* of size $o(N)$. The number of small tribes is of constant order, the distance between them is of order a_N, and the distance between particles within a tribe is $o(a_N)$. We think of particles in a tribe (big or small) as descendants of a single particle that made a big jump (see Section 1.2.1).

We will see that the positions of the tribes do not change too much in ℓ_N time, but the number of particles changes in each tribe. In particular the original big tribe will eventually die out and one of the small tribes will grow to become the new big tribe. In order to start a new tribe, a particle needs to make a big jump of order a_N. If a particle performs a big jump and becomes the rightmost particle or 'leader', there is a good chance that its descendants will take over the population (i.e. there will be $N - o(N)$ of these descendants) and become the new big tribe.

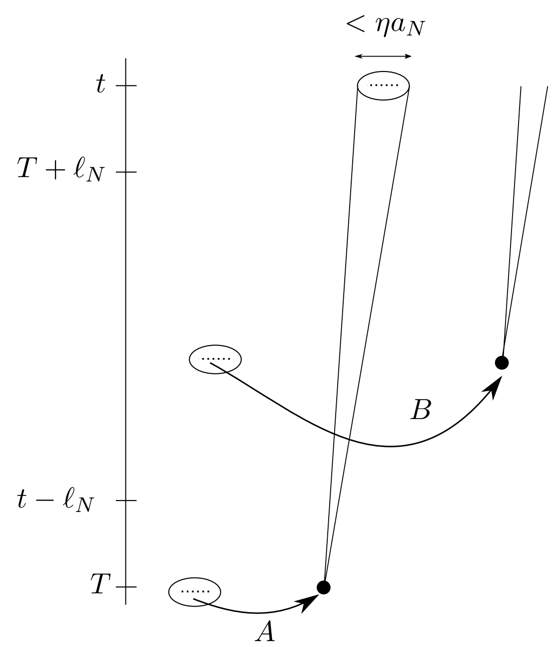

Figure 1-3: A particle that makes a big jump of order a_N at time T is the common ancestor of almost the whole population at time t. The vertical axis represents time, and the particles' locations are depicted horizontally, increasing from left to right. The black dots represent particles. Horizontal dotted lines in an ellipse show where the majority of the population (the big tribe) is. The arrows represent jumps from the big tribe. At the top of the figure we indicate that the every particle in the big tribe is within distance ηa_N for some small $\eta > 0$. The events labelled A and B are described in the main text.

In Figure 1-3 we illustrate (with A) that a particle makes a big jump and becomes the new leader at some time $T \in [t - 2\ell_N, t - \ell_N]$. With B we show a particle that also starts a new tribe, but since this big jump happens significantly after time $t - \ell_N$, the tribe created with that jump will still be a small tribe at time t. In contrast, the tribe started at A has enough time to grow to a big tribe of size $N - o(N)$ by time t. Once we have shown that $N - o(N)$ particles descend from a time-T particle with $T \in [t - 2\ell_N, t - \ell_N]$, we will be able to see that with high probability a sample of M particles are all from the big tribe, and so they indeed have a common ancestor at time t. We will give a more detailed heuristic picture in Chapter 2.

By a *path* we mean a sequence of jumps between a particle and one of its descendants. A key ingredient of the proof of our result will be a large deviation result on jumps restricted to being smaller than a_N multiplied by a small constant; the result says that paths without big jumps move very little on the a_N space scale. This will be the reason for tribes not moving too much during a time interval of length roughly ℓ_N, in particular for the big tribe staying close to the position of the original big jump it descends from.

The tribe heuristics will give us an idea of why the time to coalescence should be less than $2\ell_N$ in the genealogy, but to prove that the coalescent is in fact star shaped, we also need to prove that there is no coalescence significantly after the common ancestor makes its big jump at time T. To do this, we will use concentration inequalities to prove that the time-$(T+\varepsilon_N\ell_N)$ descendants of the common ancestor have fairly similar numbers of time-t descendants.

1.3 Stretched exponential case

1.3.1 Speed of the particle cloud

In Chapter 3 we describe the behaviour of the speed of the particle cloud in the case when the jump distribution has stretched exponential tails. The motivation to investigate this question comes from the fact that the behaviour of the speed is very different in a light-tailed and in a heavy-tailed setting (see Sections 1.4.2 and 1.4.3). Therefore, we were interested in filling the gap in between these two cases by looking at the intermediate jump distribution with stretched exponential tails.

In order to state our theorem, we need a previous result of Bérard and Gouéré from [2]. In that work the branching and selection steps of the process were the same as in the N-BRW we defined earlier, but negative jumps were also allowed. It was shown that if the jump distribution has some exponential moments, then for any fixed N the particle cloud has a finite deterministic asymptotic speed as time goes to infinity. In this thesis we assume that jumps are non-negative, and in that case we can state this result for any (non-negative) jump distribution with finite mean. We discuss the proof of this result in Chapter 3.

Proposition 1.3.1. *[2, Proposition 2] Consider an N-BRW with arbitrary initial configuration and with a jump distribution given by the non-negative random variable X. Assume that $\mathbb{E}[X]<\infty$. Then for any fixed $N\in\mathbb{N}$, there exists $v_N\in\mathbb{R}$ such that*

$$\lim_{n\to\infty}\frac{\mathcal{X}_1(n)}{n}=\lim_{n\to\infty}\frac{\mathcal{X}_N(n)}{n}=v_N, \tag{1.3.1}$$

almost surely and in L^1, where v_N depends on the jump distribution.

Now we will state our result on the asymptotic behaviour of v_N in the stretched exponential case of the N-BRW. We will assume the jump distribution

$$\mathbb{P}(X > x) = e^{-x^\beta} \quad \text{for all } x \geq 0, \tag{1.3.2}$$

for some fixed $\beta \in (0,1)$. In Chapter 3 we will consider a more general jump distribution, where we replace x^β by a regularly varying function in the exponent. For any $N \in \mathbb{N}$, let v_N denote the asymptotic speed given by (1.3.1) for the jump distribution in (1.3.2). For two positive sequences a_N and b_N we say that $a_N \sim b_N$ as $N \to \infty$, if $a_N/b_N \to 1$ as $N \to \infty$.

Theorem 1.3.2. *Consider an N-BRW with arbitrary initial configuration and with a jump distribution given by* (1.3.2)*. Then*

$$v_N \sim (\log 2)(\log N)^{1/\beta - 1} \text{ as } N \to \infty.$$

The intuition behind this behaviour is the following. First, Theorem 1.3.2 says that the particle cloud moves a distance $(\log N)^{1/\beta}(1 + o(1))$ in $\log_2 N$ time. Now notice that with the jump distribution in (1.3.2), we have

$$\mathbb{P}\left(X > (\log N)^{1/\beta}\right) = 1/N.$$

Since the number of jumps at each time step is $2N$, we will be able to prove that for any $\varepsilon > 0$, there will be a jump larger than $(1-\varepsilon)(\log N)^{1/\beta}$ at time 0 with high probability. Then we can also prove that the population catches up with the particle that made this jump in $\log_2 N$ time (because the number of descendants of this particle will double at each time step until the whole population is to the right of the position the particle jumped to). With this argument we will be able to show that the particle cloud moves at least a distance larger than $(1-\varepsilon)(\log N)^{1/\beta}$ in $\log_2 N$ time, and hence we have the lower bound

$$v_N > (1-\varepsilon)(\log 2)(\log N)^{1/\beta - 1}.$$

It is not hard to see that the largest jump in $A \log N$ time is smaller than $(1+\varepsilon)(\log N)^{1/\beta}$ for any constant (independent of N) $A > 0$ with the jump distribution in (1.3.2). We prove this property in Chapter 3, and then we show the other key step, which says that smaller jumps will not increase the speed of the cloud.

We see this by proving a large deviation result for random walks with stretched exponential tailed jump distribution, but with jumps restricted to being less than $(1+\varepsilon)(\log N)^{1/\beta}$. For this large deviation result we use ideas from work of Gantert [27]. The large deviation result will allow us to conclude that no particle can get further than $(1+\varepsilon)(\log 2)A(\log N)^{1/\beta}$ in $A \log N$ time. From there we will be able to conclude the upper bound

$$v_N < (1+\varepsilon)(\log N)^{1/\beta}.$$

1.3.2 Genealogy

Our motivation to study the genealogy of the N-BRW in the stretched exponential case came again from the fact that the polynomial and light-tailed cases behave very differently. In Section 1.2 we discussed our result which said that the time to coalescence in the N-BRW with polynomial tails is of order $\log N$, and that the coalescent is star-shaped. In the light-tailed case the conjecture is that the genealogy is given by the Bolthausen-Sznitman coalescent and the time to coalescence is of order $(\log N)^3$ (see Section 1.4.4). We are now interested in what happens in between these two cases.

In Chapter 4 we investigate the genealogy with the jump distribution given by (1.3.2), for $\beta \in (0, 1/2)$. Recall the notation $\ell_N = \lceil \log_2 N \rceil$. Our result says that for any large time t, with at least probability of order $(\log N)^{-1/2}$, there exists a time $k \in [t - 2\ell_N, t - \ell_N]$, such that a positive proportion of the time-t population descends from a time-k particle. Chapter 4 provides a summary of the proof, in several cases omitting the details. The chapter also aims to give an explanation of why the result cannot be improved with our current method.

Now we state the main result of Chapter 4. For $k \in \mathbb{N}_0$, $n \geq k$ and $i \in [N]$, let $\mathcal{N}_{i,k}(n)$ denote the set of time-n particles descended from the ith particle from the left at time k (we introduce this notation formally in Chapter 2).

Theorem 1.3.3. *Consider the N-BRW with a jump distribution given by* (1.3.2) *with $\beta \in (0, 1/2)$. There exist $C > 0$ and $c > 0$ such that for N sufficiently large, for all $t > 2\ell_N$,*

$$\mathbb{P}\left(\exists k \in [t - 2\ell_N, t - \ell_N],\ i \in [N] : |\mathcal{N}_{i,k}(t)| \geq cN\right) \geq \frac{C}{(\log N)^{1/2}}.$$

An initial guess for the behaviour in the stretched exponential case can be, that for small values of β it is more similar to the behaviour in the heavy-tailed case, and for β closer to 1, it is more similar to the light-tailed case. In the future we aim to study the case when $\beta \in [1/2, 1)$ and see whether there is a change of behaviour as we change the value of β.

For $\beta \in (0, 1/2)$, the behaviour of the genealogies is a more difficult question than in the polynomial tail case. One of the main reasons for this is the difference in the space scaling. Recall that $\mathcal{X}_N(k)$ and $\mathcal{X}_{N-1}(k)$ denote the positions of the rightmost and second rightmost particles at time k. In the polynomial case, for example in Figure 1-3, we have a particle that takes the lead at some time T, and position $\mathcal{X}_N(T)$. We will prove in the polynomial case that with high probability this particle will lead by a large distance, and $\mathcal{X}_N(T) - \mathcal{X}_{N-1}(T)$ will be of order a_N (see (1.2.3)). This means that $\mathcal{X}_N(T) - \mathcal{X}_{N-1}(T)$ will be of the same order as the size of the big jumps and the typical size of the diameter of the particle cloud. Moreover, the number of jumps of this order in of order $\log N$ time can be upper bounded by a constant with high probability. This means that not many particles will overtake the descendants of the particle that has taken the lead.

In contrast, in the stretched exponential case the size of the largest jump in of order $\log N$ time is of size $(1 + o(1))(\log N)^{1/\beta}$, but the expected number of jumps of size $c(\log N)^{1/\beta}$ is $2N^{1-c^\beta}$ even in a single time step, which is much larger than any constant for $c \in (0, 1)$ and N sufficiently large. As a result, we will need a more precise analysis in the stretched exponential case than in the polynomial one.

In our proof we will investigate a time interval of length $2\ell_N$, so for simplicity, let us fix $t = 2\ell_N$. We will also fix a position y depending on the (arbitrary) initial configuration in such a way, that the expected number of time-$2\ell_N$ particles to the right of position y is less than cN for some small constant $c > 0$.

By our speed result Theorem 1.3.2, we expect that the leftmost position of the cloud moves a distance roughly $2(\log N)^{1/\beta}$ in $2\ell_N$ time. We think of the typical diameter of the particle cloud as roughly $(\log N)^{1/\beta}$, so we imagine that particles need to move a distance between $(\log N)^{1/\beta}$ and roughly $2(\log N)^{1/\beta}$ to survive at time $2\ell_N$. Now note that the largest jump in $2\ell_N$ time steps is of size about $(\log N)^{1/\beta}$. Using this, we can prove that to reach position y at time $2\ell_N$, most particles need to make two 'fairly big jumps' (we will give a definition of these in Chapter 4), and at least one of the fairly big jumps should be of order $(\log N)^{1/\beta}$. We then have to group the paths leading to the right of y by time $2\ell_N$, by their starting positions, and by the times and sizes of the fairly big jumps on the path. Using this detailed analysis we give a lower bound on the probability that a single particle, which makes a large first fairly big jump at a 'good time' and from a 'good starting position', will have $c'N$ time-$2\ell_N$ descendants for some constant $c' > 0$. In Chapter 4 we explain why we initially hoped that the above lower bound on the probability would be a positive constant (independent of N), and why we eventually ended up with a factor of $(\log N)^{-1/2}$.

1.4 Related literature

The works of Bérard and Maillard [3] and Bérard and Gouéré [2] gave the main inspiration for this thesis. In Sections 1.4.2 and 1.4.3 we summarise the ideas of these two essential articles, but before that we give some background on branching random walks below.

1.4.1 Propagation of branching random walks

Consider a BRW as defined in Section 1.1. Let us denote the rightmost particle position in the BRW at time n by $\mathbf{M}(n)$ and recall the notation $\mathcal{X}_N(n)$ in the N-BRW. Observe that $\mathbf{M}(n)$ should typically be larger than $\mathcal{X}_N(n)$ for any fixed N and large n. Indeed, take for example a particle that is killed in the very first selection step in the N-BRW. If we kept that particle, then its descendants could potentially go beyond the descendants of the first N surviving particles later on, and they could lead to a larger value of the rightmost position.

This idea becomes more apparent when one considers a coupling between the N-BRW and N independent BRWs. We will construct this coupling rigorously in Chapter 2, in a similar way to how it was described in [2]. For now we give an illustration of the coupling in Figure 1-4 and an informal definition below. Recall that X denotes the jump distribution. We construct a particle system consisting of blue and red particles.

- Initially we have N blue particles with locations in $\mathbb{R}$.

Then at each time step:

- Each particle is replaced by two offspring.
- Each offspring particle performs a jump from its parents' location, independently from the other jumps, and with the same law as X.
- The N rightmost children of the blue particles are coloured blue.
- The other new offspring particles are coloured red.

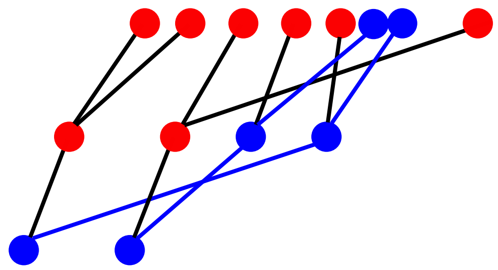

Figure 1-4: Coupling between the N-BRW and N independent BRWs with $N = 2$. The blue particles form an N-BRW.

Now the blue and red particles together form N independent BRWs (each BRW starting from an initial blue particle), and the blue particles form an N-BRW. Furthermore, most importantly, the construction shows that a path of jumps from a particle to a descendant in the N-BRW is a path in one of the BRWs. We will utilise this idea several times in our arguments, and it also shows that the rightmost position in N independent BRWs dominates the rightmost position of the N-BRW. (When we investigate $\mathbf{M}(n)$, we only have a single BRW, but when N is fixed and n is large, this does not really make a difference.)

We now mention results on the behaviour of the rightmost position $\mathbf{M}(n)$ of the BRW. First, when the jump distribution is light-tailed, i.e. has some exponential moments, then the results of Hammersley [29], Kingman [31], and Biggins [8] prove that $\lim_{n\to\infty} \mathbf{M}(n)/n$ exists and determine the limit, which is given by the logarithmic moment-generating function. We will discuss this case further in Section 1.4.3.

If the jump distribution has polynomial (or regularly varying) tails, then Durrett [21] proved that $\mathbf{M}(n)$ grows exponentially fast in n: there exists an exponentially growing sequence s_n such that $\mathbf{M}(n)/s_n$ converges in law to a non-degenerate random variable.

Assuming that the jump distribution is given by $\mathbb{P}(X > x) = \min(1, x^{-\alpha})$ for some $\alpha > 0$, one can give an easy lower bound on $\mathbf{M}(n)$ to see that the propagation should indeed be exponentially fast. For a sequence s_n and for large values of n we have

$$\mathbb{P}(\mathbf{M}(n+1) < s_n) \leq \mathbb{P}(\text{all jumps at time } n \text{ are at most of size } s_n) = (1 - s_n^{-\alpha})^{2^n}.$$

For the right-hand side to be a constant (independent of n) in $(0,1)$, we need s_n to grow as $2^{n/\alpha}$, which suggests an exponentially growing lower bound on $\mathbf{M}(n)$.

In the case when the jump distribution has stretched exponential tails, by results of Gantert [27], we see a superlinear, but polynomial propagation. Assuming that the jump distribution is given by $\mathbb{P}(X > x) = e^{-x^\beta}$ for some $\beta \in (0,1)$, Gantert's result says that, almost surely, $\lim_{n\to\infty} \mathbf{M}(n)/n^{1/\beta} = (\log 2)^{1/\beta}$.

Similarly to the polynomial case, we can give a simple lower bound on $\mathbf{M}(n)$ to have an intuition for this result. For a sequence r_n and for large values of n we have

$$\mathbb{P}(\mathbf{M}(n+1) < r_n) \leq (1 - e^{-r_n^\beta})^{2^n},$$

which will be a constant in $(0,1)$ if r_n is roughly $(n \log 2)^{1/\beta}$.

Since the above result was proved in [27], Dyszewski, Gantert and Höfelsauer have also described the second order behaviour of the rightmost particle [25] and proved a large deviation principle for $\mathbf{M}(n)/n^{1/\beta}$ [24].

Recall the result giving existence of an asymptotic speed for the N-BRW in Proposition 1.3.1. This statement implies that in the N-BRW, the propagation of the cloud of particles is linear in time for the stretched exponential case (where $\mathbb{P}(X > x) = e^{-x^\beta}$ for some $\beta \in (0,1)$) and also for the polynomial case (where $\mathbb{P}(X > x) = \min(1, x^{-\alpha})$) for $\alpha > 1$. Furthermore, in [3] Bérard and Maillard also proved that when $\alpha \in (0,1)$, the propagation is superlinear but at most polynomial in time. We therefore see that with these jump distributions the propagation of the N-BRW is much slower than that of the BRW.

In the case of a light-tailed jump distribution we do not see such a huge difference. The speed of the rightmost particle in the BRW converges to a finite limit in this case. By Proposition 1.3.1, if the jump distribution has positive mean, for any fixed N, the N-BRW also has a positive asymptotic speed as time goes to infinity. In fact, the asymptotic speed v_N of the particle cloud in the N-BRW converges (slowly) to the speed of the rightmost particle in the BRW as N goes to infinity (See Section 1.4.3 for more details).

1.4.2 The limiting process of N-particle branching random walk with polynomial tails [3]

Consider the N-BRW with the initial condition that all particles start from the origin (i.e. $\mathcal{X}_i(0) = 0$ for all $i \in [N]$) and with jump distribution given by

$$\mathbb{P}(X > x) = \min(1, x^{-\alpha}) \quad \text{for all } x \geq 0, \tag{1.4.1}$$

for some $\alpha > 1$. We use this jump distribution to discuss the main ideas of the work of Bérard and Maillard [3], but we remark that the results of the paper are more general: it describes the limiting process and the behaviour of the speed of the N-BRW when the tail of the jump distribution is regularly varying with index α for any $\alpha > 0$. We will make the same more general assumptions on the jump distribution in Chapter 2.

The tribe heuristics (see also Section 1.2.3) suggest that typically we will see the following phenomenon. There will be a big jump which takes the lead, and ℓ_N time later its tribe (consisting of its descendants) becomes the new big tribe, which is close to the leftmost particle position. So we expect the leftmost position to follow the same trajectory as the rightmost, with a lag of ℓ_N generations. The rightmost position only changes significantly when big jumps happen, and then stays the same for a while, until another big jump takes the lead. Therefore, we expect the rightmost trajectory to look like a step function (and therefore the leftmost as well).

Stairs process

The first main result in [3] describes the limiting trajectories of the rightmost and leftmost particles by the so-called *stairs process*, $(\mathcal{R}(t))_{t\geq 0}$. The second main result is about the limiting behaviour of the speed of the particle cloud, which follows from the behaviour of the stairs process.

In order to introduce $\mathcal{R}(t)$, we first explain what we mean by a space-time Poisson point process on $\{(t,x) : \ t, x > 0\}$. Let μ be a σ-finite non-zero measure on $\mathbb{R}_+$ such that $\mu([a, +\infty)) < \infty$ for all $a > 0$. We will call μ the *stairs measure.* Let ν denote the product measure $dt \otimes \mu$ on $\mathbb{R}_+ \times \mathbb{R}_+$; that is, ν is the unique measure such that $\nu([0,t] \times B) = t\mu(B)$ for all $t \geq 0$ and Borel sets $B \subseteq \mathbb{R}_+$.

Let $\Pi \subset \mathbb{R}_+ \times \mathbb{R}_+$ be a random set of points and let $M(A) := |A \cap \Pi|$ denote the number of points from Π falling into A for any Borel set $A \subseteq \mathbb{R}_+ \times \mathbb{R}_+$. Then M is a space-time Poisson point process with intensity measure ν, if for any collection of disjoint Borel sets $(A_k, k \geq 1)$ in $\mathbb{R}_+ \times \mathbb{R}_+$ with $\nu(A_k) < \infty$ for all k,

(1) the random variables $M(A_k)$, $k \geq 1$ are independent, and

(2) $M(A_k)$ has Poisson distribution with parameter $\nu(A_k)$ for all $k \geq 1$.

Let us define the process $(\xi_t)_{t\geq 0}$ by letting $\xi_t = x$ if $(x,t) \in \Pi$ and $\xi_t = 0$ otherwise. Now we are ready to define the stairs process $(\mathcal{R}(t))_{t\in\mathbb{R}}$ with stairs measure μ. The process is constructed inductively. First, for $t \leq 0$, we let $\mathcal{R}(t) = 0$. Now take $n \in \mathbb{N}_0$, and suppose that $\mathcal{R}(t)$ is defined for all $t \leq n$. Then for $t \in (n, n+1]$ we define $\mathcal{R}(t)$ by

$$\mathcal{R}(t) = \max_{s \in [0,1]} (\mathcal{R}(t-1-s) + \xi_{t-s}).$$

Looking at Figure 1-5, this definition is equivalent to the following. Suppose $\mathcal{R}(t)$ is defined for all $t \leq n$. Note that the graph of $\mathcal{R}(t-1)$ (the dashed line in Figure 1-5) is the graph of $\mathcal{R}(t)$ shifted by 1 to the right, and the values of $(\mathcal{R}(t-1))_{t \in (n,n+1]}$ are known if $\mathcal{R}(t)$ is already defined for all $t \leq n$.

Consider the points of Π whose time coordinates fall into the interval $(n, n+1]$. The x coordinate of each of these points is ξ_t for some $t \in (n, n+1]$. Now shift these points vertically by an amount given by the dashed line in Figure 1-5; that is, a point with time coordinate t will be shifted along the (vertical) x axis by $\mathcal{R}(t-1)$. This is how we construct the points in Figure 1-5. Then the function $(\mathcal{R}(t))_{t \in (n,n+1]}$ is given by the record process of the shifted points.

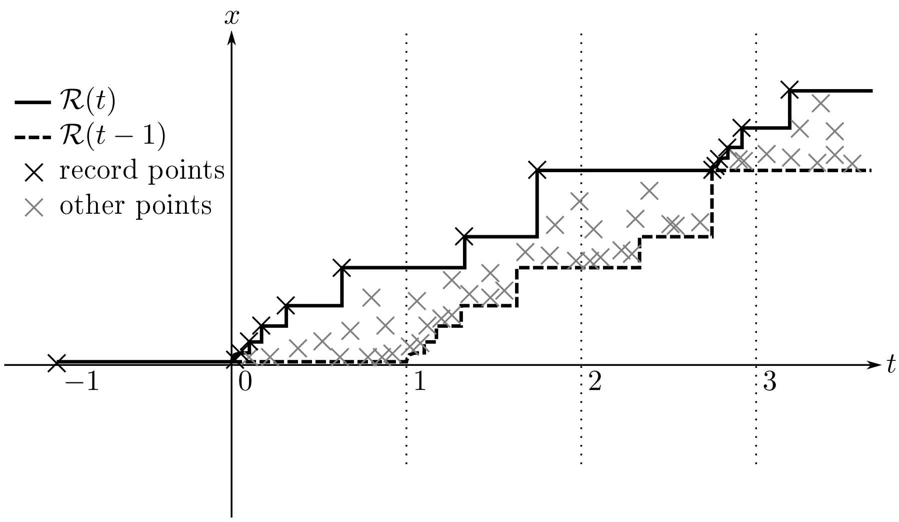

Figure 1-5: Graphical representation of the stairs process. This figure is Figure 1 from [3]. We have a record point at (x, t) if (x, t) is a point in Π shifted by the dashed line (that is, if $\mathcal{R}(t) = x = \mathcal{R}(t-1) + \xi_t$) and if (x, t) is a record; that is, $\mathcal{R}(t) = x > \mathcal{R}(s)$ for all $s < t$. The other points are points of Π shifted by the dashed line which are not record points.

Results

Let $\mathcal{R}^\alpha$ denote a stairs process with stairs measure $\mu = \mu_\alpha$, where $\mu_\alpha([x, \infty)) = x^{-\alpha}$. Recall the definition of a_N from (1.2.3). In the first main result of [3], Bérard and Maillard consider the N-BRW with the scaling where time is sped up by a factor of $\log_2 N$, and space is shrunk by a factor of a_N. Theorem 1.4.1 below says that the scaling limit of the rightmost particle's trajectory is the process $\mathcal{R}^\alpha$ (solid line in Figure 1-5) and that of the leftmost particle's trajectory is given by the $\mathcal{R}^\alpha$ shifted by 1 to the right (dashed line in Figure 1-5). The result is stated as a convergence in law in the J_1 and SM_1 topologies, whose definitions can be found in Chapters 3 and 12 in [45].

Theorem 1.4.1. *[3, Theorem 1.1] Consider the N-BRW with jump distribution given by* (1.4.1) *and with initial configuration $\mathcal{X}_i(0) = 0$ for all $i \in [N]$. Then, as $N \to \infty$,*

$$(a_N^{-1}\mathcal{X}_N(\lfloor t \log_2 N \rfloor))_{t\geq 0} \Longrightarrow (\mathcal{R}^\alpha(t))_{t\geq 0}, \qquad \textit{in the } J_1\textit{-topology},$$
$$(a_N^{-1}\mathcal{X}_N(\lfloor t \log_2 N \rfloor), a_N^{-1}\mathcal{X}_1(\lfloor t \log_2 N \rfloor))_{t\geq 0} \Longrightarrow (\mathcal{R}^\alpha(t), \mathcal{R}^\alpha(t-1))_{t\geq 0} \qquad \textit{in the } SM_1\textit{-topology.}$$

The second main result of [3] describes the limiting behaviour of the speed of the rightmost and leftmost particles, when we let first time and then the number of particles N go to infinity. This is considered to be the main result of the paper.

Theorem 1.4.2. *[3, Theorem 1.2] Consider the N-BRW with jump distribution given by* (1.4.1) *and with initial configuration $\mathcal{X}_i(0) = 0$ for all $i \in [N]$. Then the limit*

$$v_N = \lim_{n\to\infty} \frac{\mathcal{X}_N(n)}{n} = \lim_{n\to\infty} \frac{\mathcal{X}_1(n)}{n}$$

exists almost surely and in L^1, and it satisfies

$$v_N \sim \rho_\alpha \frac{a_N}{\log_2 N},$$

where the limit $\rho_\alpha = \lim_{t\to\infty} \mathcal{R}^\alpha(t)/t$ exists almost surely and in L^1.

For the proofs of both theorems the most important idea is to couple the N-BRW with the stairs process, and investigate the limiting behaviour of the stairs process. Then the theorem follows by putting together the coupling results and the results on the stairs process. Below we summarise the heuristic ideas from [3] which describe how one should think of the stairs process in terms of the N-BRW.

Tribes and the stairs process

The heuristic argument in [3] says (and in Chapter 2 we prove) that at a typical time, on the a_N space scale, $N - o(N)$ particles are very close to the leftmost particle forming a big tribe, and the other particles are in small tribes of size $o(N)$ further to the right of the big tribe. Recall that new tribes are created when a particle makes a big jump of size at least of order a_N. The heuristics also suggest that the points of the Poisson point process Π correspond to the big jumps of the N-BRW after rescaling. More precisely, the record points (see Figure 1-5) correspond to particles which perform a big jump of order a_N, and whose descendants eventually take over the population. The other points correspond to big jumps whose descendants die out before their tribe reaches size of order N. The dashed line shows the leftmost trajectory and also the trajectory of the big tribe. We expect big jumps to come from the big tribe (since there are very few particles outside of the big tribe), therefore the big jumps (i.e. points of the Poisson point process Π) are shifted by the location of the big tribe (i.e. the dashed line). The solid line shows the rightmost trajectory.

Note that we should indeed expect that on the a_N space scale the location of the big tribe at some time t is where the rightmost or 'leader' particle was at time $t - \ell_N$, since the time-t leader has enough time to build a tribe of size roughly N by time $t - \ell_N$, whereas particles that break the record between times $t - \ell_N$ and t do not. We have illustrated this idea in Figure 1-3.

We also remark that Theorem 1.4.1 justifies the heuristics for the trajectories of the leftmost and rightmost particles, but the other parts of the heuristic picture are not proven in [3]. In Chapter 2 we prove that the descendants of a particle that takes the lead with a big jump can indeed take over the population $\log_2 N$ time later, and that the location of the big tribe at some time t is indeed close to the leftmost particle at time t, and also close to the rightmost particle at time $t - \ell_N$. We did not prove however, that the tribes of the big jumps whose descendants take over the population correspond to the record points of the Poisson point process introduced in [3]. One of the reasons is that, to prove our genealogy result, we needed to precisely work out the details of the N-BRW in a time interval of length $2\ell_N$, as opposed to comparing the whole process to the Poisson point process.

1.4.3 Brunet-Derrida behavior of branching-selection particle systems on the line [2]

Bérard and Gouéré's paper describes the limiting behaviour of the speed of the particle cloud when the jump distribution is light-tailed in the N-BRW. In this section we will specify what we mean by a light-tailed jump distribution. The main result of [2] says that v_N, the asymptotic speed given by (1.3.1), converges to a finite limit as N goes to infinity with a very slow rate of $(\log N)^{-2}$. The title of the paper refers to this slow convergence rate, which was predicted by Brunet and Derrida, and which has appeared in other models as well, suggesting that this is a universal behaviour. We will expand on this topic in Section 1.4.4.

Consider an N-BRW as described in Section 1.1.1 but with the modification that negative jumps are allowed, i.e. the jump distribution X is not necessarily non-negative. We now describe the assumptions on the jump distribution in [2]. Let p denote the probability measure that describes the jump distribution:

$$\mathbb{P}\left(X > a\right) = p((a, \infty)) \quad \text{for } a \in \mathbb{R}.$$

Furthermore, let $\Lambda(t)$ denote the logarithmic moment-generating function of p,

$$\Lambda(t) := \log \int_{\mathbb{R}} \exp(tx) dp(x).$$

We make the following assumptions on p:

(A) The number $\sigma := \sup\{t \geq 0;\ \Lambda(-t) < +\infty\}$ is positive.

(B) The number $\zeta := \sup\{t \geq 0;\ \Lambda(t) < +\infty\}$ is positive.

(C) There exists $t^* \in (0, \zeta)$ such that $t^*\Lambda'(t^*) - \Lambda(t^*) = \log 2$.

It can be checked that under these assumptions, both

$$\chi(p) := \frac{\pi^2}{2} t^* \Lambda''(t^*) \text{ and } v(p) := \Lambda'(t^*) \tag{1.4.2}$$

are well-defined and satisfy $0 < \chi(p) < \infty$ and $v(p) \in \mathbb{R}$. Cases when the above assumptions are satisfied include for example when p is the uniform distribution on $[0, 1]$, when p is the standard Gaussian distribution, and when p is Bernoulli with $q \in (0, 1/2)$ and $p = q\delta_1 + (1 - q)\delta_0$. Interestingly, for $q \geq 1/2$, assumption (C) does not hold in the Bernoulli case, and it is shown in [2] that the behaviour indeed changes for $q \geq 1/2$: we see faster convergence for v_N than the Brunet-Derrida behaviour.

The first result in [2] says that a finite deterministic asymptotic speed $v_N(p)$ as in (1.3.1) exists, if the assumptions (A)-(B)-(C) hold (and the jump distribution is not necessarily non-negative). The main result states the limit and convergence rate of $v_N(p)$ to the limit $v(p)$ as defined in (1.4.2).

Theorem 1.4.3. *[2, Theorem 1] Assume that (A)-(B)-(C) hold. Then,*

$$v(p) - v_N(p) \sim \chi(p)(\log N)^{-2} \quad \textit{as } N \to \infty.$$

The limiting speed $v(p)$ coincides with the asymptotic speed of the rightmost particle of a BRW without selection as time goes to infinity. The intuition for the speed of the BRW without selection comes from Cramér's large deviation theorem (see e.g. Theorem 2.2.3 in [18]). Let us use the informal notation $\approx$ to indicate that the two sides are close to each other in some sense. Let X_i be i.i.d. random variables distributed as X, and let $S_n := \sum_{i=1}^n X_i$. Then the message of Cramér's theorem is that for fixed $x > \mathbb{E}[X]$, for large n,

$$\mathbb{P}(S_n > xn) \approx \exp(-n\Lambda^*(x)),$$

where $\Lambda^*(x) = \sup_{t\in\mathbb{R}}(tx - \Lambda(t))$.

If we choose x in such a way that the expected number of paths arriving to the right of position nx is about 1 at time n in the BRW, then x should be roughly the speed of the rightmost particle in the BRW, and so we gain an intuition for the definition of $v(p)$ in (1.4.2). Consider x such that

$$2^n \mathbb{P}(S_n > nx) = 1. \tag{1.4.3}$$

Let $t^{**} := \arg\max_{t\in\mathbb{R}}(tx-\Lambda(t))$. Then t^{**} satisfies $\Lambda'(t^{**}) = x$. Hence (1.4.3) and Cramér's theorem imply

$$2^n \approx e^{n(t^{**}\Lambda'(t^{**})-\Lambda(t^{**}))}.$$

That is, t^{**} satisfies $\log 2 = t^{**}\Lambda'(t^{**}) - \Lambda(t^{**})$, which is the equation in (C), and thus $t^{**} = t^*$. Therefore, the speed should be indeed $x = \Lambda'(t^{**}) = \Lambda'(t^*)$, where t^* satisfies (C), i.e. the speed should be given by (1.4.2).

Ideas of the proof of Theorem 1.4.3

The coupling between the N-BRW and N independent BRWs is one of the most important tools for proving that the speed of the rightmost particle in the N-BRW converges to the same speed as the rightmost particle of a BRW. Furthermore, the proof focuses on paths which move with a consistent speed rather than on single big jumps like in the previous heavier tailed cases. The main concept in the proof of the $(\log N)^{-2}$ convergence rate is as follows.

Assume we have a path which consists of i.i.d. jumps of size $X_1, \dots, X_n$. We say that the path is consistently above the line of slope $v > 0$ until time $n \in \mathbb{N}$ if we have $S_k > kv$ for all $k \in [n]$. Let us call such a path an (n, v)-good path.

An essential result that is used in the proof of [2] is an estimate for the probability $\rho(n, \varepsilon)$ of the existence of an $(n, v(p) - \varepsilon)$-good path in a BRW between times 0 and n, where $v(p)$ is the speed defined in (1.4.2). The result in [28] says that for all large n,

$$\rho(n, \varepsilon) \approx e^{-\sqrt{\chi(p)/\varepsilon}} = e^{-\sqrt{\frac{\chi(p)}{\theta}}n^{1/3}}, \tag{1.4.4}$$

if $\varepsilon := \theta n^{-2/3}$ for some well-chosen $\theta > 0$, and where $\chi(p)$ is defined in (1.4.2). We remark here that this is not exactly the message of the main result of [28]; this statement follows from their proof, and was cited as a key theorem in [2] (where the theorem was stated rigorously).

Let $\mathbf{A}_n^N$ denote the event that there exists an $(n, v(p) - \varepsilon)$-good path between times 0 and n in at least one of N independent BRWs. Then

$$\mathbb{P}\left(\mathbf{A}_n^N\right) = 1 - (1 - \rho(n, \varepsilon))^N,$$

and using (1.4.4), this is of constant order in N if $n = n_N \sim \left(\frac{\theta}{\chi(p)}\right)^{3/2} (\log N)^3$ and so $\varepsilon = \varepsilon_N \sim \chi(p)(\log N)^{-2}$.

The idea to connect these heuristics to the speed of the N-BRW is the following. If we take $\tilde{\varepsilon}_N$ significantly larger than ε_N above, then we expect that the number of $(n_N, v(p) - \tilde{\varepsilon}_N)$-good paths in N independent BRWs will be large if N is large. Now using this fact and the coupling from Section 1.4.1, it is shown in [2] that the trajectory of the leftmost position $(\mathcal{X}_1(n))_{n\geq 0}$ in the N-BRW will never stay consistently *below* the line of slope $v(p) - \tilde{\varepsilon}_N$, which will imply that the average speed is larger than $v(p) - \tilde{\varepsilon}_N$. (The proofs of these steps are far from straightforward.)

On the other hand, if we take $\hat{\varepsilon}_N$ significantly smaller than ε_N above, then we expect that the number of $(n_N, \hat{\varepsilon}_N)$-good paths in N independent BRWs will be small if N is large. Then making use of the coupling again, it can be shown that a speed of $v(p) - \hat{\varepsilon}_N$ is not sustainable in the N-BRW.

1.4.4 Further related models

We end this section by discussing some other models, results and conjectures that are closely related to the N-BRW. Several results in the area of branching processes are stated in continuous time, for the branching Brownian motion (BBM) rather than for branching random walks. The (one-dimensional) BBM is defined as follows (using the definition e.g. from [43]):

- At time $t = 0$ there is a single particle at the origin, which starts to move as a standard one-dimensional Brownian motion.
- The lifetime of the particle is random, and has exponential distribution with mean 1.
- When the particle dies, it produces two new particles (i.e. splits into two).
- The new particles move as independent Brownian motions, each having an exponentially distributed lifetime with mean 1.
- The same splitting rule applies for the new particles as for the original particle.
- The system goes on indefinitely.

The continuous time analogue of the N-BRW is called the N-particle branching Brownian motion (N-BBM) proposed by Maillard [34]. The N-BBM is defined as a BBM, in which each time the number of particles exceeds N, only the N rightmost particles survive and the others are killed instantaneously.

In the following we review some of the conjectures and results on discrete and continuous time branching processes with selection. Brunet and Derrida [14, 15] introduced a discrete time particle system with selection which is similar to the N-BRW but uses a slightly different selection mechanism. We will discuss some of the authors' solved and open conjectures related to this and similar models. We also note that the behaviours described below are conjectured to be universal and should not depend on the details of the specific model.

Connection with the FKPP equation

The FKPP (Fisher, Kolmogorov, Petrovsky and Piscounov) equation was first introduced to model the spread of favourable genes in a population [26, 33], and is given by

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1-u),$$

where $u = u(x,t)$, $x \in \mathbb{R}$ and $t \geq 0$. This reaction-diffusion equation, which admits travelling wave solutions, is closely related to branching particle systems. The connection originates from the duality relation by Skorohod [44] and McKean [36] between the solution of the FKPP equation and the position of the rightmost particle of a BBM. Several results on the FKPP equation and the BBM can be found in [13].

One of the conjectures of Brunet and Derrida is that the large-scale behaviour of the branching-selection system in [15] is given by an analogue of the FKPP equation. In particular, the authors argue that the fraction of particles to the right of a position x at time t should behave like the solution of an FKPP type equation.

Since then, several rigorous results have been proved on the relation of particle systems with selection and free boundary problems with travelling wave solutions, such as [23] and [17, 5]. The result in [17] is about the N-BBM: it says that the empirical measure of the N-BBM converges to the solution of a free boundary problem of FKPP type; then the global existence of solutions to this free boundary problem is shown in [5]. The work [23] is similarly about the convergence of the empirical measure, in this case for a branching random walk with selection (slightly different from the N-BRW that we are investigating).

Minimal velocity solution

Brunet and Derrida also had conjectures on the speed of the particle cloud in the branching-selection system they examined in [14, 15]. They argue that the speed of the rightmost particle should converge to an asymptotic speed v_N as time increases, and then v_N should converge to a limit as N goes to infinity. This limit is given by the solution of an analogue of the FKPP equation with minimal velocity, and the rate of convergence is very slow, $(\log N)^{-2}$. These conjectures have been proved in [2] as we have discussed in Section 1.4.3. A similar result has been proved by Pain on the continuous time model called the L-BBM, in which particles are killed when they are at distance L from the rightmost particle [39].

The original heuristic reasoning for the convergence rate of $(\log N)^{-2}$ in [14, 15] involved considering the FKPP equation with cut-off:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1-u)\mathbb{1}_{\{u \geq 1/N\}}.$$

The cut-off accounts for the fact that the number of particles is finite, and so a proportion of particles cannot be in $(0, 1/N)$ if the number of particles is N. The authors give a heuristic argument which shows that the travelling wave solutions of the FKPP equation with cut-off are shifted from those of the FKPP equation by $(\log N)^{-2}$. A few years later, this statement was proved rigorously [20]. Furthermore, there have been rigorous results showing the $(\log N)^{-2}$ convergence rate for the FKPP equation with random noise as well [38].

Coalescent processes

The third type of conjecture made by Brunet and Derrida is about the genealogies of the branching-selection system. The genealogies of a population can be described by coalescent processes; stochastic processes whose state space is the set of partitions of $\mathbb{N}$ or $[M]$ for some $M \in \mathbb{N}$. The blocks of the partitions represent groups of particles with a common ancestor. In a coalescent process, at a given time, a set of blocks merge into one block if the particles in all the blocks have the same ancestor at that time.

For example, in Figure 1-2, if we label the ancestral lines by $1, 2, \ldots, 6$, then the state of the coalescent process at time t is $\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$, after the first merging event it is $\{\{1\}, \{2, 3\}, \{4\}, \{5\}, \{6\}\}$, and by time T it is $\{\{1, 2, 3, 4, 5, 6\}\}$.

A classical example of a coalescent process is Kingman's coalescent [32], which is a continuous time Markov chain on the set of partitions, in which the only allowed transitions are that two of the present blocks merge into one, and this occurs independently at rate one for each pair of blocks. Kingman's coalescent describes the genealogies of classical models such as the Moran [37] and Wright-Fisher models (see e.g. [6]); and in general, for models in which the population size is constant, particles typically have few offspring, and no selection is involved, Kingman's coalescent is expected to be a universal scaling limit (see e.g. [6]).

A more general setting is necessary, if selection is included in the model. Kingman's coalescent only allows two blocks to merge at a time. However, we have seen in our genealogy result in Section 1.2.2, that on a $\log N$ scale all coalescences of the lineages occurred very close to each other in time, thus in the limit as $N \to \infty$ we saw multiple ancestral lines merging at the same time.

The more general setting is the so-called Λ-coalescent [41], which allows multiple mergers. The canonical example of a Λ-coalescent is the Bolthausen-Sznitman coalescent [10], which is conjectured to appear as a universal scaling limit of population models with selection. For the genealogy of the N-BRW with light-tailed jump distribution (and for other similar branching-selection systems), the papers of Brunet, Derrida, Mueller and Munier [11, 12] arrived at the following conjecture (see also [34] by Maillard). If we pick two particles at random in a generation, then the number of generations we need to go back to find a common ancestor of the two particles is of order $(\log N)^3$. Furthermore, if we take a uniform sample of k particles in a generation and trace back their ancestral lines, the coalescence of their lineages is described by the Bolthausen-Sznitman coalescent, if time is scaled by $(\log N)^3$. For the N-BRW and its continuous time analogue, the N-BBM, this conjecture is still open.

There are, however, rigorous results in the literature on convergence to the Bolthausen-Sznitman coalescent for models with different branching or selection rules. The papers [11, 12] show this property for an exactly solvable model in which particles reproduce according to a Poisson point process, and in the selection step the N rightmost offspring of all individuals are kept. Cortines and Mallein proved a generalisation of this result with random selection [16].

J. Berestycki, N. Berestycki and Schweinsberg proved an important result on convergence to the Bolthausen-Sznitman coalescent for the model called BBM with absorption [4], where particles are killed when hitting a deterministic moving boundary, which is defined in such a way that the process has approximately N particles at large times.

Furthermore, Schweinsberg [42] also proved convergence to the Bolthausen-Sznitman coalescent for a model with fixed sized population, in which individuals gain beneficial mutations at a certain rate, which increases their fitness, and individuals with larger fitness are more likely to reproduce.

1.5 Outline

Chapter 2 is the article [40] about the result discussed in Section 1.2, which is joint work with Sarah Penington and Matthew Roberts. Chapters 3 and 4 are based on joint work with Sarah Penington; in Chapter 3 we prove the result from Section 1.3.1, in Chapter 4 we give a summary of the proof of the result stated in Section 1.3.2, and discuss some questions for the future.

Chapter 2

Genealogy and spatial distribution of the N-particle branching random walk with polynomial tails

Inspired by work of J. Bérard and P. Maillard, we examine the long term behaviour of the N-BRW in the case where the jump distribution has regularly varying tails and the number of particles is large. We prove that at a typical large time the genealogy of the population is given by a star-shaped coalescent, and that almost the whole population is near the leftmost particle on the relevant space scale. This is joint work with Sarah Penington and Matthew Roberts and appears in [40].

2.1 Introduction

2.1.1 The N-BRW model

We investigate a particle system called N-particle branching random walk (N-BRW). In this discrete time stochastic process, at each time step, we have N particles located on the real line. We say that the particles at the nth time step or at time n belong to the nth generation. The locations of the particles change at every time step according to the following rules. Every particle has two offspring. The offspring particles have random independent displacements from their parents' locations, according to some prescribed displacement distribution supported on the non-negative real numbers. Then from the $2N$ offspring particles, only the N particles with the rightmost positions survive to form the next generation. That is, at each time step we have a *branching step* in which the $2N$ offspring particles move, and we have a *selection step*, in which N out of the $2N$ offspring are killed. Ties are decided arbitrarily. We describe the process more formally in Section 2.2.1.

We will use the notation $[N] := \{1, \ldots, N\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ throughout. A pair (i, n) with $i \in [N]$ and $n \in \mathbb{N}_0$ will represent the ith particle from the left in generation n. We also refer to the rightmost particle (N, n) as the *leader* at time n. Furthermore, we will denote the locations of the N particles in the nth generation by the ordered set of N real numbers

$$\mathcal{X}(n) = \{\mathcal{X}_1(n) \leq \cdots \leq \mathcal{X}_N(n)\}, \tag{2.1.1}$$

where $\mathcal{X}_i(n)$ is the location of particle (i,n). We sometimes call $\mathcal{X}(n)$ the *particle cloud.*

The long term behaviour of the N-BRW heavily depends on the tail of the displacement distribution. Motivated by the work of Bérard and Maillard [3], we investigate the N-BRW in the case where the displacement distribution is regularly varying, and N is large.

We say that a function f is regularly varying with index $\alpha \in \mathbb{R}$ if for all $y > 0$,

$$\frac{f(xy)}{f(x)} \to y^\alpha \text{ as } x \to \infty. \tag{2.1.2}$$

Let X be a random variable and let the function h be defined by

$$\mathbb{P}(X > x) = \frac{1}{h(x)} \text{ for } x \geq 0. \tag{2.1.3}$$

We assume throughout that $\mathbb{P}(X \geq 0) = 1$, that h is regularly varying with index $\alpha > 0$, and that the displacement distribution of the N-BRW is given by (2.1.3). These are the same assumptions under which the results of [3] were proved. The reader may wish to think of the particular regularly varying function given by $h(x) = x^\alpha$ for $x \geq 1$ and $h(x) = 1$ for $x \in [0,1)$. We do not expect significant change in the behaviour of the N-BRW if jumps of negative size are allowed, but we do not prove this; we use the assumption that the jumps are non-negative several times in our argument.

2.1.2 Time and space scales

Before explaining our main result, we describe the time and space scales we will be working with. We define

$$\ell_N := \lceil \log_2 N \rceil, \tag{2.1.4}$$

for $N \geq 2$; this is the time scale we will be using throughout. To avoid trivial cases we always assume that $N \geq 2$. The time scale ℓ_N is the time it takes for the descendants of one particle to take over the whole population, if none are killed in selection steps.

For the space scale we choose

$$a_N := h^{-1}(2N\ell_N), \tag{2.1.5}$$

where h is as in (2.1.3), and h^{-1} denotes the generalised inverse of h defined by

$$h^{-1}(x) := \inf\{y \geq 0 : \ h(y) > x\}. \tag{2.1.6}$$

It is worth thinking of the particular case $h(x) = x^\alpha$ for $x \geq 1$, for which we have $a_N = (2N\ell_N)^{1/\alpha}$ and $h(a_N) = 2N\ell_N$.

With the choice of a_N in (2.1.5), for any positive constant c, the expected number of jumps which are larger than ca_N in a time interval of length ℓ_N is of constant order, as N goes to infinity. The heuristic picture in [3] says that jumps of order a_N govern the speed, the spatial distribution, and the genealogy of the population for N large. Besides the main result of [3] on the asymptotic speed of the particle cloud, it is conjectured that at a typical time the majority of the population is close to the leftmost particle, and that the genealogy of the population is given by a star-shaped coalescent. In this paper we prove these conjectures.

2.1.3 The main result (in words)

Stating our main result precisely involves introducing some more notation and defining some rather intricate events. We will do this in Section 2.2. In this section we instead aim to explain the main message of the theorem. When we say 'with high probability', we mean with probability converging to 1 as $N \to \infty$.

For all $\eta > 0$, $M \in \mathbb{N}$ and $t > 4\ell_N$, the N-BRW has the following properties with high probability:

- **Spatial distribution:** *At time t there are $N - o(N)$ particles within distance ηa_N of the leftmost particle, i.e. in the interval $[\mathcal{X}_1(t), \mathcal{X}_1(t) + \eta a_N]$.*
- **Genealogy:** *The genealogy of the population on an ℓ_N time scale is asymptotically given by a star-shaped coalescent, and the time to coalescence is between ℓ_N and $2\ell_N$.*

  That is, there exists a time $T \in [t - 2\ell_N, t - \ell_N]$ such that with high probability, if we choose M particles uniformly at random at time t, then every one of these particles descends from the rightmost particle at time T. Furthermore, with high probability no two particles in the sample of size M have a common ancestor after time $T + \varepsilon_N \ell_N$, where ε_N is any sequence satisfying $\varepsilon_N \to 0$ and $\varepsilon_N \ell_N \to \infty$, as $N \to \infty$.

The star-shaped genealogy might seem counter-intuitive because every particle has only two descendants. Indeed, if we take a sample of $M > 2$ particles at time t, and look at the lineages of these particles, they certainly cannot coalesce in one time step. Our result says that all coalescences of the lineages of the sample occur within $o(\ell_N)$ time. Therefore, looking on an ℓ_N time scale the coalescence appears instantaneous.

2.1.4 Heuristic picture

We construct our heuristic picture based on the tribe heuristics for the N-BRW with regularly varying tails described in [3]. The tribe heuristics say that at a typical large time there are $N - o(N)$ particles close to the leftmost particle if we look on the a_N space scale. We call this set of particles the *big tribe*. Furthermore, there are *small tribes* of size $o(N)$ to the right of the big tribe. The number of such small tribes is $O(1)$. While the position of the big tribe moves very little on the a_N space scale, the number of particles in the small tribes doubles at each time step. As a result, the big tribe eventually dies out, and one of the small tribes grows to become the new big tribe and takes over the population.

To escape the big tribe and create a new tribe that takes over the population, a particle must make a big jump of order a_N. As we explained in Section 2.1.2, jumps of this size occur on an ℓ_N time scale, and ℓ_N is the time needed for a new tribe to grow to a big tribe of size N.

Take $t > 4\ell_N$. Building on the tribe heuristics, we describe the following picture. Assume that a particle becomes the leader with a big jump of order a_N. We claim that this particle will have of order N surviving descendants ℓ_N time after the big jump. Moreover, the particle that makes the last such jump before time $t_1 := t - \ell_N$ will be the common ancestor of the majority of the population at time t. We denote the generation of this ancestor particle by T, and assume that $T \in [t_1 - \ell_N, t_1]$. In Figure 2-1 we illustrate how a new tribe is formed at time T, and how it grows to a big tribe by time t. We will prove the main result described in Section 2.1.3 by showing that the picture in Figure 2-1 develops with high probability.

We introduce the notation

$$t_i := t - i\ell_N, \tag{2.1.7}$$

for $t, i \in \mathbb{N}$. The message of Figure 2-1, which we will prove later, is that the following occurs with high probability.

A: At time $T \in [t_2, t_1]$, particle (N, T) has taken a big jump of order a_N and escaped the big tribe. It now leads by a large distance, and its descendants will be the leaders at least until time t_1.

There are two main reasons for this. First, we define T as the last time before time t_1 when a big jump of order a_N creates a new leader, so particles with big jumps in the time interval $[T, t_1]$ cannot become leaders. Second, particles with smaller jumps not descending from particle (N, T) are unlikely to catch up with the leading tribe, because paths with small jumps move very little on the a_N space scale. This is an important property of random walks with regularly varying tails, which we will state and prove in Lemma 2.4.3 and apply in Corollary 2.4.5.

B: After time t_1, there might be particles which do not descend from particle (N, T), but which, by making a big jump of order a_N, move beyond the tribe of particle (N, T). However, these particles have substantially less than ℓ_N time to produce descendants by time t, and so each of them can only have $o(N)$ descendants at time t. Particles which do not descend from (N, T) are unlikely to move beyond the tribe of particle (N, T) without making a big jump.

There will only be $O(1)$ big jumps of order a_N between times t_1 and t, because jumps of order a_N happen with frequency of order $1/\ell_N$. Therefore, until time t, the total number of particles to the right of the tribe of particle (N, T) is at most $o(N)$.

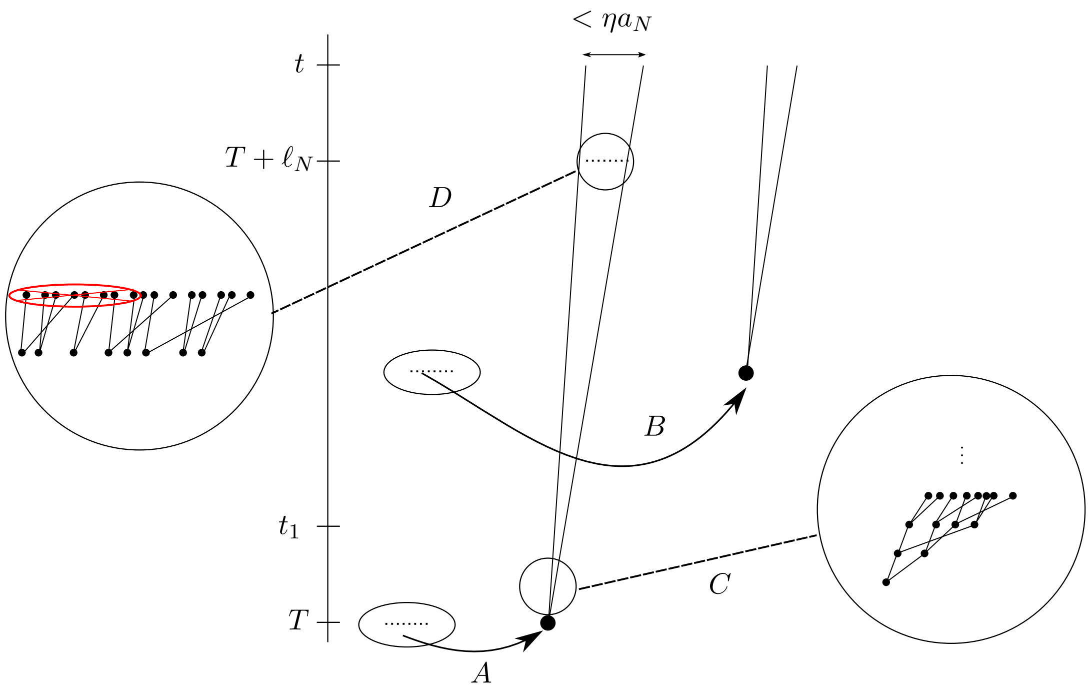

Figure 2-1: A particle that makes a big jump of order a_N at time T is the common ancestor of almost the whole population at time t. The vertical axis represents time, and the particles' locations are depicted horizontally, increasing from left to right. The black dots represent particles. Horizontal dotted lines in an ellipse or circle show where the majority of the population (the big tribe) is. The arrows represent jumps from the big tribe. We use circles to zoom in on the population. The particles circled in red are killed in the selection step. The events labelled A to D are described in the main text.

C: The tribe of particle (N, T) doubles in size at each step up to (almost) time $T + \ell_N$. Selection does not affect these particles significantly, because the number of particles to the right of this tribe is at most $o(N)$ before time $T + \ell_N$, as we explained in part B.

D: At time $T + \ell_N$ there are N particles to the right of the position of particle (N, T). This is an elementary property of the N-BRW, following from the non-negativity of the jump sizes. The N particles are mainly in the tribe of particle (N, T), and there may be $o(N)$ particles ahead of the tribe. From this point on, the N leftmost offspring particles in the tribe of particle (N, T) do not survive.

Then, between times $T + \ell_N$ and t, the number of particles in the tribe of particle (N, T) will remain $N - o(N)$, where the $o(N)$ part doubles at each time step but does not reach order N by time t. Therefore, almost every particle at time t descends from particle (N, T).

Furthermore, as the number of descendants of particle (N, T) only reaches order N at (roughly) time $T+\ell_N$, the descendants of particle (N,T) are unlikely to make big jumps of order a_N before time $T+\ell_N$. We will prove this property (and many others) in Lemma 2.4.6. Only $O(1)$ descendants of particle (N,T) make big jumps of order a_N between times T and t, and these big jumps are likely to happen after time $T+\ell_N$, and so significantly after time t_1. Therefore, most time-t descendants of particle (N,T) will not have an ancestor which made a big jump between times T and t, thus they will not move far from their ancestor's position $\mathcal{X}_N(T)$ on the a_N space scale.

In order to prove our statements in Section 2.1.3 we also need to show that there is at least one particle which becomes the new leader with a jump of order a_N during the time interval $[t_2,t_1]$. The existence of such a particle will imply that indeed there exists $T \in [t_2,t_1]$ as in Figure 2-1. We give a heuristic argument for this in Section 2.2.3, where we also explain the idea for proving that if we take a sample of M particles at time t then the coalescence of the ancestral lineages of these particles happens within a time window of width $o(\ell_N)$.

2.1.5 Optimality of our main result

In order to show that our main result is more or less optimal, we will prove two additional results.

Spatial distribution: Our main theorem says that *most* particles in the population are likely to be within distance ηa_N of the leftmost at time t, for arbitrarily small $\eta > 0$ when N and t are large. We will show that this is not true of *all* particles: the distance between the leftmost and rightmost particles is typically of order a_N, and is arbitrarily large on the a_N space scale with positive probability. Therefore our result that most particles are close to the leftmost particle on the a_N space scale gives meaningful information on the shape of the particle cloud at a typical time. We state this formally in Proposition 2.2.2 and then prove it in Section 2.6.

Genealogy: Our main theorem says that the generation T of the most recent common ancestor of a sample from the population at time t is between times t_2 and t_1 with high probability. We will prove that this is the strongest possible result in the sense that for any subinterval of $[t_2,t_1]$ with length of order ℓ_N there is a positive probability that T is in that subinterval. This will be the main message of Proposition 2.2.1, which we prove in Section 2.6.

We also mention here that the precise statement of our main result, Theorem 2.2.1, implies that the distribution of the rescaled time to coalescence, $(t-T)/\ell_N$, has no atom at 1 or 2 in the limit $N \to \infty$.

2.1.6 Related work

The N-BRW shows dramatically different behaviours with different jump distributions; this includes the speed at which the particle cloud moves to the right, the spatial distribution within the population, and the genealogy. Below we discuss existing results and conjectures on these properties of the N-BRW. We start by summarising the results of Bérard and Maillard, who studied the speed of the particle cloud when the displacement distribution is heavy-tailed.

Heavy-tailed displacement distribution

Bérard and Maillard [3] introduced the stairs process, the record process of a shifted space-time Poisson point process. They proved that it describes the scaling limit of the pair of trajectories of the leftmost and rightmost particles' positions $(\mathcal{X}_1(n), \mathcal{X}_N(n))_{n\in\mathbb{N}_0}$ when the jump distribution has polynomial tails. The correct scaling is to speed up time by $\log_2 N$ and to shrink the space scale by a_N. Using the relation between the N-BRW and the stairs process they prove their main result: the speed of the particle cloud grows as $a_N/\log_2 N$ in N, and the propagation is linear or superlinear (but at most polynomial) in time. The propagation is linear if the jump distribution has finite expectation, and superlinear otherwise; the asymptotics follow from the behaviour of the stairs process. This behaviour is different from that of the classical branching random walk without selection, where the propagation is exponentially fast in time in a heavy-tailed setting [21].

The tribe heuristics in [3] predict—but do not prove—that the majority of the population is located close to the leftmost particle, that the genealogy should be star-shaped, and that the relevant time scale for coalescence of ancestral lineages is ℓ_N. We will prove the above properties in Theorem 2.2.1, and therefore the present paper and [3] together provide a comprehensive picture of the N-BRW with regularly varying tails, including the behaviour of the speed, spatial distribution and genealogy.

Light-tailed displacement distribution

Particle systems with selection have been studied with light-tailed displacement distribution in the physics literature as a microscopic stochastic model for front propagation. First Brunet and Derrida [14, 15], and later Brunet, Derrida, Mueller and Munier [12, 11] made predictions on the behaviour of particle systems with branching and selection.

Speed: For the N-BRW, Bérard and Gouéré [2] proved the existence of the asymptotic speed of the particle cloud as time goes to infinity, which in fact applies for any jump distribution with finite expectation. They also proved that the asymptotic speed converges to a finite limiting speed as the number of particles N goes to infinity, with a surprisingly slow rate $(\log N)^{-2}$, which was predicted by Brunet and Derrida [14, 15]. The limiting speed is the same as the speed of the rightmost particle in a classical branching random walk without selection with exponentially decaying tails [29, 31, 8].

Spatial distribution: The spatial distribution in the light-tailed case is also predicted in [14, 15]. The authors argue that the fraction of particles to the right of a given position at a given time should evolve according to an analogue of the FKPP equation. The FKPP equation is a reaction-diffusion equation admitting travelling wave solutions. Rigorous results on the relation between particle systems with selection and free boundary problems with travelling wave solutions have been proved in [23] and [5, 17].

Genealogy: On the genealogy of the N-BRW with light-tailed displacement distribution, the papers [12, 11] arrived at the following conjecture (see also [34]). If we pick two particles at random in a generation, then the number of generations we need to go back to find a common ancestor of the two particles is of order $(\log N)^3$. Furthermore, if we take a uniform sample of k particles in a generation and trace back their ancestral lines, the coalescence of their lineages is described by the Bolthausen-Sznitman coalescent, if time is scaled by $(\log N)^3$. This property has been shown for a continuous time model, a branching Brownian motion (BBM) with absorption [4], where particles are killed when hitting a deterministic moving boundary. For the N-BRW and its continuous time analogue, the N-BBM, no rigorous proof has yet been given.

Displacement distribution with stretched exponential tail

As we have seen, the behaviour of the N-BRW is significantly different in the light-tailed and heavy-tailed cases. It is then a natural question to ask what happens in an intermediate regime, where the jump distribution has stretched exponential tails. Random walks and branching random walks with stretched exponential tails have been investigated in the literature [19, 27], but questions about the N-BRW with such a jump distribution, such as asymptotic speed, spatial distribution, and genealogy, remain open. In the future we intend to investigate the N-BRW in the stretched exponential case.

2.1.7 Organisation of the paper

In Section 2.2 we state Theorem 2.2.1 and Propositions 2.2.1 and 2.2.2, our main results, which we have explained in Sections 2.1.3 and 2.1.5. Furthermore, we give a heuristic argument for the proof of Theorem 2.2.1, introduce the notation we will be using throughout, and carry out the first step towards proving Theorem 2.2.1 in Lemma 2.2.4. As a result, the proof of Theorem 2.2.1 will be reduced to proving Propositions 2.2.5 and 2.2.6. We prove the former in Sections 2.3 and 2.4 and the latter in Section 2.5.

In Section 2.3 we give a deterministic argument for the existence of a common ancestor between times t_1 and t_2 of almost the whole population at time t. The argument will also imply that almost every particle in the population at time t is near the leftmost particle. Then in Section 2.4 we check that the events of the deterministic argument occur with high probability. A key step in the proof is to see that paths cannot move a distance of order a_N in ℓ_N time without making at least one jump of order a_N. We prove a large deviation result to show this, taking ideas from [21] and [27]. The other important tool, which we will use to estimate probabilities, is Potter's bound for regularly varying functions.

In Section 2.5 we prove that the genealogy is star-shaped. We will use concentration results from [35] to see that a single particle at time $T + \varepsilon_N \ell_N$ cannot have more than of order $N^{1-\varepsilon_N}$ surviving descendants at time t, which will be enough to conclude the result.

In Section 2.6 we prove Propositions 2.2.1 and 2.2.2 using some of our ideas from the deterministic argument in Section 2.3.

Section 2.7 is a glossary of notation, where we collect the notation most frequently used in this paper with a brief explanation, and with a reference to the section or equation where the notation is defined. In Section 2.7 we also list the most important intermediate steps of the proof of our main result.

We note here that sometimes we explain or justify an equation or inequality shortly *after* the statement appears; we encourage any reader who is struggling to understand a logical step to read a few lines ahead.

2.2 Genealogy and spatial distribution result

2.2.1 Formal definition of the N-BRW

Let $X_{i,b,n}$, $i \in [N]$, $b \in \{1,2\}$, $n \in \mathbb{N}_0$ be i.i.d. random variables with common law given by (2.1.3). Each $X_{i,b,n}$ stands for the jump size of the bth offspring of particle (i,n). Let $\mathcal{X}(0) = \{\mathcal{X}_1(0) \leq \ldots \leq \mathcal{X}_N(0)\}$ be any ordered set of N real numbers, which represents the initial locations of the N particles. Now we describe inductively how $\mathcal{X}(0)$ and the random variables $X_{i,b,n}$, $i \in [N]$, $b \in \{1,2\}$, $n \in \mathbb{N}_0$ determine the N-BRW, that is, the sequence of locations of the N particles, $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$.

We start with the initial configuration of particles $\mathcal{X}(0)$. Once $\mathcal{X}(n)$ has been determined for some $n \in \mathbb{N}_0$, then $\mathcal{X}(n+1)$ is defined as follows. Each particle has two offspring, each of which performs a jump from the location of its parent. The $2N$ independent jumps at time n are then given by the i.i.d. random variables $X_{i,b,n}$, $i \in [N]$, $b \in \{1,2\}$ as above. After the jumps, only the N rightmost offspring particles survive; that is, $\mathcal{X}(n+1) = \{\mathcal{X}_1(n+1) \leq \cdots \leq \mathcal{X}_N(n+1)\}$ is given by the N largest numbers from the collection $(\mathcal{X}_i(n) + X_{i,b,n})_{i\in[N],b\in\{1,2\}}$. Ties are decided arbitrarily.

Note that since the jumps are non-negative, the sequences $\mathcal{X}_i(n)$ are non-decreasing in n for all $i \in [N]$. Indeed, at time n there are at least $N-i+1$ particles to the right of or at position $\mathcal{X}_i(n)$, and so there are at least $\min(N, 2(N-i+1))$ particles to the right of or at $\mathcal{X}_i(n)$ at time $n+1$, so we must have $\mathcal{X}_i(n+1) \geq \mathcal{X}_i(n)$. We refer to this property as *monotonicity* throughout.

2.2.2 Statement of our main result

We explained the message of our main result in Section 2.1.3. In this section we provide the precise statement in Theorem 2.2.1. First we introduce the setup for the theorem.

For $n, k \in \mathbb{N}_0$ and $i \in [N]$ we will denote the index of the time-n ancestor of the particle $(i, n+k)$ by

$$\zeta_{i,n+k}(n),$$

i.e. particle $(\zeta_{i,n+k}(n), n)$ is the ancestor of $(i, n+k)$. Recall that the relevant space scale for our process is a_N, defined in (2.1.5). For $r \geq 0$ and $n \in \mathbb{N}_0$, let $L_{r,N}(n)$ denote the number of particles which are within distance ra_N of the leftmost particle at time n:

$$L_{r,N}(n) := \max\{i \in [N] : \mathcal{X}_i(n) \leq \mathcal{X}_1(n) + ra_N\}. \tag{2.2.1}$$

Define a sequence $(\varepsilon_N)_{N\in\mathbb{N}}$ such that $\varepsilon_N \ell_N$ is an integer for all $N \geq 1$, and which satisfies

$$\varepsilon_N \ell_N \to \infty \text{ and } \varepsilon_N \to 0 \text{ as } N \to \infty. \tag{2.2.2}$$

We introduce two events which describe the spatial distribution and the genealogy of the population at a given time t. Our main result, Theorem 2.2.1, says that these two events occur with high probability. We define the events for all $N \geq 2$ and $t > 4\ell_N$. For $\eta > 0$ and $\gamma \in (0,1)$, the first event says that at least $N - N^{1-\gamma}$ particles (i.e. almost the whole population if N is large) are within distance ηa_N of the leftmost particle at time t. We let

$$\mathcal{A}_1 = \mathcal{A}_1(t, N, \eta, \gamma) := \left\{L_{\eta,N}(t) \geq N - N^{1-\gamma}\right\}. \tag{2.2.3}$$

Recall the notation t_i from (2.1.7). We illustrate the second event in Figure 2-2. We sample $M \in \mathbb{N}$ particles uniformly at random from the population at time t. Let $\mathcal{P} = (\mathcal{P}_1, \ldots, \mathcal{P}_M)$ be the index set of the sampled particles. The event says that there exists a time T between t_2 and t_1 such that all of the particles in the sample have a common ancestor at time T, but no pair of particles in the sample have a common ancestor at time $T + \varepsilon_N \ell_N$. Moreover, the common ancestor at time T is the leader particle (N, T). Additionally, the event says that the time T is not particularly close to t_1 or t_2, in that $T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]$ for some $\delta > 0$. We let

$$\mathcal{A}_2 = \mathcal{A}_2(t, N, M, \delta) := \left\{ \begin{array}{l} \exists T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil] : \zeta_{\mathcal{P}_i,t}(T) = N \ \forall i \in [M] \text{ and} \\ \zeta_{\mathcal{P}_i,t}(T + \varepsilon_N \ell_N) \neq \zeta_{\mathcal{P}_j,t}(T + \varepsilon_N \ell_N) \, \forall i,j \in [M], \ i \neq j \end{array} \right\}. \tag{2.2.4}$$

For convenience, we will often write $\mathcal{A}_1$ and $\mathcal{A}_2$ for the two events above, omitting the arguments. We will prove the following result.

Theorem 2.2.1. *For all $\eta > 0$ and $M \in \mathbb{N}$ there exist $\gamma, \delta \in (0,1)$ such that for all $N \in \mathbb{N}$*

sufficiently large and $t \in \mathbb{N}_0$ *with* $t > 4\ell_N$,

$$\mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2) > 1 - \eta,$$

where ℓ_N *is given by* (2.1.4)*, and* $\mathcal{A}_1 = \mathcal{A}_1(t, N, \eta, \gamma)$ *and* $\mathcal{A}_2 = \mathcal{A}_2(t, N, M, \delta)$ *are defined in* (2.2.3) *and* (2.2.4) *respectively.*

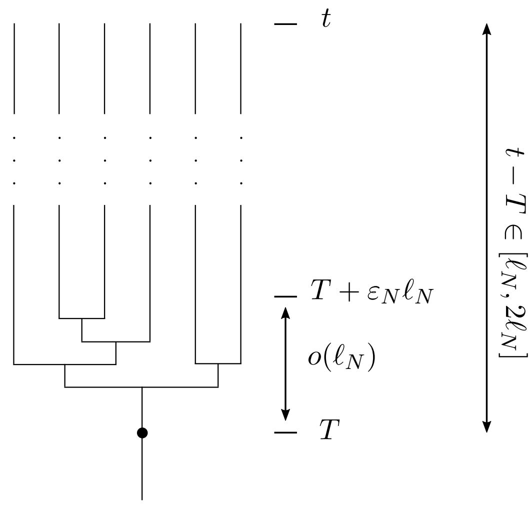

Figure 2-2: Coalescence of the ancestral lineages of $M = 6$ particles. We go backwards in time from top to bottom in the figure. To each particle in the sample we associate a vertical line, representing its ancestral line. Two lines coalesce into one when the particles they are associated with have a common ancestor for the first time going backwards from time t. All coalescences of the lineages of the sample happen within a time window of size $o(\ell_N)$. Time T is the generation of the most recent common ancestor of the majority of the whole population at time t. The three dots in each line indicate that the picture is not proportional: the time between t and T is of order ℓ_N, whereas the time between all coalescences and T is $o(\ell_N)$.

We explained two additional results in Section 2.1.5 which show the optimality of Theorem 2.2.1. We state these results precisely below.

We define the event $\mathcal{A}_2'$ as a modification of the event $\mathcal{A}_2$. Whereas $\mathcal{A}_2$ said that the coalescence time T is roughly in $[t_2, t_1]$, the event $\mathcal{A}_2'$ says that T is in the smaller interval $[t_2 + \lceil s_1\ell_N \rceil, t_2 + \lceil s_2\ell_N \rceil]$ for $0 < s_1 < s_2 < 1$; and whereas $\mathcal{A}_2$ occurs with high probability, we will show that $\mathcal{A}_2'$ occurs with probability bounded away from 0. For $M \in \mathbb{N}$ and $0 < s_1 < s_2 < 1$, we define

$$\begin{aligned} \mathcal{A}_2' &= \mathcal{A}_2'(t, N, M, s_1, s_2) \\ &:= \left\{ \begin{array}{l} \exists T \in [t_2 + \lceil s_1\ell_N \rceil, t_2 + \lceil s_2\ell_N \rceil] : \zeta_{\mathcal{P}_i,t}(T) = N \;\; \forall i \in [M] \text{ and} \\ \zeta_{\mathcal{P}_i,t}(T + \varepsilon_N\ell_N) \neq \zeta_{\mathcal{P}_j,t}(T + \varepsilon_N\ell_N) \;\; \forall i, j \in [M],\; i \neq j \end{array} \right\}. \end{aligned} \tag{2.2.5}$$

Proposition 2.2.1 below says that for all $0 < s_1 < s_2 < 1$ and $r > 0$, with probability bounded below by a constant depending on r and $s_2 - s_1$, the event $\mathcal{A}_2'$ occurs and the diameter at time t_1 is at least ra_N. The diameter of the particle cloud at time n will be denoted by $d(\mathcal{X}(n))$; that is,

$$d(\mathcal{X}(n)) := \mathcal{X}_N(n) - \mathcal{X}_1(n). \tag{2.2.6}$$

Proposition 2.2.1. *For all $0 < s_1 < s_2 < 1$, $M \in \mathbb{N}$ and $r > 0$, there exists $\pi_{r,s_2-s_1} > 0$ such that for N sufficiently large and $t > 4\ell_N$,*

$$\mathbb{P}\left(\mathcal{A}_2' \cap \{d(\mathcal{X}(t_1)) \geq ra_N\}\right) > \pi_{r,s_2-s_1},$$

where $\mathcal{A}_2'(t, N, M, s_1, s_2)$ is defined in (2.2.5).

Our second result about the diameter says that for all r, the probability that $d(\mathcal{X}(n)) \geq ra_N$ is bounded away from zero, and it tends to 1 as $r \to 0$, and tends to 0 as $r \to \infty$, if N is sufficiently large and $n > 3\ell_N$. This shows that the probability that after a long time the diameter is not of order a_N is small, and therefore the part of Theorem 2.2.1 that says most of the population is within distance ηa_N of the leftmost particle with high probability, for arbitrarily small $\eta > 0$, is meaningful.

Proposition 2.2.2. *There exist $0 < p_r \leq q_r \leq 1$ such that $q_r \to 0$ as $r \to \infty$ and $p_r \to 1$ as $r \to 0$, and for all $r > 0$,*

$$0 < p_r \leq \mathbb{P}(d(\mathcal{X}(n)) \geq ra_N) \leq q_r,$$

for N sufficiently large and $n > 3\ell_N$.

2.2.3 Heuristics for the proof of Theorem 2.2.1

We first prove a simple lemma which will be helpful in the course of the proof of Theorem 2.2.1 and also helpful for understanding the heuristics. The lemma says that the number of particles that are to the right of a given position at least doubles at every time step until it reaches N. The statement follows deterministically from the definition of the N-BRW. The proof serves as a warm-up for several more deterministic arguments to come. For $x \in \mathbb{R}$ and $n \in \mathbb{N}_0$, we write the set of particles to the right of position x at time n as

$$G_x(n) := \{i \in [N] : \mathcal{X}_i(n) \geq x\}. \tag{2.2.7}$$

Lemma 2.2.3. *Let $x \in \mathbb{R}$ and $n, k \in \mathbb{N}_0$. Then*

$$|G_x(n+k)| \geq \min\left(N,\ 2^k|G_x(n)|\right).$$

Proof. The statement is clearly true when $G_x(n) = \emptyset$. Now assume that $G_x(n) \neq \emptyset$. Let us first consider the case in which every descendant of the particles in $G_x(n)$ survives until time $n+k$. Since there are $2^k|G_x(n)|$ such descendants, each of which is to the right of x since all jumps are non-negative, in this case we have $|G_x(n+k)| \geq 2^k|G_x(n)|$.

Now let us consider the case in which not every descendant of the particles in $G_x(n)$ survives until time $n+k$. This means that there exist $m \in [n, n+k-1]$, $j \in [N]$ and $b \in \{1,2\}$ such that (j,m) is a descendant of a particle in $G_x(n)$ and

$$\mathcal{X}_j(m) + X_{j,b,m} \leq \mathcal{X}_1(m+1).$$

Since particle (j,m) descends from $G_x(n)$, and all jumps are non-negative, we also have $x \leq \mathcal{X}_j(m) + X_{j,b,m}$, and therefore $x \leq \mathcal{X}_1(m+1) \leq \mathcal{X}_1(n+k)$, and the result follows. □

Now we turn to the heuristics for the proof of Theorem 2.2.1. The heuristic picture to keep in mind when thinking about both the statement and the proof is Figure 2-1. As in Section 2.1.4, we let T denote the last time at which a particle takes the lead with a big jump of order a_N before time t_1. In Section 2.1.4, we argued that if $T \in [t_2, t_1]$ then with high probability, particle (N,T) will be the common ancestor of almost every particle in the population at time t, and almost the whole population at time t is close to $\mathcal{X}_N(T)$ on the a_N space scale. We will use a rigorous version of this heuristic argument to show that the event $\mathcal{A}_1$ occurs with high probability, and that the time T satisfies the first line in the event $\mathcal{A}_2$ with high probability. That is, every particle from a uniform sample of fixed size M at time t descends from particle (N,T) with high probability.

If T is as described above, then we can only have $T \in [t_2, t_1]$ if there is a particle which takes the lead with a jump of order a_N in the time interval $[t_2, t_1]$. It is not straightforward to show that this happens with high probability. It could be the case that the diameter is large on the a_N space scale during the time interval $[t_2, t_1]$, say greater than Ca_N, where $C > 0$ is large. In this situation, if the jumps of order a_N in the time interval $[t_2, t_1]$ come from close to the leftmost particle, and they are all smaller than Ca_N, then these jumps will not make a new leader, and time T will not be in the time interval $[t_2, t_1]$. We will prove that this is unlikely. A key property which is helpful in seeing this is the following. If no particle takes the lead with a big jump of order a_N for ℓ_N time, e.g. between times $s \in \mathbb{N}$ and $s + \ell_N$, then the diameter of the particle cloud will be very small on the a_N space scale at time $s + \ell_N$. Indeed, all the N particles, including the leftmost, are to the right of position $\mathcal{X}_N(s)$ at time $s+\ell_N$ by Lemma 2.2.3. But with high probability, particles cannot move far to the right from this position without making big jumps of order a_N. We will prove this in Corollary 2.4.5. Therefore, provided that no unlikely event happens, if no particle takes the lead with a big jump between times s and $s+\ell_N$, then every particle will be near the position $\mathcal{X}_N(s)$ at time $s+\ell_N$. We formally prove this in Lemma 2.3.9.

We will be able to use this property for $s = t_2 - c'\ell_N$ with small $c' > 0$. We will conclude that if no particle takes the lead with a jump of order a_N in the time interval $[t_2 - c'\ell_N, t_1 - c'\ell_N]$ then the diameter at time $t_1 - c'\ell_N$ is likely to be small on the a_N space scale, i.e. $d(\mathcal{X}(t_1 - c'\ell_N)) < ca_N$, for some $c > 0$ which we can choose to be much smaller than c'. If the diameter is less than ca_N, then any particle performing a jump larger than ca_N becomes the new leader.

The expected number of jumps larger than ca_N in $c'\ell_N$ time is $c'\ell_N 2Nh(ca_N)^{-1}$, because there are $2N$ jumps at each time step and the jump distribution is given by (2.1.3), which is roughly c'/c^α for N sufficiently large. If c^α is much smaller than c', then with high probability there will be a jump of size greater than ca_N in the time interval $[t_1 - c'\ell_N, t_1]$, and the particle performing it will become the new leader. Therefore the last time before time t_1 when a particle becomes the leader with a jump of order a_N will be after time t_2, which gives us $T \in [t_2, t_1]$.

The above idea works for the case where no particle takes the lead with a big jump of order a_N in the time interval $[t_2 - c'\ell_N, t_2]$ for some small $c' > 0$. If instead there is such a particle then we will argue that in a short interval of length $c'\ell_N$ it is likely that the jump made by this particle will not be too large on the a_N scale and therefore the particle's descendants will be surpassed by larger jumps of order a_N at some point in the much longer time interval $[t_2, t_1]$.

In order to show that the coalescence is star shaped, we also need the second line of the event $\mathcal{A}_2$, which says that all coalescences of the lineages of a sample of M particles at time t happen within a time window of size $\varepsilon_N \ell_N$; that is, instantaneously on the ℓ_N time scale (see Figure 2-2).

To prove that no pair of particles in the sample of M have a common ancestor at time $T + \varepsilon_N \ell_N$, it will be enough to prove that every particle at time $T + \varepsilon_N \ell_N$ has a number of time-t descendants which is at most a very small proportion of the total population size N (we will check this in Lemma 2.2.4). With high probability, most of the population at time t descends from the leading $2^{\varepsilon_N \ell_N} \approx N^{\varepsilon_N}$ particles at time $T + \varepsilon_N \ell_N$ (the descendants of particle (N, T)). If these particles share their time-t descendants fairly evenly, then a particle in this leading tribe will have roughly $N^{1-\varepsilon_N} = o(N)$ descendants. Indeed, we will prove using concentration results from [35] that with high probability the number of time-t descendants of a particle from the leading tribe at time $T + \varepsilon_N \ell_N$ will not exceed the order of $N^{1-\varepsilon_N}$.

2.2.4 Notation

We now introduce the notation we will be using throughout the proof of Theorem 2.2.1. We recall from Section 2.2.1 that the jump of the ith particle's bth offspring at time n will be referred to using the random variable $X_{i,b,n}$, and that these contain all the randomness in the system, with ties between two particles with the same position broken using some arbitrary but deterministic rule. We define the filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ by letting $\mathcal{F}_n$ be the σ-algebra generated by the random variables $(X_{i,b,m},\ i \in [N], b \in \{1,2\}, m < n)$. Since $\mathcal{X}(n)$ is defined in such a way that it only depends on jumps performed before time n, the process $(\mathcal{X}(n))_{n \in \mathbb{N}_0}$ is adapted to the filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Since $(X_{i,b,m}, i \in [N], b \in \{1,2\}, m \in \mathbb{N}_0)$ are i.i.d., the jumps $(X_{i,b,n}, i \in [N], b \in \{1,2\})$ are independent of the σ-algebra $\mathcal{F}_n$. In

Theorem 2.2.1 we assume that $t > 4\ell_N$, as in the proof we will examine the process in the time interval $[t_4, t]$, where t_4 is given by (2.1.7). Since jumps at time t are not $\mathcal{F}_t$-measurable, we will be interested in jumps performed in the time interval $[t_4, t-1]$.

In order to study the genealogy of the N-BRW system, we will need notation which says when two particles are related. We introduce the partial order $\lesssim$ on the set of pairs $\{(i,n), i \in [N], n \in \mathbb{N}_0\}$. First, for $i \in [N]$ and $n \in \mathbb{N}_0$ we say that $(i,n) \lesssim (i,n)$ and, for $j \in [N]$, we write $(i,n) \lesssim (j, n+1)$ if and only if the jth particle at time $n+1$ is an offspring of the ith particle at time n. Then in general, for $n, k \in \mathbb{N}_0$ and $i_0, i_k \in [N]$ we write $(i_0, n) \lesssim (i_k, n+k)$ if and only if particle $(i_k, n+k)$ is a descendant of particle (i_0, n):

$$(i_0,n) \lesssim (i_k, n+k) \iff \exists i_1, \dots, i_{k-1} : (i_{j-1}, n+j-1) \lesssim (i_j, n+j), \quad \forall j \in [k]. \tag{2.2.8}$$

Then the particles $((i_j, n+j),\ j \in [k])$ represent the ancestral line between (i_0, n) and $(i_k, n+k)$. Recall that for $n, k \in \mathbb{N}_0$ and $i \in [N]$ we denote the index of the time-n ancestor of the particle $(i, n+k)$ by $\zeta_{i,n+k}(n)$. Thus, using our partial order above, we can write for $j \in [N]$,

$$\zeta_{i,n+k}(n) = j \iff (j,n) \lesssim (i, n+k). \tag{2.2.9}$$

We also introduce a slightly different (strict) partial order $\lesssim_b$, which will be convenient later on. For $i_0, i_k \in [N]$, $n \in \mathbb{N}_0$ and $k \in \mathbb{N}$ we write $(i_0, n) \lesssim_b (i_k, n+k)$ if and only if the bth offspring of particle (i_0, n) is the time-$(n+1)$ ancestor of particle $(i_k, n+k)$. Note that if $(i_0, n) \lesssim_b (i_k, n+k)$ then there exists $i_1 \in [N]$ such that

$$\mathcal{X}_{i_1}(n+1) = \mathcal{X}_{i_0}(n) + X_{i_0,b,n} \text{ and } (i_1, n+1) \lesssim (i_k, n+k).$$

Using the above partial order, we define the *path* between particles (i_0, n) and $(i_k, n+k)$ (and between positions $\mathcal{X}_{i_0}(n)$ and $\mathcal{X}_{i_k}(n+k)$), as the sequence of jumps connecting the two particles. For $i_0, i_k \in [N]$ and $n \in \mathbb{N}_0$, if $k \in \mathbb{N}$ and $(i_0, n) \lesssim (i_k, n+k)$, we let

$$P^{i_k, n+k}_{i_0, n} := \left\{ (i_j, b_j, n+j) :\ j \in \{0, \dots, k-1\} \text{ and } (i_j, n+j) \lesssim_{b_j} (i_k, n+k) \right\}, \tag{2.2.10}$$

and we let $P^{i_k,n+k}_{i_0,n} := \emptyset$ otherwise. Then if $k \in \mathbb{N}$ and $(i_0, n) \lesssim (i_k, n+k)$,

$$\mathcal{X}_{i_k}(n+k) = \mathcal{X}_{i_0}(n) + \sum_{(j,b,m) \in P^{i_k,n+k}_{i_0,n}} X_{j,b,m}. \tag{2.2.11}$$

For $i \in [N]$ and $n, k \in \mathbb{N}_0$ with $n \le k$, let $\mathcal{N}_{i,n}(k)$ denote the set of descendants of particle (i,n) at time k:

$$\mathcal{N}_{i,n}(k) := \{ j \in [N] :\ (i,n) \lesssim (j,k) \}, \tag{2.2.12}$$

and if $n < k$, for $b \in \{1,2\}$, let $\mathcal{N}^b_{i,n}(k)$ be the set of time-k descendants of the bth offspring of particle (i,n):

$$\mathcal{N}^b_{i,n}(k) := \{j \in [N] : (i,n) \lesssim_b (j,k)\}. \tag{2.2.13}$$

(Note that the sets $\mathcal{N}_{i,n}(k)$ and $\mathcal{N}^b_{i,n}(k)$ may be empty.) We write $|\mathcal{N}_{i,n}(k)|$ and $|\mathcal{N}^b_{i,n}(k)|$ for the number of descendants in each case.

Finally, as time is discrete, it will be useful to introduce a notation for the set of integers in an interval; for $0 \le s_1 \le s_2$, we let

$$[\![s_1, s_2]\!] := [s_1, s_2] \cap \mathbb{N}_0.$$

2.2.5 Big jumps and breaking the record

As discussed in Section 2.1.4, the common ancestor of the majority of the population at time t is a particle which made an unusually big jump, of order a_N, between times t_2 and t_1. The set of unusually big jumps will play an essential role in the proof of Theorem 2.2.1. We will be particularly interested in particles which become 'leaders' after performing such jumps. These particles are the candidates to become the common ancestor of almost the whole population at time t.

We now introduce the necessary notation for the above concepts. In the definitions we will indicate the dependence on a new parameter $\rho \in (0,1)$, as the choice of ρ will be important later on. Furthermore, everything we define will depend on N and t, which we do not always indicate.

For $\rho \in (0,1)$ we introduce the term *big jump* for jumps of size greater than ρa_N, and we denote the set of big jumps on an interval $[s_1, s_2] \subseteq [t_4, t-1]$ by $B_N^{[s_1,s_2]}$:

$$B_N^{[s_1,s_2]} = B_N^{[s_1,s_2]}(\rho) := \{(k,b,s) \in [N] \times \{1,2\} \times [\![s_1,s_2]\!] : X_{k,b,s} > \rho a_N\}, \tag{2.2.14}$$

where a_N is given by (2.1.5). We also let

$$B_N := B_N^{[t_4,t-1]}. \tag{2.2.15}$$

We say a particle *breaks the record* if it takes the lead with a big jump. If one of the current leader's descendants makes a small jump (that is, a non-big jump) to become the leader, then that does not count as breaking the record in our terminology. Let $\mathbf{S}_N$ denote the set of times when the record is broken by a big jump between times t_4 and t:

$$\mathbf{S}_N = \mathbf{S}_N(\rho) := \left\{ \begin{array}{l} s \in [\![t_4, t-1]\!] : \exists (k,b) \in [N] \times \{1,2\} \text{ such that} \\ (k,s) \lesssim_b (N, s+1) \text{ and } X_{k,b,s} > \rho a_N \end{array} \right\}. \tag{2.2.16}$$

Next, we define T as the last time when the leader broke the record with a big jump before time t_1, if there is any such time. We let

$$T = T(\rho) := 1 + \max\{\mathbf{S}_N(\rho) \cap [t_4, t_1 - 1]\}, \tag{2.2.17}$$

and let $T = 0$ if $\mathbf{S}_N(\rho) \cap [t_4, t_1 - 1] = \emptyset$. Note that the big jump which takes the lead happens at time $T-1$, and T is the time right after the jump. In the proof it turns out that with high probability, $T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]$ for some $\delta > 0$, and particle (N, T) is the common ancestor of almost the whole population at time t.

We will have a separate notation, $\hat{\mathbf{S}}_N$, for the times when the leader is surpassed by a particle which performs a big jump. Note that this is not exactly the same set of times as $\mathbf{S}_N$: it might happen that a particle (i, s) has an offspring $(j, s+1)$, which beats the current leader (N, s) with a big jump, but it does not become the next leader at time $s+1$ because it is beaten by another offspring particle which did not make a big jump. We define

$$\hat{\mathbf{S}}_N = \hat{\mathbf{S}}_N(\rho) := \left\{ \begin{array}{l} s \in [\![t_4, t-1]\!] : \exists (k,b) \in [N] \times \{1,2\} \text{ such that} \\ X_{k,b,s} > \rho a_N \text{ and } \mathcal{X}_k(s) + X_{k,b,s} > \mathcal{X}_N(s) \end{array} \right\}. \tag{2.2.18}$$

We will see in Corollary 2.3.8 below that with high probability, $\mathbf{S}_N$ and $\hat{\mathbf{S}}_N$ coincide on certain time intervals. Sometimes we will also need to refer to the set of times when big jumps do not take the lead or beat the current leader. Therefore, with a slight abuse of notation, we will write $\mathbf{S}_N^c$ and $\hat{\mathbf{S}}_N^c$ to denote the sets of times $[\![t_4, t-1]\!] \setminus \mathbf{S}_N$ and $[\![t_4, t-1]\!] \setminus \hat{\mathbf{S}}_N$ respectively.

2.2.6 Reformulation

In this section, we break down the event $\mathcal{A}_2$ of Theorem 2.2.1. Our ultimate goal is to show, for a suitable choice of ρ, that $T = T(\rho)$, as defined in (2.2.17), has the properties required in $\mathcal{A}_2$. To this end we introduce new events which imply $\mathcal{A}_2$ with high probability, and only involve T and the number of time-t descendants of particle (N, T) and of the particles at time $T + \varepsilon_N \ell_N$. We will use the following notation:

$$T^{\varepsilon_N} = T^{\varepsilon_N}(\rho) := T(\rho) + \varepsilon_N \ell_N, \tag{2.2.19}$$

where ε_N is defined in (2.2.2). Recalling (2.2.12), for $i \in [N]$, we write

$$\mathcal{N}_i := \mathcal{N}_{i, T^{\varepsilon_N}}(t) \tag{2.2.20}$$

for the set of time-t descendants of the ith particle at time T^{ε_N}, and

$$D_i = D_{i, T^{\varepsilon_N}}(t) := |\mathcal{N}_{i, T^{\varepsilon_N}}(t)| \tag{2.2.21}$$

for the size of this set.

For $\gamma, \delta, \rho \in (0, 1)$, we introduce the event

$$\mathcal{A}_3 = \mathcal{A}_3(t, N, \delta, \rho, \gamma) := \{T(\rho) \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]\} \cap \left\{ |\mathcal{N}_{N, T(\rho)}(t)| \geq N - N^{1-\gamma} \right\}. \tag{2.2.22}$$

This event says that almost the whole population at time t descends from particle (N, T), which will imply with high probability that each particle in the uniform sample of M particles in the event $\mathcal{A}_2$ is a descendant of (N,T). The final part of the definition of the event $\mathcal{A}_2$ says that no two particles at time t in the uniform sample of M particles share an ancestor at time T^{ε_N}. We now define an event which says that every time-T^{ε_N} particle has at most a very small proportion of the N surviving descendants at time t, so that with high probability none of them have two descendants in the sample of M particles. For $\nu > 0$ and $\rho \in (0,1)$, we let

$$\mathcal{A}_4(\nu) = \mathcal{A}_4(t,N,\rho,\nu) := \left\{ \max_{i \in \mathcal{N}_{N,T}(T^{\varepsilon_N})} D_{i,T^{\varepsilon_N}}(t) \leq \nu N \right\}. \tag{2.2.23}$$

Note that in the definition of $\mathcal{A}_4(\nu)$ we take the maximum only over the time-T^{ε_N} descendants of particle (N,T). It will be easy to deal with the remaining particles at time T^{ε_N}, because the event $\mathcal{A}_3$ implies that for $\nu > 0$, if N is large, particles not descended from (N,T) cannot have more than νN descendants at time t. In the following result, we reduce the proof of Theorem 2.2.1 to showing that $\mathcal{A}_1$, $\mathcal{A}_3$ and $\mathcal{A}_4(\nu)$ occur with high probability.

As part of the proof we show that the probability that two particles in the sample of M at time t have a common ancestor at time T^{ε_N} can be upper bounded by little more than the sum of the probabilities of the events $\mathcal{A}_3^c$ and $\mathcal{A}_4(\nu)^c$ when ν is small. We will use this intermediate result in another argument later on in Section 2.6, so we state it as part of Lemma 2.2.4 below.

Lemma 2.2.4. *Take $M \in \mathbb{N}$ and $\gamma, \delta, \rho, \eta \in (0,1)$, and let $0 < \nu < \eta/M^2$. Then for all N sufficiently large and $t > 4\ell_N$,*

$$\mathbb{P}(\exists j,l \in [M],\ j \neq l :\ \zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N}) = \zeta_{\mathcal{P}_l,t}(T^{\varepsilon_N})) \leq \mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta/2,$$

and

$$\mathbb{P}(\mathcal{A}_2^c) \leq 2\mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta,$$

where $\mathcal{A}_2(t,N,M,\delta)$, $\mathcal{A}_3(t,N,\delta,\rho,\gamma)$ and $\mathcal{A}_4(t,N,\rho,\nu)$ are defined in (2.2.4), (2.2.22) and (2.2.23) respectively, $\mathcal{P}_j$ is the index of a particle in the uniform sample of M particles at time t, and $\zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N})$ is the index of the time-T^{ε_N} ancestor of particle $(\mathcal{P}_j,t)$, defined in (2.2.9).

Proof. Fix $M \in \mathbb{N}$ and $\gamma, \delta, \rho, \eta \in (0,1)$. Note that by the definition of $\mathcal{A}_2$ in (2.2.4),

$$\begin{aligned}\{T \in [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]\} \cap \{\zeta_{\mathcal{P}_j,t}(T) = N\ \forall j \in [M]\} \\ \cap \left\{\zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N}) \neq \zeta_{\mathcal{P}_l,t}(T^{\varepsilon_N})\ \forall j,l \in [M],\ j \neq l\right\} \subseteq \mathcal{A}_2.\end{aligned} \tag{2.2.24}$$

First we aim to show that for N sufficiently large,

$$\mathbb{P}(\{T \notin [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]\} \cup \{\exists j \in [M] :\ \zeta_{\mathcal{P}_j,t}(T) \neq N\}) \leq \mathbb{P}(\mathcal{A}_3^c) + \eta/2. \tag{2.2.25}$$

Note that if $\mathcal{A}_3$ occurs then $T \in [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]$, and $\mathcal{A}_3$ is $\mathcal{F}_t$-measurable, so

$$\begin{aligned}\mathbb{P}(\{T \notin [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]\} \cup \{\exists j \in [M] :\, \zeta_{\mathcal{P}_j,t}(T) \neq N\}) \\ \leq \mathbb{E}\left[\mathbb{1}_{\mathcal{A}_3}\mathbb{P}(\exists j \in [M] :\, \zeta_{\mathcal{P}_j,t}(T) \neq N \mid \mathcal{F}_t)\right] + \mathbb{P}(\mathcal{A}_3^c).\end{aligned}$$

Now, on the event $\mathcal{A}_3$, at most $N^{1-\gamma}$ time-t particles are not descended from (N, T), and therefore a union bound on the uniformly chosen sample (which is not $\mathcal{F}_t$-measurable) gives that the above is at most $MN^{1-\gamma}/N + \mathbb{P}(\mathcal{A}_3^c)$. This implies (2.2.25) for N sufficiently large.

Now fix $\nu \in (0, \eta/M^2)$. Our second step is to prove that for N sufficiently large,

$$\mathbb{P}(\exists j, l \in [M],\ j \neq l :\, \zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N}) = \zeta_{\mathcal{P}_l,t}(T^{\varepsilon_N})) \leq \mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta/2, \qquad (2.2.26)$$

which is the first part of the statement of the lemma. The event on the left-hand side means that there is a particle at time T^{ε_N} which has at least two descendants in the sample of M particles at time t. That is

$$\begin{aligned}\mathbb{P}(\exists j, l \in [M],\ j \neq l :\, \zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N}) = \zeta_{\mathcal{P}_l,t}(T^{\varepsilon_N})) \\ = \mathbb{P}(\exists i \in [N],\ j, l \in [M],\ j \neq l :\, \{\mathcal{P}_j, \mathcal{P}_l\} \subseteq \mathcal{N}_i).\end{aligned} \qquad (2.2.27)$$

We will use that if all the $\mathcal{N}_i$ sets have size smaller than νN then it is unlikely that two particles of the uniformly chosen sample will fall in the same $\mathcal{N}_i$ set. Since D_i is $\mathcal{F}_t$-measurable for all i, a union bound gives

$$\begin{aligned}&\mathbb{P}(\exists i \in [N],\ j, l \in [M],\ j \neq l :\, \{\mathcal{P}_j, \mathcal{P}_l\} \subseteq \mathcal{N}_i) \\ &\leq \mathbb{E}\left[\mathbb{1}_{\{\max_{i\in[N]} D_i \leq \nu N\}} \sum_{i=1}^{N} \sum_{1\leq j<l\leq M} \mathbb{P}(\{\mathcal{P}_j, \mathcal{P}_l\} \subseteq \mathcal{N}_i \mid \mathcal{F}_t)\right] + \mathbb{P}\left(\max_{i\in[N]} D_i > \nu N\right).\end{aligned} \qquad (2.2.28)$$

Since the sample is chosen uniformly at random, the first term on the right-hand side is equal to

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{\{\max_{i\in[N]} D_i \leq \nu N\}} \sum_{i=1}^{N} \binom{M}{2} \frac{\binom{D_i}{2}}{\binom{N}{2}}\right] &\leq \mathbb{E}\left[\mathbb{1}_{\{\max_{i\in[N]} D_i \leq \nu N\}} \max_{j\in[N]} D_j \binom{M}{2} \frac{\sum_{i=1}^{N} D_i}{N(N-1)}\right] \\ &\leq \binom{M}{2} \frac{\nu N}{N-1},\end{aligned} \qquad (2.2.29)$$

where in the second inequality we exploit the indicator and use that $\sum_{i=1}^{N} D_i = N$. In order to deal with the second term on the right-hand side of (2.2.28), note that the maximum is taken over all particles at time T^{ε_N} (because of the definition of D_i in (2.2.21)). Suppose N is sufficiently large that $N^{1-\gamma} \leq \nu N$. Then if the event $\mathcal{A}_3$ occurs, particles not descended from particle (N,T) (i.e. particles not in $\mathcal{N}_{N,T}(T^{\varepsilon_N})$) have at most νN descendants at time t. Therefore, by the definition of $\mathcal{A}_4(\nu)$,

$$\mathbb{P}\left(\max_{i\in[N]} D_i > \nu N\right) \leq \mathbb{P}(\mathcal{A}_4(\nu)^c) + \mathbb{P}\left(\max_{i\in[N]\setminus\mathcal{N}_{N,T}(T^{\varepsilon_N})} D_i > \nu N\right) \leq \mathbb{P}(\mathcal{A}_4(\nu)^c) + \mathbb{P}(\mathcal{A}_3^c), \tag{2.2.30}$$

for N sufficiently large.

Putting (2.2.27)-(2.2.30) together, since we chose $\nu < \eta/M^2$ we have that (2.2.26) holds for N sufficiently large. By (2.2.24), (2.2.25) and (2.2.26), the result follows. □

We now state the two main intermediate results in the proof of Theorem 2.2.1, which say that, for well-chosen γ, δ, and ρ, the events $\mathcal{A}_1$, $\mathcal{A}_3$ and $\mathcal{A}_4(\nu)$ occur with high probability. In Sections 2.3 and 2.4 we give the proof of Proposition 2.2.5, and in Section 2.5 we prove Proposition 2.2.6.

Proposition 2.2.5. *For $\eta \in (0,1]$ there exist $0 < \gamma < \delta < \rho < \eta$ such that for N sufficiently large and $t > 4\ell_N$,*

$$\mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_3) > 1 - \eta,$$

where $\mathcal{A}_1(t,N,\eta,\gamma)$ and $\mathcal{A}_3(t,N,\delta,\rho,\gamma)$ are defined in (2.2.3) *and* (2.2.22) *respectively.*

Proposition 2.2.6. *Let $\eta \in (0,1]$ and $\nu > 0$. Then for $\rho \in (0,\eta)$ as in Proposition 2.2.5, for N sufficiently large and $t > 4\ell_N$,*

$$\mathbb{P}(\mathcal{A}_4(\nu)) > 1 - 2\eta,$$

where $\mathcal{A}_4(t,N,\rho,\nu)$ is defined in (2.2.23).

Proof of Theorem 2.2.1. Lemma 2.2.4, Proposition 2.2.5 and Proposition 2.2.6 immediately imply Theorem 2.2.1. □

2.2.7 Strategies for the proofs of Propositions 2.2.5 and 2.2.6

Our strategy for the proof of Proposition 2.2.5 is based on the picture in Figure 2-1. For $t > 4\ell_N$, we will show that the following happens between times t_2 and t with probability close to 1.

1. There will be particles which lead by a large distance at times in $[t_2, t_1]$. The last such particle will be at time $T \in [t_2 + \lceil\delta\ell_N\rceil, t_1 - \lceil\delta\ell_N\rceil]$ with position $\mathcal{X}_N(T)$.

2. The descendants of this particle are close together and far away from the the rest of the population at time t_1, forming a small (size $o(N)$) leader tribe.

3. At time t, the descendants of the small leader tribe from time t_1 form a big tribe of $N - o(N)$ particles, which descend from particle (N,T) and are close to the leftmost particle.

The first part of the proof is a deterministic argument given in Section 2.3, which shows that if 'all goes well' between times t_4 and t, then steps 1-2-3 above roughly describe what happens, which will imply that the events $\mathcal{A}_1$ and $\mathcal{A}_3$ in Proposition 2.2.5 occur. For the deterministic argument we will introduce a number of events, which will describe sufficient criteria for $\mathcal{A}_1$ and $\mathcal{A}_3$ to happen. Once we have shown that the intersection of these events is contained in $\mathcal{A}_1 \cap \mathcal{A}_3$, it is enough to prove that the probability of this intersection is close to 1. This part will be carried out in Section 2.4, and consists of checking that 'all goes well' with high probability.

We describe our strategy for showing Proposition 2.2.6 in detail in Section 2.5.1. The main idea is to give a lower bound on the position of the leftmost particle at time t with high probability, and then use concentration inequalities from [35] to bound the number of time-t descendants of each particle in $\mathcal{N}_{N,T}(T^{\varepsilon_N})$ which can reach that lower bound by time t. A key intermediate step will be to see that with high probability, particles can reach the lower bound only if they have an ancestor which made a jump larger than a certain size.

2.3 Deterministic argument for the proof of Proposition 2.2.5

In this section we provide the main component of the proof of Proposition 2.2.5. We follow the plan explained in the previous section; we define new events and show that they imply $\mathcal{A}_1$ and $\mathcal{A}_3$. In Section 2.4, we will prove that the new events occur with high probability. The events describe a strategy designed to make sure that the majority of the population at time t has a common ancestor at some time between t_2 and t_1; that is, to ensure that $\mathcal{A}_3$ occurs. The strategy will also show that most of the particles descended from particle (N, T) cannot move too far from position $\mathcal{X}_N(T)$ by time t. Thus it will be easy to see that these descendants are near the leftmost particle at time t, and so $\mathcal{A}_1$ must occur. So although the strategy is designed for the event $\mathcal{A}_3$, it will imply $\mathcal{A}_1$ too.

In the course of the proof we will use several constants. We first give a guideline, which shows how the constants should be thought of throughout the rest of the paper, then we describe the specific assumptions we need for the rest of this section. Recall that we fixed $\alpha > 0$ as in (2.1.3) and that we have $\eta \in (0, 1]$ from the statement of Proposition 2.2.5. The other constants can be thought of as

$$0 < \gamma < \delta \ll \rho \ll c_1 \ll c_2 \ll c_3 \ll c_4 \ll c_5 \ll c_6 \ll \eta < 1 \quad \text{and} \quad K \gg \rho^{-\alpha}. \tag{2.3.1}$$

As everything is constant in (2.3.1), we only use $\ll$ as an informal notation to say that the left-hand side is much smaller than the right-hand side.

More specifically, for the rest of this section we fix the constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, \eta$ and K, and assume that they satisfy

$$0 < \gamma < \delta < \rho, \tag{2.3.2}$$

$$10\rho < c_1, \tag{2.3.3}$$

$$10c_j < c_{j+1} < \eta < 1, \quad j = 1, \dots, 5, \tag{2.3.4}$$

$$K > \rho^{-\alpha}. \tag{2.3.5}$$

We will have additional conditions on these constants in Section 2.4, which will be consistent with the assumptions (2.3.2)-(2.3.5).

Every event we introduce below will depend on N, t (with $t > 4\ell_N$) and on some of the constants above. In the definitions we will not indicate this dependence explicitly. Note furthermore that in the statement of Proposition 2.2.5, taking N sufficiently large may depend on γ, δ, or ρ.

2.3.1 Breaking down event $\mathcal{A}_3$

We begin by breaking down the event $\mathcal{A}_3$ from Proposition 2.2.5 into two other events. Then we will define a strategy for showing that these two events occur. The first event describes the particle system at time t_1; it says that there is a small leader tribe of size less than $2N^{1-\delta}$, and every other particle is at least $c_2 a_N$ to the left of this tribe. Moreover, each particle in the leading tribe descends from the same particle, (N, T). The common ancestor (N, T) is the last particle which breaks the record with a big jump before time t_1 (see (2.2.17) and also Figure 2-1). We also require $T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]$, which is part of the event $\mathcal{A}_3$.

To keep track of the size of the leader tribe we introduce notation for the number of particles which are within distance εa_N of the leader at time n:

$$R_{\varepsilon,N}(n) := \max \{ i \in [N] : \mathcal{X}_{N-i+1}(n) \geq \mathcal{X}_N(n) - \varepsilon a_N \}, \text{ for } n \in \mathbb{N}_0 \text{ and } \varepsilon > 0. \tag{2.3.6}$$

Note that if $R_{\varepsilon,N}(t_1) < N$ then particle $(N - R_{\varepsilon,N}(t_1) + 1, t_1)$ is within distance εa_N of the leader, but particle $(N - R_{\varepsilon,N}(t_1), t_1)$ is not. In the event we introduce below, we set $\varepsilon = c_1$ and require the distance between these two particles to be at least $c_2 a_N$, showing that there is a gap between the leader tribe and the other particles. The event is defined as follows:

$$\mathcal{B}_1 := \left\{ \begin{array}{l} R_{c_1,N}(t_1) \leq \min \left\{ N-1, 2N^{1-\delta} \right\}, \\ \mathcal{X}_{N-R_{c_1,N}(t_1)}(t_1) \leq \mathcal{X}_{N-R_{c_1,N}(t_1)+1}(t_1) - c_2 a_N, \\ T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil] \text{ and } \mathcal{N}_{N,T}(t_1) = \{N - R_{c_1,N}(t_1) + 1, \dots, N\} \end{array} \right\}, \tag{2.3.7}$$

where $T = T(\rho)$ and $\mathcal{N}_{N,T}(t_1)$ are given by (2.2.17) and (2.2.12) respectively.

In the description of Figure 2-1 in Section 2.1.4, we explained that the descendants of particle (N, T) are likely to lead at time t_1. The event $\mathcal{B}_1$ requires more; it also says that the leading tribe leads by a large distance, which is important to ensure that no other tribes can interfere with our heuristic picture and will be useful in Section 2.3.2. The most involved part of the deterministic argument in the remainder of Section 2.3 is to break up the event $\mathcal{B}_1$ into other events which happen with probability close to 1.

We now define another event which says that particles which are not in the leading tribe at time t_1 have at most $N^{1-\gamma}$ (i.e. much less than N for N large) descendants in total at time t. This will imply that the leading tribe at time t_1 will dominate the population at time t. We let

$$\mathcal{B}_2 := \left\{ \sum_{j=1}^{N-R_{c_1,N}(t_1)} |\mathcal{N}_{j,t_1}(t)| \leq N^{1-\gamma} \right\}, \tag{2.3.8}$$

where $\mathcal{N}_{j,t_1}(t)$ is given by (2.2.12). The events which we will introduce to break down the event $\mathcal{B}_1$ will easily imply $\mathcal{B}_2$ as well. Before defining the new events we check that $\mathcal{B}_1$ and $\mathcal{B}_2$ indeed imply $\mathcal{A}_3$.

Lemma 2.3.1. *Let $\mathcal{A}_3$, $\mathcal{B}_1$ and $\mathcal{B}_2$ be the events given by* (2.2.22), (2.3.7) *and* (2.3.8) *respectively. Then for all $N \geq 2$ and $t > 4\ell_N$,*

$$\mathcal{B}_1 \cap \mathcal{B}_2 \subseteq \mathcal{A}_3.$$

Proof. On the event $\mathcal{B}_1$, the descendants of particle (N, T) are the $R_{c_1,N}(t_1)$ rightmost particles at time t_1. Thus $\mathcal{N}_{N,T}(t)$ is a disjoint union of the sets $\mathcal{N}_{j,t_1}(t)$ for $j \in [\![N - R_{c_1,N}(t_1) + 1, N]\!]$. We deduce that on the event $\mathcal{B}_1 \cap \mathcal{B}_2$,

$$|\mathcal{N}_{N,T}(t)| = \sum_{j=N-R_{c_1,N}(t_1)+1}^{N} |\mathcal{N}_{j,t_1}(t)| \geq N - N^{1-\gamma}.$$

Since $T \in [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]$ on the event $\mathcal{B}_1$, the result follows. □

2.3.2 Breaking down events $\mathcal{B}_1$ and $\mathcal{B}_2$

We now break down the events $\mathcal{B}_1$ and $\mathcal{B}_2$ into new events $\mathcal{C}_1$ to $\mathcal{C}_7$ whose probabilities will be easier to estimate. The majority of the work in this section consists of showing that the intersection of the new events implies $\mathcal{B}_1$. We can then quickly conclude that the intersection implies both $\mathcal{B}_2$ and $\mathcal{A}_1$. One of the new events will need to be further broken down in Section 2.3.3.

New events $\mathcal{C}_1$ to $\mathcal{C}_7$

Recall that $[\![s_1, s_2]\!]$ denotes the set of integers in the interval $[s_1, s_2]$ and that the constants γ, δ, ρ, $c_1, c_2, \ldots, c_6$, η and K satisfy (2.3.2)-(2.3.5). We first introduce τ_1 to denote the first time after t_2 when a gap of size $2c_3a_N$ appears between the leader and the second rightmost particle:

$$\tau_1 := \inf\left\{s \geq t_2 + 1 : \mathcal{X}_N(s) > \mathcal{X}_{N-1}(s) + 2c_3a_N\right\}. \tag{2.3.9}$$

The first new event we define says that such a gap appears by time t_1, that is

$$\mathcal{C}_1 := \{\tau_1 \in [\![t_2+1, t_1]\!]\}\,. \tag{2.3.10}$$

The next event $\mathcal{C}_2$ ensures that the current leading tribe keeps distance from the other tribes during the time interval $[\tau_1, t_1]$. This is important, since $\mathcal{B}_1$ requires a gap behind the leading tribe at time t_1. The event $\mathcal{C}_2$ says that if a particle is far away (at least $c_3 a_N$) from the leader, then it cannot jump to within distance $2c_2 a_N$ of the leader's position with a single big jump (recall from (2.3.1) that $c_2 \ll c_3$). That is, a particle far from the leader either stays at least $2c_2 a_N$ behind the leader, or it beats the leader by more than $2c_2 a_N$. Jumping close to the leader would require a large jump, of size greater than $c_3 a_N$, restricted to an interval of size $4c_2 a_N$, which is much smaller than the size of the jump. We will see in Section 2.4 that the probability that such a jump occurs between times t_3 and t_1 is small. Let $Z_i(s)$ denote the gap between the rightmost and the ith particle at time s:

$$Z_i(s) := \mathcal{X}_N(s) - \mathcal{X}_i(s), \quad \text{for } s \in \mathbb{N}_0 \text{ and } i \in [N]. \tag{2.3.11}$$

Now we can define our next event

$$\mathcal{C}_2 := \left\{ \begin{array}{l} \nexists (i,b,s) \in [N] \times \{1,2\} \times [\![t_3, t-1]\!] \text{ such that} \\ Z_i(s) \geq c_3 a_N \text{ and } X_{i,b,s} \in (Z_i(s) - 2c_2 a_N, Z_i(s) + 2c_2 a_N] \end{array} \right\}. \tag{2.3.12}$$

We need to introduce several more events to make sure that 'all goes well'; that is, particles which we do not expect to make big jumps indeed do not make big jumps, and smaller jumps do not make too much difference on the a_N space scale. The next event says that if a particle makes a big jump, then it will not have a descendant which makes another big jump within ℓ_N time:

$$\mathcal{C}_3 := \left\{ \begin{array}{l} B_N \cap P^{k_2,s_2}_{k_1,s_1} = \{(k_1,b_1,s_1)\} \\ \forall (k_1,b_1,s_1) \in B_N,\ \forall s_2 \in [\![s_1+1, \min\{s_1+\ell_N+1, t\}]\!],\ \forall k_2 \in \mathcal{N}^{b_1}_{k_1,s_1}(s_2) \end{array} \right\}, \tag{2.3.13}$$

where B_N, $P^{k_2,s_2}_{k_1,s_1}$ and $\mathcal{N}^{b_1}_{k_1,s_1}(s_2)$ are defined in (2.2.15), (2.2.10) and (2.2.13) respectively.

The next event says the following. Take any path between two particles in the time interval $[t_4, t]$. If we omit the big jumps from the path then it does not move more than distance $c_1 a_N$. In particular, if there are no big jumps at all then the path moves at most $c_1 a_N$. The event is given by

$$\mathcal{C}_4 := \left\{ \begin{array}{l} \sum_{(i,b,s)\in P^{k_2,s_2}_{k_1,s_1}} X_{i,b,s} \mathbb{1}_{\{X_{i,b,s} \leq \rho a_N\}} \leq c_1 a_N \\ \forall (k_1,s_1) \in [N] \times [\![t_4, t-1]\!],\ \forall s_2 \in [\![s_1+1, t]\!],\ \forall k_2 \in \mathcal{N}_{k_1,s_1}(s_2) \end{array} \right\}, \tag{2.3.14}$$

where $P^{k_2,s_2}_{k_1,s_1}$ and $\mathcal{N}_{k_1,s_1}(s_2)$ are defined in (2.2.10) and (2.2.12) respectively.

The last three events are simple. On $\mathcal{C}_5$, two big jumps cannot happen at the same time:

$$\mathcal{C}_5 := \{|B_N \cap \{(k,b,s) : (k,b) \in [N] \times \{1,2\}\}| \le 1 \, \forall s \in [\![t_4, t-1]\!]\}. \tag{2.3.15}$$

Then $\mathcal{C}_6$ excludes big jumps which happen either right after time t_2 or very close to time t_1:

$$\mathcal{C}_6 := \left\{ B_N^{[t_2, t_2 + \lceil \delta \ell_N \rceil]} \cup B_N^{[t_1 - \lceil \delta \ell_N \rceil, t_1 + \lceil \delta \ell_N \rceil]} = \emptyset \right\} \tag{2.3.16}$$

where $B_N^{[s_1,s_2]}$ is defined in (2.2.14). Finally, $\mathcal{C}_7$ gives a bound on the number of big jumps:

$$\mathcal{C}_7 := \{|B_N| \le K\}, \tag{2.3.17}$$

where we recall that we chose K to be a positive constant at the start of Section 2.3.

Now we can state the main result of this subsection. It says that on the events $\mathcal{C}_1$ to $\mathcal{C}_7$ the events $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{A}_1$ occur, and therefore $\mathcal{A}_3$ occurs as well. We have an additional event in Proposition 2.3.2 below, which says that the diameter of the particle cloud at time t_1 is larger than $\frac{3}{2}c_3 a_N$. As part of the proposition we also show that $\mathcal{C}_1$ to $\mathcal{C}_7$ imply this event, because it will be useful in another argument later on in Section 2.6.

Proposition 2.3.2. *Let $\eta \in (0,1]$, and assume that the constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ satisfy* (2.3.2)-(2.3.5). *Then for N sufficiently large that $2KN^{-\delta} < N^{-\gamma} < 1$ and $t > 4\ell_N$,*

$$\bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \mathcal{B}_1 \cap \mathcal{B}_2 \cap \mathcal{A}_1 \cap \left\{ d(\mathcal{X}(t_1)) \ge \tfrac{3}{2}c_3 a_N \right\} \subseteq \mathcal{A}_1 \cap \mathcal{A}_3 \cap \left\{ d(\mathcal{X}(t_1)) \ge \tfrac{3}{2}c_3 a_N \right\},$$

where $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{A}_1$ and $\mathcal{A}_3$ are defined in (2.3.7), (2.3.8), (2.2.3) *and* (2.2.22) *respectively, and $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_7$ are given by* (2.3.10) *and* (2.3.12)–(2.3.17).

Note that the second inclusion in Proposition 2.3.2 follows directly from Lemma 2.3.1.

$\mathcal{C}_1$ to $\mathcal{C}_7$ imply $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{A}_1$: proof of Proposition 2.3.2

We start by proving some easy lemmas which hold on the event $\bigcap_{j=1}^{7} \mathcal{C}_j$, and which will be applied in the course of the proof of Proposition 2.3.2.

The first lemma gives another way of writing the event $\mathcal{C}_4$, which will be more convenient to use in this section. (The definition of $\mathcal{C}_4$ will be easier to work with when we show, in Section 2.4, that $\mathcal{C}_4$ occurs with high probability.) The lemma says that on the event $\mathcal{C}_4$, if a path moves more than $c_1 a_N$ then it must contain a big jump.

Lemma 2.3.3. *On the event $\mathcal{C}_4$, for all $(k_1, s_1) \in [N] \times [\![t_4, t-1]\!]$, $s_2 \in [\![s_1+1, t]\!]$ and $k_2 \in \mathcal{N}_{k_1,s_1}(s_2)$,*

$$\mathcal{X}_{k_2}(s_2) > \mathcal{X}_{k_1}(s_1) + c_1 a_N \Longrightarrow B_N \cap P_{k_1,s_1}^{k_2,s_2} \neq \emptyset,$$

where B_N, $\mathcal{N}_{k_1,s_1}(s_2)$ and $P_{k_1,s_1}^{k_2,s_2}$ are defined in (2.2.15), (2.2.12) *and* (2.2.10) *respectively.*

Proof. Let $(k_1, s_1) \in [N] \times [\![t_4, t-1]\!]$, $s_2 \in [\![s_1+1, t]\!]$, and $k_2 \in \mathcal{N}_{k_1,s_1}(s_2)$. Assume that $B_N \cap P^{k_2,s_2}_{k_1,s_1} = \emptyset$, and the event $\mathcal{C}_4$ occurs. Then by (2.2.11),

$$\begin{aligned}\mathcal{X}_{k_2}(s_2) &= \mathcal{X}_{k_1}(s_1) + \sum_{(i,b,s)\in P^{k_2,s_2}_{k_1,s_1}} X_{i,b,s} = \mathcal{X}_{k_1}(s_1) + \sum_{(i,b,s)\in P^{k_2,s_2}_{k_1,s_1}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\le \rho a_N\}} \\ &\le \mathcal{X}_{k_1}(s_1) + c_1 a_N\end{aligned}$$

by the definition of the event $\mathcal{C}_4$, which completes the proof. □

The next lemma says that on the event $\mathcal{C}_3 \cap \mathcal{C}_4$, if a path of length at most ℓ_N starts with a big jump then it moves distance at most $c_1 a_N$ after the big jump.

Lemma 2.3.4. *On the event $\mathcal{C}_3 \cap \mathcal{C}_4$, for all $(k_1, b_1, s_1) \in B_N$, $s_2 \in [\![s_1+1, \min\{s_1+\ell_N, t\}]\!]$ and $k_2 \in \mathcal{N}^{b_1}_{k_1,s_1}(s_2)$,*

$$\mathcal{X}_{k_2}(s_2) \le \mathcal{X}_{k_1}(s_1) + X_{k_1,b_1,s_1} + c_1 a_N,$$

where B_N and $\mathcal{N}^{b_1}_{k_1,s_1}(s_2)$ are defined in (2.2.15) *and* (2.2.13) *respectively.*

Proof. Let $l \in [N]$ be such that $(k_1, s_1) \lesssim_{b_1} (l, s_1+1)$, so that

$$\mathcal{X}_l(s_1+1) = \mathcal{X}_{k_1}(s_1) + X_{k_1,b_1,s_1}. \tag{2.3.18}$$

If $s_2 = s_1+1$ then we are done; from now on assume $s_2 \ge s_1+2$. Since X_{k_1,b_1,s_1} is a big jump, on the event $\mathcal{C}_3$ there are no further big jumps on the path between particles (l, s_1+1) and (k_2, s_2), that is $B_N \cap P^{k_2,s_2}_{l,s_1+1} = \emptyset$. Therefore, by Lemma 2.3.3 we have $\mathcal{X}_{k_2}(s_2) \le \mathcal{X}_l(s_1+1) + c_1 a_N$, which, together with (2.3.18), completes the proof. □

In the next lemma, we describe how we can exploit the fact that on the event $\mathcal{C}_5$ there are never two big jumps at the same time. First, the event $\mathcal{C}_5$ tells us that if a particle makes a big jump, then the other particles move very little at the time of the jump. Second, it also implies that if a particle significantly beats the current leader with a big jump, then it becomes the new leader, and the gap behind this new leader will be roughly the distance by which it beat the previous leader. Both statements follow immediately from the setup, but will be useful for example in the proofs of Corollaries 2.3.7 and 2.3.8 below, and later on in the proofs of Propositions 2.3.11 and 2.2.1 as well.

Lemma 2.3.5. *On the event $\mathcal{C}_5$, for all $(k, b, s) \in B_N$,*

(a) $\mathcal{X}_j(s+1) \le \mathcal{X}_N(s) + \rho a_N$ for all $j \in [N] \setminus \mathcal{N}^b_{k,s}(s+1)$, and

(b) if $\mathcal{X}_k(s) + X_{k,b,s} > \mathcal{X}_N(s) + c a_N$ for some $c > \rho$, then $(k, s) \lesssim_b (N, s+1)$ and $\mathcal{X}_N(s+1) - \mathcal{X}_{N-1}(s+1) > (c-\rho)a_N$.

Proof. Assume that $\mathcal{C}_5$ occurs and fix k, b, s as in the statement. Let $j \in [N] \setminus \mathcal{N}^b_{k,s}(s+1)$ be arbitrary. Assume that $i \in [N]$ and $b_i \in \{1, 2\}$ are such that $(i, s) \lesssim_{b_i} (j, s+1)$, and so $\mathcal{X}_j(s+1) = \mathcal{X}_i(s) + X_{i,b_i,s}$, with $(i,b_i) \in ([N]\times\{1,2\}) \setminus \{(k,b)\}$. By the definition of the event $\mathcal{C}_5$, $X_{k,b,s}$ is the only big jump at time s. Thus we have $X_{i,b_i,s} \le \rho a_N$, and by bounding the ith particle's position at time s by the rightmost position at time s we get

$$\mathcal{X}_j(s+1) = \mathcal{X}_i(s) + X_{i,b_i,s} \le \mathcal{X}_N(s) + \rho a_N,$$

which completes the proof of part (a). Furthermore, if the condition in (b) holds, then we also have

$$\mathcal{X}_j(s+1) \le \mathcal{X}_N(s) + \rho a_N < \mathcal{X}_k(s) + X_{k,b,s} - (c-\rho)a_N. \tag{2.3.19}$$

Since (2.3.19) holds for any $j \in [N] \setminus \mathcal{N}^b_{k,s}(s+1)$ and we are assuming $c > \rho$, we conclude that $(k,s) \lesssim_b (N, s+1)$, and the result follows by taking $j = N-1$ in (2.3.19). □

The next lemma says that if $\mathcal{C}_3 \cap \mathcal{C}_4$ occurs then all big jumps in the time interval $[t_3, t-1]$ come from close to the leftmost particle. Our heuristics suggest this should be true, because we expect most particles to be close to the leftmost particle at a typical time. However, the proof only relies on the assumption that the events $\mathcal{C}_3$ and $\mathcal{C}_4$ occur.

Lemma 2.3.6. *On the event $\mathcal{C}_3 \cap \mathcal{C}_4$,*

$$\mathcal{X}_k(s) \le \mathcal{X}_1(s) + c_1 a_N \quad \forall (k,b,s) \in B_N^{[t_3,t-1]}.$$

Proof. Take $s \in [\![t_3, t-1]\!]$, $k \in [N]$ and $b \in \{1,2\}$, and assume that we have $X_{k,b,s} > \rho a_N$. Let $i_k = \zeta_{k,s}(s-\ell_N)$ be the time-$(s-\ell_N)$ ancestor of particle (k,s) (recall (2.2.9)). Since $(k,b,s) \in B_N$, by the definition of the event $\mathcal{C}_3$, we must have $B_N \cap P^{k,s}_{i_k, s-\ell_N} = \emptyset$. Then by Lemma 2.3.3 we have

$$\mathcal{X}_k(s) \le \mathcal{X}_{i_k}(s-\ell_N) + c_1 a_N. \tag{2.3.20}$$

Furthermore, at time s every particle is to the right of $\mathcal{X}_{i_k}(s-\ell_N)$, by Lemma 2.2.3. This means $\mathcal{X}_{i_k}(s-\ell_N) \le \mathcal{X}_1(s)$, and so $\mathcal{X}_k(s) \le \mathcal{X}_1(s) + c_1 a_N$ by (2.3.20). □

We will use Lemma 2.3.6 to prove the next result, which says that on the event $\bigcap_{j=2}^5 \mathcal{C}_j$, if the diameter of the cloud of particles is large and a particle makes a big jump, then either it takes the lead and will be significantly ahead of the second rightmost particle, or it stays significantly behind the leader.

Corollary 2.3.7. *On the event $\bigcap_{j=2}^5 \mathcal{C}_j$, if $(k,b,s) \in B_N^{[t_3,t-1]}$ and $d(\mathcal{X}(s)) \ge (c_3 + c_1)a_N$ then*

(a) if $X_{k,b,s} > Z_k(s)$ then $\mathcal{X}_N(s+1) = \mathcal{X}_k(s) + X_{k,b,s} > \mathcal{X}_{N-1}(s+1) + (2c_2 - \rho)a_N$, and

(b) if $X_{k,b,s} \le Z_k(s)$ then $\mathcal{X}_k(s) + X_{k,b,s} \le \mathcal{X}_N(s) - 2c_2 a_N$,

where $Z_k(s)$ and $\mathcal{C}_2, \dots, \mathcal{C}_5$ are given by (2.3.11)–(2.3.15).

Proof. Since $X_{k,b,s}$ is a big jump, by Lemma 2.3.6 and the fact that $d(\mathcal{X}(s)) = \mathcal{X}_N(s) - \mathcal{X}_1(s) \geq (c_3 + c_1)a_N$,

$$\mathcal{X}_k(s) \leq \mathcal{X}_1(s) + c_1 a_N \leq \mathcal{X}_N(s) - c_3 a_N.$$

Hence the gap between the kth particle and the rightmost particle is bounded below by $c_3 a_N$:

$$Z_k(s) \geq c_3 a_N. \tag{2.3.21}$$

It follows that if $X_{k,b,s} > Z_k(s)$, by the definition of the event $\mathcal{C}_2$ we have $X_{k,b,s} > Z_k(s) + 2c_2 a_N$, which implies that

$$X_{k,b,s} + \mathcal{X}_k(s) > \mathcal{X}_N(s) + 2c_2 a_N.$$

Since $2c_2 > \rho$ by (2.3.3) and (2.3.4), Lemma 2.3.5(b) implies the statement of part (a). If instead $X_{k,b,s} \leq Z_k(s)$, then by (2.3.21) and the definition of $\mathcal{C}_2$, we have $X_{k,b,s} \leq Z_k(s) - 2c_2 a_N$, which completes the proof. □

The next result says that on the event $\bigcap_{j=2}^{5} \mathcal{C}_j$, if the diameter of the cloud of particles is big at some time s, then if at time s or $s-1$ a particle makes a big jump which beats the current leader, this particle becomes the new leader.

Corollary 2.3.8. *On the event $\bigcap_{j=2}^{5} \mathcal{C}_j$, for all $s \in [\![t_3+1, t-1]\!]$, if $d(\mathcal{X}(s)) \geq \frac{3}{2} c_3 a_N$ then*

$$s \in \mathbf{S}_N \iff s \in \hat{\mathbf{S}}_N \quad \textit{and} \quad s-1 \in \mathbf{S}_N \iff s-1 \in \hat{\mathbf{S}}_N,$$

where $\mathbf{S}_N$ and $\hat{\mathbf{S}}_N$ are defined in (2.2.16) *and* (2.2.18).

Proof. Take $s \in [\![t_3+1, t-1]\!]$ and suppose $d(\mathcal{X}(s)) \geq \frac{3}{2} c_3 a_N$.

If $s \in \mathbf{S}_N$, then there exists $(k,b,s) \in B_N$ such that $\mathcal{X}_k(s) + X_{k,b,s} = \mathcal{X}_N(s+1) \geq \mathcal{X}_N(s)$, where we used monotonicity for the inequality. To show that $s \in \hat{\mathbf{S}}_N$, we need to show that in fact $\mathcal{X}_k(s) + X_{k,b,s} > \mathcal{X}_N(s)$, i.e. the inequality is strict, but this follows from Corollary 2.3.7(b), which applies since $d(\mathcal{X}(s)) \geq \frac{3}{2} c_3 a_N \geq (c_1 + c_3)a_N$ by (2.3.4).

Now suppose $s \in \hat{\mathbf{S}}_N$. Since $d(\mathcal{X}(s)) \geq \frac{3}{2} c_3 a_N \geq (c_1 + c_3)a_N$, and by the definition of $\hat{\mathbf{S}}_N$, the conditions of Corollary 2.3.7(a) hold for (k,b,s), for some $(k,b) \in [N] \times \{1,2\}$. Then Corollary 2.3.7(a) implies that $s \in \mathbf{S}_N$, and therefore the first equivalence in the statement holds.

If $d(\mathcal{X}(s-1)) \geq (c_3 + c_1)a_N$, then we can repeat the proof of the first equivalence to show that $s-1 \in \mathbf{S}_N \iff s-1 \in \hat{\mathbf{S}}_N$.

If instead $d(\mathcal{X}(s-1)) < (c_3 + c_1)a_N$ we argue as follows. Suppose $s-1 \in \mathbf{S}_N$. Then there exists $(k,b,s-1) \in B_N$ such that

$$\mathcal{X}_k(s-1) + X_{k,b,s-1} = \mathcal{X}_N(s) \geq \mathcal{X}_N(s-1),$$

which means $X_{k,b,s-1} \geq Z_k(s-1)$. Now $X_{k,b,s-1} = Z_k(s-1)$ is impossible because, with the assumption that $d(\mathcal{X}(s-1)) < (c_3+c_1)a_N$, it would imply

$$\mathcal{X}_N(s) = \mathcal{X}_k(s-1) + Z_k(s-1) = \mathcal{X}_N(s-1) < \mathcal{X}_1(s-1) + (c_3+c_1)a_N < \mathcal{X}_1(s) + \tfrac{3}{2}c_3 a_N$$

by monotonicity and (2.3.4). This contradicts the assumption $d(\mathcal{X}(s)) \geq \frac{3}{2}c_3 a_N$ from the statement of this corollary. Hence, we must have $X_{k,b,s-1} > Z_k(s-1)$, and so $s-1 \in \hat{\mathbf{S}}_N$.

Now suppose $s-1 \in \hat{\mathbf{S}}_N$, and take $(k,b,s-1) \in B_N$ such that $\mathcal{X}_k(s-1) + X_{k,b,s-1} > \mathcal{X}_N(s-1)$. Then by Lemma 2.3.5(a) and the assumption on $d(\mathcal{X}(s-1))$, for all $j \in [N] \setminus \mathcal{N}^b_{k,s-1}(s)$ we have

$$\mathcal{X}_j(s) \leq \mathcal{X}_N(s-1) + \rho a_N < \mathcal{X}_1(s-1) + (c_3+c_1+\rho)a_N.$$

By monotonicity, (2.3.3) and (2.3.4) this is strictly smaller than $\mathcal{X}_1(s) + \frac{3}{2}c_3 a_N$. Thus, at time s, all particles not in $\mathcal{N}^b_{k,s-1}(s)$ are closer than distance $\frac{3}{2}c_3 a_N$ to the leftmost particle. Hence, since we assumed that $d(\mathcal{X}(s)) \geq \frac{3}{2}c_3 a_N$, we must have $(k,s-1) \lesssim_b (N,s)$, which means that $s-1 \in \mathbf{S}_N$. □

The last property we state before the proof of Proposition 2.3.2 says the following. First, if no particle beats the leader with a big jump for a time interval of length at most ℓ_N, then the leader's position does not change much during this time. We will use the extra condition that the diameter is not too small to prove this easily; if the diameter is too small then jumps that are "almost big" could complicate matters. Second, the lemma says that if the diameter becomes small at some point, then it cannot become too large within ℓ_N time, if there is no particle which beats the leader with a big jump. Recall the definition of $\hat{\mathbf{S}}_N$ from (2.2.18).

Lemma 2.3.9. *On the event $\bigcap_{j=2}^5 \mathcal{C}_j$, for all $s \in [\![t_3, t_1-1]\!]$ and $\Delta s \in [\ell_N]$, if $s+\Delta s \leq t_1$ and $[\![s, s+\Delta s-1]\!] \subseteq \hat{\mathbf{S}}_N^c$ then the following statements hold:*

(a) If $d(\mathcal{X}(r)) \geq \frac{3}{2}c_3 a_N$ for all $r \in [\![s, s+\Delta s-1]\!]$, then $\mathcal{X}_N(s+\Delta s) \leq \mathcal{X}_N(s) + c_1 a_N$. In particular, if $\Delta s = \ell_N$ then $d(\mathcal{X}(s+\ell_N)) \leq c_1 a_N$.

(b) If there exists $r \in [\![s, s+\Delta s-1]\!]$ such that $d(\mathcal{X}(r)) \leq \frac{3}{2}c_3 a_N$, then $d(\mathcal{X}(s+\Delta s)) \leq \frac{3}{2}c_3 a_N + 2c_1 a_N$.

Proof. First we prove part (a). Let $i,j \in [N]$ with $(i,s) \lesssim (j, s+\Delta s)$. Assume that there is a big jump on the path between $\mathcal{X}_i(s)$ and $\mathcal{X}_j(s+\Delta s)$ at time $s' \in [\![s, s+\Delta s-1]\!]$, i.e. there exists $(k',b',s') \in B_N \cap P^{j,s+\Delta s}_{i,s}$. Since we assume $s' \in \hat{\mathbf{S}}^c_N$, we have $\mathcal{X}_{k'}(s') + X_{k',b',s'} \leq \mathcal{X}_N(s')$. Then since we assume $d(\mathcal{X}(s')) \geq \frac{3}{2}c_3 a_N > (c_3+c_1)a_N$ by (2.3.4), we can apply Corollary 2.3.7(b) to obtain

$$\mathcal{X}_{k'}(s') + X_{k',b',s'} \leq \mathcal{X}_N(s') - 2c_2 a_N. \tag{2.3.22}$$

Therefore, first by Lemma 2.3.4, second by (2.3.22), and third by monotonicity and (2.3.4) we get

$$\mathcal{X}_j(s+\Delta s) \le \mathcal{X}_{k'}(s') + X_{k',b',s'} + c_1 a_N \le \mathcal{X}_N(s') - 2c_2 a_N + c_1 a_N < \mathcal{X}_N(s+\Delta s).$$

Hence $j \neq N$, which means that the leader at time $s+\Delta s$ must be a particle which does not have an ancestor which made a big jump in the time interval $[s, s+\Delta s - 1]$. That is, $B_N \cap P_{i,s}^{N,s+\Delta s} = \emptyset$ for all $i \in [N]$. But then by Lemma 2.3.3 we must have

$$\mathcal{X}_N(s+\Delta s) \le \mathcal{X}_N(s) + c_1 a_N,$$

which shows the first statement of part (a). By Lemma 2.2.3 we also have $\mathcal{X}_1(s+\ell_N) \ge \mathcal{X}_N(s)$, and the second statement of part (a) follows.

Now we prove part (b). Let τ_d denote the last time before $s+\Delta s$ when the diameter is at most $\frac{3}{2}c_3 a_N$, that is

$$\tau_d = \sup\left\{r \le s+\Delta s : \ d(\mathcal{X}(r)) \le \tfrac{3}{2}c_3 a_N\right\}.$$

By our assumption in part (b) we have $\tau_d \ge s$.

If $\tau_d = s+\Delta s$ then we are done. Assume instead that $\tau_d < s+\Delta s$. Then we can estimate the leftmost particle position at time $s+\Delta s$ using monotonicity and the definition of τ_d:

$$\mathcal{X}_1(s+\Delta s) \ge \mathcal{X}_1(\tau_d) \ge \mathcal{X}_N(\tau_d) - \tfrac{3}{2}c_3 a_N. \tag{2.3.23}$$

To estimate the rightmost position, we first use the fact that $\tau_d \in [\![s, s+\Delta s-1]\!] \subseteq \hat{\mathbf{S}}_N^c$ and $d(\mathcal{X}(\tau_d+1)) > \frac{3}{2}c_3 a_N$ by the definition of τ_d. Hence, the second equivalence of Corollary 2.3.8 implies that $\tau_d \in \mathbf{S}_N^c$; that is, no big jump takes the lead at time $\tau_d + 1$. Thus, for some $(k,b) \in [N] \times \{1,2\}$ we have

$$\mathcal{X}_N(\tau_d+1) = \mathcal{X}_k(\tau_d) + X_{k,b,\tau_d} \le \mathcal{X}_N(\tau_d) + \rho a_N. \tag{2.3.24}$$

Now (2.3.23), (2.3.24) and (2.3.3) show that if $\tau_d = s+\Delta s-1$ then we are done. Assume instead that $\tau_d < s+\Delta s-1$. Then we can apply part (a) for the time interval $[\tau_d+1, s+\Delta s]$, because $d(\mathcal{X}(r)) > \frac{3}{2}c_3 a_N$ $\forall r \in [\![\tau_d+1, s+\Delta s]\!]$ by the definition of τ_d. So by part (a) and then by (2.3.24) we have

$$\mathcal{X}_N(s+\Delta s) \le \mathcal{X}_N(\tau_d+1) + c_1 a_N \le \mathcal{X}_N(\tau_d) + (\rho + c_1)a_N. \tag{2.3.25}$$

Now (2.3.25), (2.3.23) and (2.3.3) yield part (b). $\square$

Proof of Proposition 2.3.2. The main effort of this proof is in showing that the $\mathcal{C}_i$ events imply $\mathcal{B}_1$. So we want to see a leader tribe at time t_1 in which all the particles are descended from particle (N,T), and are significantly to the right of all the particles not descended from particle (N,T). We begin by giving an outline of how this will be proved.

Outline of proof that $\mathcal{C}_1$ to $\mathcal{C}_7$ imply $\mathcal{B}_1$

Assume the event $\cap_{j=1}^{7}\mathcal{C}_j$ occurs. On the event $\mathcal{C}_1$ there will be a time $\tau_1 \in [t_2+1, t_1]$ when the leader, particle (N, τ_1), is a distance more than $2c_3a_N$ ahead of the second rightmost (and every other) particle. Having this gap at time τ_1 will ensure that the back of the population is further than $\frac{3}{2}c_3a_N$ away from the leader at all times up to t_1. That is, the diameter cannot be too small after time τ_1, and so we will be able to apply Corollary 2.3.7.

It is a possibility that on the time interval $[\tau_1, t_1]$, every particle not descended from (N, τ_1) stays further than roughly $2c_2a_N$ to the left of the tribe descending from (N, τ_1). Then we will have the desired leader tribe with a gap behind it at time t_1. Alternatively, the tribe of particle (N, τ_1) may be surpassed by other particles. But then, by Corollary 2.3.7(a), the leader must be beaten by at least roughly $2c_2a_N$. The new leader's descendants might be surpassed too, but again by at least $2c_2a_N$. Then, after the last time T when a tribe is surpassed before t_1 (i.e. the last time when a big jump takes the lead, see (2.2.17)), no particle will make a big jump that gets closer to the leader tribe than $2c_2a_N$, by Corollary 2.3.7(b). We will see that this implies that at time t_1, the leader tribe will be further away than c_2a_N from all the other particles. This argument works if the particles of the tribes do not move far from the position of their ancestor which made a big jump. We have this property due to Lemma 2.3.4.

Therefore, the proof will expand on the following steps:

(i) The record is broken by a big jump at time τ_1. Therefore time T, the last time when the record is broken by a big jump before time t_1, is either at time τ_1 or later.

(ii) The diameter is at least $\frac{3}{2}c_3a_N$ between times τ_1 and t_1.

We will show that the back of the population stays far behind $\mathcal{X}_N(\tau_1)$, because of the small number of big jumps compared to the number of particles. This is useful, because most of the lemmas and corollaries above will apply if the diameter is not too small.

(iii) At time T, the last time before t_1 when a particle takes the lead with a big jump, there will be a gap of size at least $\frac{3}{2}c_2a_N$ between the leader (N, T) and the second rightmost particle $(N-1, T)$.

This step follows by Corollary 2.3.7(a), which we can apply because of step (ii). If the diameter is big and the leader is beaten, then the new leader will lead by a large distance.

(iv) Every other particle stays at least distance c_2a_N behind the descendants of particle (N, T) until time t_1.

This is mainly due to steps (ii) and (iii) and Corollary 2.3.7(b): if the diameter is big and the leader is not beaten by a big jump, then big jumps will arrive far behind the leader. Therefore, the gap behind the leader tribe created in step (iii) will remain until time t_1.

(v) The leading tribe has the size required by the event $\mathcal{B}_1$, and thus the event $\mathcal{B}_1$ occurs.

Proof that $\mathcal{C}_1$ to $\mathcal{C}_7$ imply $\mathcal{B}_1$
We now give a detailed proof, following steps (i)-(v) above, that

$$\bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \mathcal{B}_1. \tag{2.3.26}$$

Assume that $\bigcap_{j=1}^{7} \mathcal{C}_j$ occurs. We first check that we have $T \in [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]$ by proving the following statement.

Step (i). We have $t_2 + \lceil \delta\ell_N \rceil < \tau_1 \le T \le t_1 - \lceil \delta\ell_N \rceil$, where τ_1 and T are defined in (2.3.9) *and* (2.2.17).

In order to see this, we will use the following simple property:

$$\mathcal{X}_j(s-1) \le \mathcal{X}_{N-1}(s) \quad \forall s \in \mathbb{N} \text{ and } j \in [N]. \tag{2.3.27}$$

Indeed, since all jumps are non-negative, and particle $(N, s-1)$ has two offspring, there are at least two particles to the right of (or at) position $\mathcal{X}_N(s-1)$ at time s. Thus $\mathcal{X}_N(s-1) \le \mathcal{X}_{N-1}(s)$, which shows (2.3.27).

By the definition of the event $\mathcal{C}_1$, we have $\tau_1 \in [\![t_2+1, t_1]\!]$. Let $(\hat{J}, \hat{b}) \in [N] \times \{1,2\}$ be such that $(\hat{J}, \tau_1 - 1) \lesssim_{\hat{b}} (N, \tau_1)$, and so $\mathcal{X}_N(\tau_1) = \mathcal{X}_{\hat{J}}(\tau_1 - 1) + X_{\hat{J},\hat{b},\tau_1-1}$. It also follows from (2.3.27) that $\mathcal{X}_{\hat{J}}(\tau_1 - 1) \le \mathcal{X}_{N-1}(\tau_1)$. Hence the definition of τ_1 in (2.3.9) implies that $X_{\hat{J},\hat{b},\tau_1-1} > 2c_3 a_N$, which means that $X_{\hat{J},\hat{b},\tau_1-1}$ is a big jump, and so cannot happen on the time interval $[t_2, t_2 + \lceil \delta\ell_N \rceil]$ by the definition of $\mathcal{C}_6$. This implies the first inequality in Step (i). We also notice that $X_{\hat{J},\hat{b},\tau_1-1}$ is a big jump which takes the lead at time τ_1, that is $\tau_1 - 1 \in \mathbf{S}_N$ (see (2.2.16)). Then we have $T \ge \tau_1$ by the definition of T in (2.2.17), which shows the second inequality of Step (i). Furthermore, the definition of T also shows that $T > t_1 - \lceil \delta\ell_N \rceil$ is not possible on $\mathcal{C}_6$, which concludes the third inequality and the proof of Step (i).

Since we now know that $T \neq 0$, particle (N, T) is the last particle which broke the record with a big jump before time t_1. Take $(J, b^*) \in [N] \times \{1, 2\}$ such that $(J, T-1) \lesssim_{b^*} (N, T)$, so

$$\mathcal{X}_N(T) = \mathcal{X}_J(T-1) + X_{J,b^*,T-1}, \tag{2.3.28}$$

with $X_{J,b^*,T-1} > \rho a_N$. That is, at time $T-1$ the Jth particle's b^*th offspring performed a big jump $X_{J,b^*,T-1}$, with which it became the leader at time T at position $\mathcal{X}_N(T)$. We will show that at time t_1 there is a leader tribe in which every particle descends from particle (N, T). Our next step towards this statement is to show that the diameter is large between times τ_1 and t_1.

Step (ii). We have $d(\mathcal{X}(s)) \ge \frac{3}{2} c_3 a_N$ for all $s \in [\![\tau_1, t_1]\!]$.

We prove Step (ii) by showing that the number of particles within distance $\frac{3}{2}c_3 a_N$ of the leader is strictly smaller than N at all times in $[\![\tau_1, t_1]\!]$.

Let $s \in [\![\tau_1, t_1]\!]$. Consider an arbitrary particle (i, s) in the population at time s. We first claim that if

$$\mathcal{X}_i(s) > \mathcal{X}_N(\tau_1) - \tfrac{3}{2}c_3 a_N, \tag{2.3.29}$$

then particle (i, s) has an ancestor which made a big jump at some time $\tilde{s} \in [\![\tau_1 - 1, s-1]\!]$. That is, if (2.3.29) holds then

$$B_N \cap P^{i,s}_{j,\tau_1-1} \neq \emptyset, \quad \text{for some } j \in [N]. \tag{2.3.30}$$

To see this, we notice that

$$\mathcal{X}_j(\tau_1 - 1) \leq \mathcal{X}_{N-1}(\tau_1) < \mathcal{X}_N(\tau_1) - 2c_3 a_N \quad \forall j \in [N], \tag{2.3.31}$$

where the first inequality follows by (2.3.27), and the second from the definition of τ_1. Therefore, by (2.3.29), (2.3.31) and (2.3.4), we have

$$\mathcal{X}_i(s) > \mathcal{X}_j(\tau_1 - 1) + c_1 a_N \quad \forall j \in [N]. \tag{2.3.32}$$

In particular, this holds for $j \in [N]$ such that $(j, \tau_1 - 1) \lesssim (i, s)$. Therefore (2.3.30) must hold by Lemma 2.3.3, showing that our claim is true.

Thus, every particle which is to the right of $\mathcal{X}_N(\tau_1) - \frac{3}{2}c_3 a_N$ at time s has an ancestor which made a big jump between times $\tau_1 - 1$ and $s - 1$. This gives us the following bound:

$$\#\left\{i \in [N] : \mathcal{X}_i(s) > \mathcal{X}_N(\tau_1) - \tfrac{3}{2}c_3 a_N\right\} \leq \sum_{(l,b,r)\in B_N^{[\tau_1-1,s-1]}} |\mathcal{N}^b_{l,r}(s)|, \tag{2.3.33}$$

where $\mathcal{N}^b_{l,r}(s)$ and $B_N^{[\tau_1-1,s-1]}$ are defined in (2.2.13) and (2.2.14) respectively. On the right-hand side we sum the number of descendants of all particles which made a big jump between times $\tau_1 - 1$ and $s-1$. We want to show that this is smaller than N, because that means that there must be at least one particle to the left of (or at) $\mathcal{X}_N(\tau_1) - \frac{3}{2}c_3 a_N$ at time s.

Since $[\tau_1 - 1, s] \subseteq [t_2 + \lceil \delta \ell_N \rceil, t_1]$ by Step (i), any particle at a time in $[\tau_1 - 1, s-1]$ has at most $2^{t_1-(t_2+\lceil\delta\ell_N\rceil)}$ descendants at time s. Furthermore, the number of big jumps in the time interval $[\tau_1 - 1, s-1]$ is at most K, by the definition of $\mathcal{C}_7$. Hence, by (2.3.33) and then since $t_1 - t_2 = \ell_N$,

$$\#\left\{i \in [N] : \mathcal{X}_i(s) > \mathcal{X}_N(\tau_1) - \tfrac{3}{2}c_3 a_N\right\} \leq K 2^{t_1-(t_2+\lceil\delta\ell_N\rceil)} \leq 2KN^{1-\delta} < N, \tag{2.3.34}$$

by our assumption on N in the statement of Proposition 2.3.2. Therefore, by (2.3.34) and monotonicity we must have $\mathcal{X}_1(s) \leq \mathcal{X}_N(\tau_1) - \frac{3}{2}c_3 a_N \leq \mathcal{X}_N(s) - \frac{3}{2}c_3 a_N$, which concludes the proof of Step (ii).

Next we show that there is a gap between the two rightmost particles at time T.

Step (iii). We have $\mathcal{X}_{N-1}(T) + \frac{3}{2}c_2 a_N < \mathcal{X}_N(T)$.

Note that we have $\tau_1 \leq T$ by Step (i). If $T = \tau_1$ then the statement of Step (iii) holds by the definition of τ_1 and (2.3.4).

Suppose instead that $T > \tau_1$. We now check the conditions of Corollary 2.3.7(a). Recall from (2.3.28) that $X_{J,b^*,T-1}$ is a big jump. Since the particle performing the jump $X_{J,b^*,T-1}$ becomes the leader at time T, we have $X_{J,b^*,T-1} \geq Z_J(T-1)$, where $Z_J(T-1)$ is the gap between the Jth particle and the leader at time $T-1$. Also note that $(J, b^*, T-1) \in B_N^{[t_2,t_1]}$, and that by Step (ii) and (2.3.4) we have $d(\mathcal{X}(T-1)) > (c_3+c_1)a_N$. Therefore Corollary 2.3.7 (using part (a) when $X_{J,b^*,T-1} > Z_J(T-1)$; part (b) shows that we cannot have $X_{J,b^*,T-1} = Z_J(T-1)$ since otherwise the particle performing the jump $X_{J,b^*,T-1}$ would not take the lead at time T) implies

$$\mathcal{X}_N(T) = \mathcal{X}_J(T-1) + X_{J,b^*,T-1} > \mathcal{X}_{N-1}(T) + (2c_2 - \rho)a_N,$$

which together with (2.3.3) and (2.3.4) shows the statement of Step (iii).

In Step (iv) we show that every particle which does not descend from particle (N,T) is to the left of $\mathcal{X}_N(T) - c_2 a_N$ at time t_1.

Step (iv). Let $i \in [N-1]$ *and* $j \in [N]$*. If* $(i,T) \lesssim (j,t_1)$ *then* $\mathcal{X}_j(t_1) \leq \mathcal{X}_N(T) - c_2 a_N$.

First we will use Lemma 2.3.9(a) to bound $\mathcal{X}_N(t_1)$. Since T is the last time when a particle took the lead with a big jump before time t_1, we have $[\![T, t_1-1]\!] \subseteq \mathbf{S}_N^c$, where $\mathbf{S}_N$ is defined in (2.2.16). By Corollary 2.3.8 and Steps (i) and (ii), it follows that $[\![T, t_1-1]\!] \subseteq \hat{\mathbf{S}}_N^c$. Therefore the conditions of Lemma 2.3.9(a) hold with $s = T$ and $\Delta s = t_1 - T$. Then Lemma 2.3.9(a) yields

$$\mathcal{X}_N(t_1) \leq \mathcal{X}_N(T) + c_1 a_N. \tag{2.3.35}$$

Now we prove the upper bound on $\mathcal{X}_j(t_1)$ in the statement of Step (iv). Let us first consider the case in which there is no big jump in the path between particles (i,T) and (j,t_1), i.e. $B_N \cap P_{i,T}^{j,t_1} = \emptyset$. Then, by Lemma 2.3.3, Step (iii) and (2.3.4) we have

$$\mathcal{X}_j(t_1) \leq \mathcal{X}_i(T) + c_1 a_N < \mathcal{X}_N(T) - \tfrac{3}{2}c_2 a_N + c_1 a_N < \mathcal{X}_N(T) - c_2 a_N,$$

which shows that the statement of Step (iv) holds in this case.

Now suppose instead that there exists a big jump on the path between particles (i,T) and (j,t_1), so assume we have some $(l,b,r) \in B_N \cap P_{i,T}^{j,t_1}$. We will show that, even with the big jump $X_{l,b,r}$, particle (j,t_1) cannot arrive close to the leader particle (N,t_1) at time t_1. This fact together with (2.3.35) will imply Step (iv).

We know that $[\![T, t_1-1]\!] \subseteq \hat{\mathbf{S}}_N^c$, and so, in particular, the leader at time r is not beaten by the big jump $X_{l,b,r}$. Hence by the definition of $Z_l(r)$ in (2.3.11) we have $X_{l,b,r} \leq Z_l(r)$.

Therefore, because of Steps (i) and (ii) and by (2.3.4), Corollary 2.3.7(b) applies, which implies

$$\mathcal{X}_l(r) + X_{l,b,r} \le \mathcal{X}_N(r) - 2c_2a_N. \tag{2.3.36}$$

Now by Lemma 2.3.4 and since $t_1 - T < \ell_N$ by Step (i), then by (2.3.36), and finally by monotonicity,

$$\mathcal{X}_j(t_1) \le \mathcal{X}_l(r) + X_{l,b,r} + c_1a_N \le \mathcal{X}_N(r) - 2c_2a_N + c_1a_N \le \mathcal{X}_N(t_1) - 2c_2a_N + c_1a_N. \tag{2.3.37}$$

Putting (2.3.37) and (2.3.35) together and then using (2.3.4), we obtain

$$\mathcal{X}_j(t_1) \le \mathcal{X}_N(T) - 2c_2a_N + 2c_1a_N \le \mathcal{X}_N(T) - c_2a_N,$$

which finishes the proof of Step (iv).

Step (v). The event $\mathcal{B}_1$, as defined in (2.3.7), occurs.

Let us simplify the notation by writing $R = R_{c_1,N}(t_1)$, where $R_{c_1,N}(t_1)$ is given by (2.3.6). To prove that $\mathcal{B}_1$ occurs, we first show that

$$\mathcal{N}_{N,T}(t_1) = \{j \in [N] : \mathcal{X}_j(t_1) \ge \mathcal{X}_N(t_1) - c_1a_N\} = \{N - R + 1, \dots, N\}. \tag{2.3.38}$$

The second equality follows directly from the definition of R; we will prove the first equality.

Note that Step (iv) implies that every descendant of particle (N, T) survives until time t_1, that is $|\mathcal{N}_{N,T}(t_1)| = 2^{t_1-T} > 1$. Indeed, by Step (i) and our assumption on N we have $2^{t_1-T} \le 2N^{1-\delta} < N$, thus at time t_1 there are at least 2^{t_1-T} particles to the right of (or at) position $\mathcal{X}_N(T)$ by Lemma 2.2.3. By Step (iv), particles not descended from particle (N, T) are to the left of position $\mathcal{X}_N(T)$ at time t_1. Therefore, particle (N, T) must have 2^{t_1-T} surviving descendants at time t_1, since otherwise there would not be 2^{t_1-T} particles to the right of (or at) position $\mathcal{X}_N(T)$.

The above argument also implies that the leader at time t_1 must be a descendant of particle (N, T), i.e. $N \in \mathcal{N}_{N,T}(t_1)$. Furthermore, as all jumps are non-negative, and by (2.3.35), we have

$$\mathcal{X}_k(t_1) \in [\mathcal{X}_N(T), \mathcal{X}_N(T) + c_1a_N] \quad \forall k \in \mathcal{N}_{N,T}(t_1). \tag{2.3.39}$$

In particular, using the above and (2.3.35) again, we must have $\mathcal{X}_k(t_1) \ge \mathcal{X}_N(T) \ge \mathcal{X}_N(t_1) - c_1a_N$ for all $k \in \mathcal{N}_{N,T}(t_1)$.

By Step (iv) and then by monotonicity and (2.3.4),

$$\mathcal{X}_j(t_1) \le \mathcal{X}_N(T) - c_2a_N < \mathcal{X}_N(t_1) - c_1a_N \quad \forall j \in [N] \setminus \mathcal{N}_{N,T}(t_1),$$

and (2.3.38) follows.

Next we check that

$$\mathcal{X}_{N-R}(t_1) \le \mathcal{X}_{N-R+1}(t_1) - c_2 a_N. \tag{2.3.40}$$

By (2.3.38) we see that $N-R+1 \in \mathcal{N}_{N,T}(t_1)$ and $N-R \notin \mathcal{N}_{N,T}(t_1)$. Therefore, Step (iv) and (2.3.39) imply (2.3.40).

Finally, we need to show that

$$R \le 2N^{1-\delta}. \tag{2.3.41}$$

We have that

$$R = |\{N-R+1,\dots,N\}| = |\mathcal{N}_{N,T}(t_1)| \le 2^{t_1-(t_2+\lceil \delta \ell_N \rceil)} \le 2N^{1-\delta},$$

where in the second equality we used (2.3.38), and the inequality follows since $T > t_2 + \lceil \delta \ell_N \rceil$ by Step (i). Therefore by Step (i), (2.3.38), (2.3.40) and (2.3.41), $\mathcal{B}_1$ occurs, which concludes Step (v).

This completes the proof of (2.3.26).

Proof that $\mathcal{C}_1$ to $\mathcal{C}_7$ imply $\mathcal{B}_2$

Recall the definition of the event $\mathcal{B}_2$ in (2.3.8). We now prove that

$$\mathcal{B}_1 \cap \mathcal{C}_4 \cap \mathcal{C}_6 \cap \mathcal{C}_7 \subseteq \mathcal{B}_2, \tag{2.3.42}$$

which implies $\bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \mathcal{B}_2$ because of (2.3.26).

Assume that $\mathcal{B}_1 \cap \mathcal{C}_4 \cap \mathcal{C}_6 \cap \mathcal{C}_7$ occurs. Again write $R = R_{c_1,N}(t_1)$, where $R_{c_1,N}(t_1)$ is defined using (2.3.6). Take $j \in [N-R]$ and consider particle (j, t_1). Then, by the definition of the event $\mathcal{B}_1$ in (2.3.7), and since the leader at time t_1 is to the right of every particle at time t_1, we have

$$\mathcal{X}_j(t_1) \le \mathcal{X}_{N-R+1}(t_1) - c_2 a_N \le \mathcal{X}_N(t_1) - c_2 a_N. \tag{2.3.43}$$

Now suppose that the ith particle at time t is a descendant of particle (j, t_1), i.e. $i \in \mathcal{N}_{j,t_1}(t)$. Lemma 2.2.3 implies that every particle at time t is to the right of (or at) $\mathcal{X}_N(t_1)$. Thus we have $\mathcal{X}_i(t) \ge \mathcal{X}_N(t_1)$, which together with (2.3.43) and (2.3.4) implies

$$\mathcal{X}_i(t) > \mathcal{X}_j(t_1) + c_1 a_N.$$

Thus, by Lemma 2.3.3, there must be a big jump in the path between particles (j, t_1) and (i, t); that is, we must have $B_N \cap P_{j,t_1}^{i,t} \ne \emptyset$.

Therefore we can bound the number of time-t descendants of particles $(1, t_1)$, $(2, t_1)$, $\dots$, $(N-R, t_1)$ by the number of descendants of particles which made a big jump between times t_1 and $t-1$:

$$\sum_{j=1}^{N-R} |\mathcal{N}_{j,t_1}(t)| \leq \sum_{(k,b,s)\in B_N^{[t_1,t-1]}} |\mathcal{N}^b_{k,s}(t)|. \tag{2.3.44}$$

By the definition of the event $\mathcal{C}_6$, no particle makes a big jump in the time interval $[t_1 - \lceil \delta\ell_N \rceil, t_1 + \lceil \delta\ell_N \rceil]$. Hence, any particle which made a big jump between times t_1 and $t-1$ can have at most $2^{t-(t_1+\lceil \delta\ell_N \rceil)}$ descendants at time t. Furthermore, by the definition of $\mathcal{C}_7$, $|B_N^{[t_1,t-1]}| \leq K$. Putting these observations together with (2.3.44) we obtain

$$\sum_{j=1}^{N-R} |\mathcal{N}_{j,t_1}(t)| \leq 2KN^{1-\delta} < N^{1-\gamma}, \tag{2.3.45}$$

by our assumption on N in the statement of the proposition. This completes the proof of (2.3.42).

Proof that $\mathcal{C}_1$ to $\mathcal{C}_7$ imply $\mathcal{A}_1$

Recall the definition of $\mathcal{A}_1$ in (2.2.3). We now complete the proof of Proposition 2.3.2 by showing that

$$\bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \mathcal{A}_1. \tag{2.3.46}$$

Assume $\bigcap_{j=1}^7 \mathcal{C}_j$ occurs. Let $i,j \in [N]$ be such that $(j,t_1) \lesssim (i,t)$. Assume first that $B_N \cap P^{i,t}_{j,t_1} = \emptyset$. Then, by Lemma 2.3.3 and using the leader's position as an upper bound, we obtain

$$\mathcal{X}_i(t) \leq \mathcal{X}_j(t_1) + c_1 a_N \leq \mathcal{X}_N(t_1) + c_1 a_N \leq \mathcal{X}_1(t) + c_1 a_N,$$

where the last inequality follows by Lemma 2.2.3. Thus, recalling the definition of $L_{c_1,N}(t)$ in (2.2.1), we have $i \in [L_{c_1,N}(t)]$. Therefore, if $i > L_{c_1,N}(t)$ then we must have $B_N \cap P^{i,t}_{j,t_1} \neq \emptyset$. It follows that

$$N - L_{c_1,N}(t) \leq \sum_{(k,b,s)\in B_N^{[t_1,t-1]}} |\mathcal{N}^b_{k,s}(t)| < N^{1-\gamma}$$

by the same argument as for (2.3.45). Since we took $c_1 < \eta$ in (2.3.4), we now have $L_{\eta,N}(t) \geq N - N^{1-\gamma}$, which finishes the proof of (2.3.46). The proof of Proposition 2.3.2 then follows from (2.3.26), (2.3.42), (2.3.46) and Step (ii). □

2.3.3 Breaking down event $\mathcal{C}_1$

We have now broken down the events $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{A}_1$ into simpler events $\mathcal{C}_1$ to $\mathcal{C}_7$. In Section 2.4 we will be able to show directly that the events $\mathcal{C}_2$ to $\mathcal{C}_7$ occur with high probability. However, we will need to break $\mathcal{C}_1$ down further, into simpler events that we will show occur with high probability in Section 2.4. In this section we carry out the task of breaking down $\mathcal{C}_1$, which says that a gap of size $2c_3a_N$ appears behind the rightmost particle at some point during the time interval $[t_2+1, t_1]$ (see (2.3.10)), into simpler events. Recall that we assumed $t > 4\ell_N$, and that the constants $\eta \in (0,1]$, γ, δ, ρ, $c_1, c_2, \dots, c_6$, K satisfy (2.3.2)-(2.3.5).

The first event we introduce is the same as the event $\mathcal{C}_2$ in (2.3.12), except with larger gaps and jumps. That is, if a particle is more than $c_4 a_N$ away from the leader, then it does not jump to within distance $3c_3 a_N$ of the leader's position with a single big jump (recall that $c_3 \ll c_4$). We let

$$\mathcal{D}_1 := \left\{ \begin{array}{l} \nexists (i,b,s) \in [N] \times \{1,2\} \times [\![t_3, t-1]\!] \text{ such that} \\ X_{i,b,s} \in (Z_i(s) - 3c_3 a_N, Z_i(s) + 3c_3 a_N] \text{ and } Z_i(s) \geq c_4 a_N \end{array} \right\}, \tag{2.3.47}$$

where $Z_i(s)$ is the gap between the ith and the rightmost particle. The reason behind the definition of $\mathcal{D}_1$ is the following. Assume that a big jump beats the leader at a time when the diameter is fairly big ($> \frac{3}{2} c_4 a_N$). Then the event $\mathcal{D}_1$, together with the events $\mathcal{C}_3$, $\mathcal{C}_4$ and $\mathcal{C}_5$, implies that this particle must become the new leader and it will lead by at least $(3c_3 - \rho)a_N$, which will be enough to show that $\mathcal{C}_1$ occurs. We state this as a corollary below, which we will use later on in this section.

Corollary 2.3.10. *On the event $\mathcal{D}_1 \cap \mathcal{C}_3 \cap \mathcal{C}_4 \cap \mathcal{C}_5$, if $(k,b,s) \in B_N^{[t_3, t-1]}$, $d(\mathcal{X}(s)) \geq (c_4 + c_1) a_N$ and $X_{k,b,s} > Z_k(s)$, then*

$$\mathcal{X}_N(s+1) = \mathcal{X}_k(s) + X_{k,b,s} > \mathcal{X}_{N-1}(s+1) + (3c_3 - \rho) a_N,$$

where $Z_k(s), \mathcal{D}_1$ and $\mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_5$ are given by (2.3.11), (2.3.47) *and* (2.3.13)-(2.3.15) *respectively, and $B_N^{[t_3,t-1]}$ is defined in* (2.2.14).

Proof. The statement follows by exactly the same argument as for Corollary 2.3.7(a), if we replace $\mathcal{C}_2$ by $\mathcal{D}_1$, c_3 by c_4 and $2c_2$ by $3c_3$. □

The next two events will ensure that the record is broken in the time interval $[t_2+1, t_1]$. The first event says that there is a jump of size greater than $2c_4 a_N$ in every interval of length $c_5 \ell_N$ in $[t_3, t_1]$ (recall $c_4 \ll c_5$). We define

$$\mathcal{D}_2 := \{\forall s \in [\![t_3, t_1 - c_5 \ell_N]\!],\ \exists (k,b,\hat{s}) \in [N] \times \{1,2\} \times [\![s, s + c_5 \ell_N]\!] :\ X_{k,b,\hat{s}} > 2c_4 a_N\}. \tag{2.3.48}$$

The event $\mathcal{D}_2$ will be useful if at some point in the time interval $[t_2, t_1]$ the diameter is not too large ($\leq \frac{3}{2} c_4 a_N$). If $\mathcal{D}_2$ occurs then shortly after this point a jump of size larger than $2c_4 a_N$ happens. We will show that this jump breaks the record, and the particle performing this jump will lead by at least $2c_3 a_N$. The reason for this is that the jump size ($> 2c_4 a_N$) is much greater than the preceding diameter ($\leq \frac{3}{2} c_4 a_N$), and that $c_3 \ll c_4$.

The next event says that there will be a jump of size greater than $2c_6 a_N$ between times t_2 and $t_2 + \lceil \ell_N / 2 \rceil$ (recall $c_6 \gg c_5$). Let

$$\mathcal{D}_3 := \{\exists (i,b,s) \in [N] \times \{1,2\} \times [\![t_2, t_2 + \lceil \ell_N/2 \rceil]\!] :\ X_{i,b,s} > 2c_6 a_N\}. \tag{2.3.49}$$

The next event says that there is no jump of size greater than $c_6 a_N$ shortly before time t_2. We let

$$\mathcal{D}_4 := \{\nexists (i,b,s) \in [N] \times \{1,2\} \times [\![t_2 - \lceil c_5 \ell_N \rceil, t_2]\!] : X_{i,b,s} > c_6 a_N\} \tag{2.3.50}$$

(recall $c_5 \ll c_6$). Our last event excludes jumps of size in a certain small range in a certain short time interval. The starting point of this time interval will be the first time after t_2 when the diameter is at most $\frac{3}{2}c_4 a_N$:

$$\tau_2 := \inf\left\{s \geq t_2 : d(\mathcal{X}(s)) \leq \tfrac{3}{2}c_4 a_N\right\}, \tag{2.3.51}$$

and we define the event

$$\mathcal{D}_5 := \left\{ \begin{array}{l} \nexists (k,b,s) \in [N] \times \{1,2\} \times [\![\tau_2, \tau_2 + c_5 \ell_N]\!] : \\ X_{k,b,s} \in (2c_4 a_N, 2c_4 a_N + 3c_3 a_N] \end{array} \right\}. \tag{2.3.52}$$

We can now state the main result of this subsection.

Proposition 2.3.11. *Let $\eta \in (0,1]$, and assume that the constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ satisfy* (2.3.2)-(2.3.5). *For all $N \geq 2$ sufficiently large that $\ell_N - \lceil c_5 \ell_N \rceil \geq \lceil \ell_N / 2 \rceil$ and $t > 4\ell_N$,*

$$\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i \subseteq \mathcal{C}_1,$$

where $\mathcal{D}_1, \ldots, \mathcal{D}_5$ are defined in (2.3.47)-(2.3.50) *and* (2.3.52) *respectively, and $\mathcal{C}_1, \ldots, \mathcal{C}_7$ are defined in* (2.3.10) *and* (2.3.12)-(2.3.17) *respectively.*

Before giving a precise proof of Proposition 2.3.11, we give an outline of the argument, which is divided into four separate cases. Suppose $\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i$ occurs.

Case 1: Suppose there is a time $\tau_2 \in [t_2, t_1 - c_5 \ell_N]$ when the diameter is not too large (at most $\frac{3}{2}c_4 a_N$). Then shortly after time τ_2, there will be a jump of size larger than $2c_4 a_N$, by the definition of the event $\mathcal{D}_2$. We will show that the particle making this jump breaks the record and will lead by a distance larger than $2c_3 a_N$. The proof will also use the definition of the event $\mathcal{D}_5$.

Case 2(a): Suppose the diameter is larger than $\frac{3}{2}c_4 a_N$ at all times in $[t_2, t_1 - c_5 \ell_N]$, but the record is broken by a big jump at some point in this time interval. Then Corollary 2.3.10 tells us that there will be a gap of size greater than $2c_3 a_N$ behind the new record.

Case 2(b): Suppose the diameter is larger than $\frac{3}{2}c_4 a_N$ at all times in $[t_2, t_1 - c_5 \ell_N]$. If the record is not broken on the time interval $[t_2 - \lceil c_5 \ell_N \rceil, t_1 - c_5 \ell_N]$, then using Lemma 2.3.9, we can show that the diameter is less than $\frac{3}{2}c_4a_N$ at time $t_1 - \lceil c_5\ell_N \rceil$, giving us a contradiction. Thus this case is impossible.

Case 2(c): Suppose the diameter is larger than $\frac{3}{2}c_4a_N$ at all times in $[t_2, t_1 - c_5\ell_N]$. Now consider the case that the record is not broken on the time interval $[t_2, t_1 - c_5\ell_N]$, but is broken shortly before t_2, during the time interval $[t_2 - \lceil c_5\ell_N \rceil, t_2 - 1]$. By the definition of the event $\mathcal{D}_4$, this jump cannot be very big. Therefore, we will see that the new leader will be beaten by the first jump of size greater than $2c_6a_N$, if the record has not already been broken before that. There will be a jump of size greater than $2c_6a_N$ before time $t_2 + \lceil \ell_N/2 \rceil$ because of the event $\mathcal{D}_3$, so the record must be broken by a big jump before time $t_1 - c_5\ell_N$. This again gives us a contradiction, meaning that Case 2(c) is also impossible.

We now prove Proposition 2.3.11, using cases 1, 2(a), 2(b) and 2(c) as described above.

Proof of Proposition 2.3.11. Fix $\eta \in (0, 1]$ and take constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ as in (2.3.2)-(2.3.5). Let us assume that $\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i$ occurs.

Case 1: $t_2 \leq \tau_2 \leq t_1 - c_5\ell_N$.
In this case, by the definition of τ_2 we have

$$d(\mathcal{X}(\tau_2)) \leq \tfrac{3}{2}c_4a_N. \tag{2.3.53}$$

Let us now consider the first jump of size greater than $2c_4a_N$ after time τ_2; that is, let

$$s^* = \inf\{s \geq \tau_2 : \exists (k, b) \in [N] \times \{1, 2\} \text{ such that } X_{k,b,s} > 2c_4a_N\} \in [\![\tau_2, \tau_2 + c_5\ell_N]\!] \tag{2.3.54}$$

by the definition of the event $\mathcal{D}_2$ in (2.3.48). Take $(k^*, b^*) \in [N] \times \{1, 2\}$ such that $X_{k^*,b^*,s^*} > 2c_4a_N$ (there is a unique choice of the pair (k^*, b^*) by the definition of the event $\mathcal{C}_5$). We will show that the jump X_{k^*,b^*,s^*} creates a gap of size larger than $2c_3a_N$ behind the leader. We do this in two steps. First we show that the diameter is not too large right before the jump X_{k^*,b^*,s^*} occurs; then we show that a gap is created.

(i) We claim that

$$d(\mathcal{X}(s^*)) \leq 2c_4a_N + c_2a_N. \tag{2.3.55}$$

Now we prove the claim. By (2.3.53), the claim holds if $s^* = \tau_2$. Suppose on the other hand that $s^* > \tau_2$. Let $j \in [N]$ be arbitrary, and then take $i \in [N]$ such that $(i, \tau_2) \lesssim (j, s^*)$. We will show that particle (j, s^*) is within distance $(2c_4 + c_2)a_N$ of the leftmost particle at time s^*. We consider two cases, depending on whether there is a big jump on the path between $\mathcal{X}_i(\tau_2)$ and $\mathcal{X}_j(s^*)$.

- If $B_N \cap P_{i,\tau_2}^{j,s^*} = \emptyset$, then by Lemma 2.3.3, (2.3.53) and monotonicity,

$$\mathcal{X}_j(s^*) \le \mathcal{X}_i(\tau_2) + c_1 a_N \le \mathcal{X}_1(\tau_2) + \tfrac{3}{2}c_4 a_N + c_1 a_N \le \mathcal{X}_1(s^*) + \tfrac{3}{2}c_4 a_N + c_1 a_N. \tag{2.3.56}$$

- If $B_N \cap P_{i,\tau_2}^{j,s^*} \neq \emptyset$, then take $(k', b', s') \in B_N \cap P_{i,\tau_2}^{j,s^*}$. Then $\mathcal{X}_{k'}(s')$ is the position of the parent of the particle that makes the jump $X_{k',b',s'}$. Since (by (2.3.54)) X_{k^*,b^*,s^*} is the first jump of size greater than $2c_4 a_N$ after time τ_2, and since $s' < s^*$, we have $X_{k',b',s'} \le 2c_4 a_N$. Then since $s^* - s' \le s^* - \tau_2 \le c_5 \ell_N$, by Lemma 2.3.4 we have

$$\mathcal{X}_j(s^*) \le \mathcal{X}_{k'}(s') + X_{k',b',s'} + c_1 a_N \le \mathcal{X}_{k'}(s') + 2c_4 a_N + c_1 a_N.$$

  Now Lemma 2.3.6 and monotonicity imply that this is at most

$$\mathcal{X}_1(s') + 2c_4 a_N + 2c_1 a_N \le \mathcal{X}_1(s^*) + 2c_4 a_N + 2c_1 a_N. \tag{2.3.57}$$

By (2.3.56), (2.3.57) and our choice of constants in (2.3.4), we conclude that for any particle position $\mathcal{X}_j(s^*)$ in the population at time s^*, $\mathcal{X}_j(s^*) \le \mathcal{X}_1(s^*) + 2c_4 a_N + c_2 a_N$, which implies (2.3.55).

(ii) We claim that

$$\mathcal{X}_{N-1}(s^*+1) + 2c_3 a_N < \mathcal{X}_N(s^*+1). \tag{2.3.58}$$

By the definition of (k^*, b^*, s^*), we have $X_{k^*,b^*,s^*} > 2c_4 a_N$, and we also know that $X_{k^*,b^*,s^*} \notin (2c_4 a_N, 2c_4 a_N + 3c_3 a_N]$ by the definition of the event $\mathcal{D}_5$, because $s^* \in [\![\tau_2, \tau_2 + c_5 \ell_N]\!]$. Therefore we have

$$X_{k^*,b^*,s^*} > 2c_4 a_N + 3c_3 a_N. \tag{2.3.59}$$

Then by (2.3.59) and (2.3.55),

$$\mathcal{X}_{k^*}(s^*) + X_{k^*,b^*,s^*} > \mathcal{X}_1(s^*) + (2c_4 + 3c_3) a_N \ge \mathcal{X}_N(s^*) + (3c_3 - c_2) a_N. \tag{2.3.60}$$

Note that $3c_3 - c_2 > \rho$ by (2.3.3)-(2.3.4), which in particular shows that X_{k^*,b^*,s^*} must be a big jump. Hence by (2.3.60) and Lemma 2.3.5(b), we have $(k^*, s^*) \lesssim_{b^*} (N, s^*+1)$ and

$$\mathcal{X}_N(s^*+1) > \mathcal{X}_{N-1}(s^*+1) + (3c_3 - c_2 - \rho) a_N,$$

which is larger than $\mathcal{X}_{N-1}(s^*+1) + 2c_3 a_N$ by (2.3.3)-(2.3.4). This finishes the proof of (2.3.58).

Recall from (2.3.54) that $s^* \in [\![\tau_2, \tau_2 + c_5 \ell_N]\!]$. Furthermore, event $\mathcal{C}_6$ tells us that $s^* \notin [t_1 - \lceil \delta \ell_N \rceil, t_1]$. Therefore, by the assumption of Case 1 that $\tau_2 \in [t_2, t_1 - c_5 \ell_N]$, we conclude $t_2+1 \le s^*+1 \le t_1$, which together with (2.3.58) shows that $\mathcal{C}_1$ occurs. We conclude that Proposition 2.3.11 holds in Case 1.

Case 2(a): $\tau_2 > t_1 - c_5\ell_N$ and $[t_2, t_1 - c_5\ell_N] \cap \hat{\mathbf{S}}_N \neq \emptyset$, where $\hat{\mathbf{S}}_N$ is defined in (2.2.18). This means that there exists $(\hat{k}, \hat{b}, \hat{s}) \in B_N^{[t_2, t_1 - c_5\ell_N]}$ with $X_{\hat{k},\hat{b},\hat{s}} > Z_{\hat{k}}(\hat{s})$ (recall (2.3.11)). Since $\tau_2 > t_1 - c_5\ell_N$, we have $d(\mathcal{X}(\hat{s})) > \frac{3}{2}c_4 a_N$. Then by (2.3.4), we can apply Corollary 2.3.10 to obtain

$$\mathcal{X}_N(\hat{s}+1) = \mathcal{X}_{\hat{k}}(\hat{s}) + X_{\hat{k},\hat{b},\hat{s}} > \mathcal{X}_{N-1}(\hat{s}+1) + (3c_3 - \rho)a_N.$$

By our choice of constants in (2.3.3)-(2.3.4), and because $\hat{s}+1 \in [\![t_2+1, t_1]\!]$, this shows that $\mathcal{C}_1$ occurs. Therefore we are done with the proof of Proposition 2.3.11 in Case 2(a).

Case 2(b): $\tau_2 > t_1 - c_5\ell_N$ and $[t_2 - \lceil c_5\ell_N \rceil, t_1 - c_5\ell_N] \cap \hat{\mathbf{S}}_N = \emptyset$.
We will apply Lemma 2.3.9 with $s = t_2 - \lceil c_5\ell_N \rceil$ and $\Delta s = \ell_N$. By assumption we have $[s, s+\Delta s - 1] \subseteq \hat{\mathbf{S}}_N^c$, and therefore applying either part (a) or part (b) of Lemma 2.3.9 as appropriate, we have

$$d(\mathcal{X}(s+\Delta s)) = d(\mathcal{X}(t_1 - \lceil c_5\ell_N \rceil)) \le \max\left\{c_1 a_N, \tfrac{3}{2}c_3 a_N + 2c_1 a_N\right\}$$

which is smaller than $\frac{3}{2}c_4 a_N$ by (2.3.4), contradicting the assumption that $\tau_2 > t_1 - c_5\ell_N$. This shows that Case 2(b) cannot occur.

Case 2(c): $\tau_2 > t_1 - c_5\ell_N$ and $[t_2, t_1 - c_5\ell_N] \cap \hat{\mathbf{S}}_N = \emptyset$, but $[t_2 - \lceil c_5\ell_N \rceil, t_2 - 1] \cap \hat{\mathbf{S}}_N \neq \emptyset$. Define

$$\tau_3 := \inf\left\{s \le t_2 : [\![s, t_2]\!] \subseteq \hat{\mathbf{S}}_N^c\right\} \in (t_2 - \lceil c_5\ell_N \rceil, t_2]. \tag{2.3.61}$$

Suppose, aiming for a contradiction, that there exists $r \in [\![\tau_3, t_1 - c_5\ell_N]\!]$ such that $d(\mathcal{X}(r)) \le \frac{3}{2}c_3 a_N$. Then since $[\![\tau_3, t_2]\!] \subseteq \hat{\mathbf{S}}_N^c$ and $[\![t_2, t_1 - c_5\ell_N]\!] \subseteq \hat{\mathbf{S}}_N^c$, Lemma 2.3.9(b) applies with $s = r$ and $\Delta s = t_1 - \lceil c_5\ell_N \rceil - r$ (which is smaller than ℓ_N since $r \ge \tau_3 > t_2 - \lceil c_5\ell_N \rceil$), and says that $d(\mathcal{X}(t_1 - \lceil c_5\ell_N \rceil)) \le \frac{3}{2}c_3 a_N + 2c_1 a_N$. By (2.3.4), this contradicts the assumption that $\tau_2 > t_1 - c_5\ell_N$. Thus we must have

$$d(\mathcal{X}(r)) \ge \tfrac{3}{2}c_3 a_N \quad \forall r \in [\![\tau_3, t_1 - c_5\ell_N]\!]. \tag{2.3.62}$$

Now note that $\tau_3 - 1 \in \hat{\mathbf{S}}_N$. Then by (2.3.62), the second equivalence in Corollary 2.3.8 implies that in fact $\tau_3 - 1 \in \mathbf{S}_N$. Hence, by the definition of $\mathbf{S}_N$ in (2.2.16), there exists $(k, b) \in [N] \times \{1, 2\}$ such that

$$\mathcal{X}_N(\tau_3) = \mathcal{X}_k(\tau_3 - 1) + X_{k,b,\tau_3-1}, \tag{2.3.63}$$

where $X_{k,b,\tau_3-1} > \rho a_N$. Now Lemma 2.3.6 provides a bound on $\mathcal{X}_k(\tau_3 - 1)$, and the definition of $\mathcal{D}_4$ together with the fact that $\tau_3 - 1 \in [t_2 - \lceil c_5 \ell_N \rceil, t_2]$ gives us a bound on X_{k,b,τ_3-1}, so that we obtain

$$\mathcal{X}_N(\tau_3) \leq \mathcal{X}_1(\tau_3 - 1) + (c_1 + c_6)a_N. \tag{2.3.64}$$

Now, on the event $\mathcal{D}_3$, there exists $(\tilde{i}, \tilde{b}, \tilde{s}) \in [N] \times \{1,2\} \times [\![t_2, t_2 + \lceil \ell_N/2 \rceil]\!]$ such that

$$X_{\tilde{i},\tilde{b},\tilde{s}} > 2c_6 a_N > \rho a_N \tag{2.3.65}$$

by (2.3.3)-(2.3.4). We show that the particle performing this big jump beats the leader at time $\tilde{s}$. By our assumption that $\ell_N - \lceil c_5 \ell_N \rceil \geq \lceil \ell_N/2 \rceil$ and by (2.3.61), we have $[\![\tau_3, \tilde{s}]\!] \subseteq \hat{\mathbf{S}}_N^c$ and $\tilde{s} - \tau_3 \leq \ell_N$. Therefore, by (2.3.62) we can apply Lemma 2.3.9(a) with $s = \tau_3$ and $\Delta s = \tilde{s} - \tau_3$, and then by (2.3.64) we have

$$\mathcal{X}_N(\tilde{s}) \leq \mathcal{X}_N(\tau_3) + c_1 a_N \leq \mathcal{X}_1(\tau_3 - 1) + (2c_1 + c_6)a_N. \tag{2.3.66}$$

By (2.3.4), it follows that

$$\mathcal{X}_N(\tilde{s}) < \mathcal{X}_1(\tau_3 - 1) + 2c_6 a_N < \mathcal{X}_1(\tilde{s}) + X_{\tilde{i},\tilde{b},\tilde{s}} \leq \mathcal{X}_{\tilde{i}}(\tilde{s}) + X_{\tilde{i},\tilde{b},\tilde{s}},$$

where in the second inequality we use monotonicity and (2.3.65). Therefore, by the assumptions that $\tilde{s} \in [\![t_2, t_2 + \lceil \ell_N/2 \rceil]\!]$ and $\ell_N - \lceil c_5 \ell_N \rceil \geq \lceil \ell_N/2 \rceil$, and by the definition of $\hat{\mathbf{S}}_N$ in (2.2.18), we have $\tilde{s} \in \hat{\mathbf{S}}_N \cap [t_2, t_1 - c_5 \ell_N]$, which contradicts the assumption of Case 2(c).

We have now shown that if $\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i$ occurs then Cases 2(b) and 2(c) are impossible, whereas Cases 1 and 2(a) imply that $\mathcal{C}_1$ must occur. This concludes the proof of Proposition 2.3.11. □

2.4 Probabilities of the events from the deterministic argument

In the deterministic argument in Section 2.3 we have provided a strategy which ensures that the events $\mathcal{A}_1$ and $\mathcal{A}_3$ occur. In this section we check that the events $\mathcal{C}_2$ to $\mathcal{C}_7$ and $\mathcal{D}_1$ to $\mathcal{D}_5$ which make up this strategy all occur with high probability, and use this to finish the proof of Proposition 2.2.5.

When bounding the probabilities of these events, it will be useful to consider branching random walks (BRWs) without selection, where at each time step all particles have two offspring, the offspring particles make i.i.d. jumps from their parents' locations, and every offspring particle survives. Below we describe a construction of the N-BRW from N independent BRWs, which will allow us to consider our events on the probability space on which the BRWs are defined. (A similar construction was used in [2].)

2.4.1 Construction of the N-BRW from N independent BRWs

Consider a binary tree with the following labelling. Let

$$\mathcal{U}_0 := \bigcup_{n=0}^{\infty} \{1,2\}^n ,$$

and for convenience we write e.g. 121 instead of $(1,2,1)$. Then the root of the binary tree has label $\emptyset$, and for all $u \in \mathcal{U}_0$ the two children of vertex u have labels $u1$ and $u2$. We will use the partial order $\preceq$ on the set $\mathcal{U}_0$; we write $u \preceq v$ if either $u = v$ or the vertex with label u is an ancestor of the vertex with label v in the binary tree. We also write $u \prec v$ if $u \preceq v$ and $u \neq v$.

The particles of the N independent BRWs will have labels from the set $[N] \times \mathcal{U}_0$, and we have a lexicographical order on the set of labels. We also let $\mathcal{U} := \mathcal{U}_0 \setminus \{\emptyset\}$. The jumps of the BRWs will be given by random variables $(Y_{j,u})_{j\in[N],u\in\mathcal{U}}$, which are i.i.d. with common law given by (2.1.3).

The N initial particles of the N independent BRWs are labelled with the pairs $(j,\emptyset)$ with $j \in [N]$. For each $j \in [N]$, we let $\mathcal{Y}_j(\emptyset) \in \mathbb{R}$ be the initial location of particle $(j,\emptyset)$. Then, at each time step $n \in \mathbb{N}_0$, each particle (j,u) with $j \in [N]$ and $u \in \{1,2\}^n$ has two offspring labelled $(j,u1)$ and $(j,u2)$, which make jumps $Y_{j,u1}$, $Y_{j,u2}$ from the location $\mathcal{Y}_j(u)$. The locations of the offspring particles $(j,u1)$ and $(j,u2)$ will be $\mathcal{Y}_j(u1) = \mathcal{Y}_j(u) + Y_{j,u1}$ and $\mathcal{Y}_j(u2) = \mathcal{Y}_j(u) + Y_{j,u2}$. Note that for $u \prec v$, the path between particles (j,u) and (j,v) is given by the jumps $Y_{j,w}$ with $u \prec w \preceq v$, i.e. $\mathcal{Y}_j(v) - \mathcal{Y}_j(u) = \sum_{u\prec w\preceq v} Y_{j,w}$.

Now we construct the N-BRW by defining the surviving set of particles for each time $n \in \mathbb{N}_0$ as the N-element set $H_n \subseteq [N] \times \{1,2\}^n$, constructed iteratively as follows. Let $H_0 := \{(1,\emptyset),\dots,(N,\emptyset)\}$. Given H_n for some $n \in \mathbb{N}_0$, we let H'_n denote the set of offspring of the particles in the set H_n:

$$H'_n := \bigcup_{(j,u)\in H_n} \{(j,u1),(j,u2)\} .$$

Then $H_{n+1} \subseteq H'_n$ consists of the particles with the N largest values in the collection $(\mathcal{Y}_j(u))_{(j,u)\in H'_n}$, where ties are broken based on the lexicographical order of the labels. In this way an N-BRW is constructed from the initial configuration $(\mathcal{Y}_j(\emptyset))_{j\in[N]}$ and the jumps $(Y_{j,u})_{j\in[N],u\in\mathcal{U}}$.

For $n \in \mathbb{N}$, we let $\mathcal{F}'_n$ denote the σ-algebra generated by $(Y_{j,u})_{j\in[N],u\in\cup_{m=1}^{n}\{1,2\}^m}$. Note that H_n is $\mathcal{F}'_n$-measurable for each n.

Returning to our original notation in Section 2.2.1, we can say the following. For all $n \in \mathbb{N}_0$, let $\mathcal{X}(n)$ denote the ordered set which contains the values $(\mathcal{Y}_j(u))_{(j,u)\in H_n}$ in ascending order:

$$\mathcal{X}(n) = \{\mathcal{X}_1(n) \le \dots \le \mathcal{X}_N(n)\} := \{\mathcal{Y}_{j_1}(u_1) \le \dots \le \mathcal{Y}_{j_N}(u_N)\} , \tag{2.4.1}$$

where $H_n = \{(j_i, u_i) : i \in [N]\}$, and again ties are broken based on the lexicographical order of the labels. Then we define the map σ which associates the pair $(i,n) \in [N] \times \mathbb{N}_0$ with particle $(j_i, u_i) \in H_n$, where $\mathcal{Y}_{j_i}(u_i)$ has the ith position in the ordered set $\mathcal{X}(n)$. That is, for $(i,n) \in [N] \times \mathbb{N}_0$ we let

$$\sigma(i,n) = (j_i, u_i) \in H_n \subset [N] \times \mathcal{U}_0, \tag{2.4.2}$$

where (j_i, u_i) is as in (2.4.1). The jumps in our original notation are then given by

$$X_{i,1,n} := Y_{j_i,u_i 1} \quad \text{and} \quad X_{i,2,n} := Y_{j_i,u_i 2}, \tag{2.4.3}$$

if $\sigma(i,n) = (j_i, u_i)$.

Finally, recall that we introduced the partial order $\lesssim$ in (2.2.8) in Section 2.2.4 to denote that two particles are related in the N-BRW. This partial order corresponds to the partial order $\preceq$ in the N independent BRWs as follows. For all $n, k \in \mathbb{N}_0$ and $i_0, i_k \in [N]$, we have $(i_0, n) \lesssim (i_k, n+k)$ if and only if for some $j \in [N]$ and $u, v \in \mathcal{U}_0$, we have $\sigma(i_0, n) = (j, u)$, $\sigma(i_k, n+k) = (j, v)$, and $u \preceq v$. Furthermore, for $b \in \{1,2\}$ we have $(i_0, n) \lesssim_b (i_k, n+k)$ if and only if the above holds and additionally $k \geq 1$ and $ub \preceq v$.

Now we can consider the N-BRW constructed from N independent BRWs with the notation introduced in Sections 2.2.1 and 2.2.4. It follows from our construction that for any path in the N-BRW, there is a path in one of the N independent BRWs that consists of the same sequence of jumps as the path in the N-BRW. We state and prove this simple property below. Recall the notation $P^{i_2,n_2}_{i_1,n_1}$ from (2.2.10).

Lemma 2.4.1. *For all $k \in \mathbb{N}$, $i_0, i_k \in [N]$ and $n \in \mathbb{N}_0$, if $(i_0, n) \lesssim (i_k, n+k)$ with $P^{i_k,n+k}_{i_0,n} = \{(i_l, b_l, n+l) : l \in \{0, \ldots, k-1\}\}$, then there exists $j \in [N]$ and $(u_l)_{l=0}^k \subseteq \mathcal{U}_0$ such that*

(1) $(j, u_l) \in H_{n+l}$, for all $l \in \{0, \ldots, k\}$,

(2) $u_l b_l \preceq u_k$, for all $l \in \{0, \ldots, k-1\}$, and

(3) $X_{i_l,b_l,n+l} = Y_{j,u_l b_l}$, for all $l \in \{0, \ldots, k-1\}$.

Proof. Take $(i_l, b_l, n+l) \in P^{i_k,n+k}_{i_0,n}$ (with $l \in \{0, \ldots, k-1\}$). Then $(i_l, n+l) \lesssim_{b_l} (i_k, n+k)$. Thus, there exist $j \in [N]$ and $u_l, u_k \in \mathcal{U}_0$ such that $\sigma(i_l, n+l) = (j, u_l)$, $\sigma(i_k, n+k) = (j, u_k)$, and $u_l b_l \preceq u_k$. This implies $X_{i_l,b_l,n+l} = Y_{j,u_l b_l}$ (see (2.4.3)) and also $(j, u_l) \in H_{n+l}$ and $(j, u_k) \in H_{n+k}$ by the definition (2.4.2) of σ. Since $(i_l, b_l, n+l) \in P^{i_k,n+k}_{i_0,n}$ was arbitrary, the result follows. □

2.4.2 Paths with regularly varying jump distribution

One of the most important components of the deterministic argument in Section 2.3 is that paths cannot move very far without big jumps; this is the meaning of the event $\mathcal{C}_4$ defined in (2.3.14). Corollary 2.4.5 is the main result of this section and will be used to bound from below the probability that the event $\mathcal{C}_4$ occurs.

As in [3], we use Potter's bounds to give useful estimates on the regularly varying function h (with index α) defined in (2.1.3). We will use the following elementary consequence of Potter's bounds.

Lemma 2.4.2. *For $\epsilon > 0$, there exist $B(\epsilon) > 1$ and $C_1(\epsilon), C_2(\epsilon) > 0$ such that*

$$\frac{1}{h(x)} \leq C_1 x^{\epsilon-\alpha} \quad \text{and} \quad h(x) \leq C_2 x^{\alpha+\epsilon} \quad \forall x \geq B.$$

Proof. Let $\epsilon > 0$ be arbitrary. By Potter's bounds [9, Theorem 1.5.6(iii)], there exists $x_0 > 0$ depending only on ϵ such that

$$\frac{h(y)}{h(x)} \leq 2 \max\left((y/x)^{\alpha+\epsilon}, (y/x)^{\alpha-\epsilon}\right) \quad \forall x, y \geq x_0. \tag{2.4.4}$$

Let $x \geq x_0$ be arbitrary and let $y = x_0$ in (2.4.4). Then we have $y/x \leq 1$ and so $(y/x)^{\alpha+\epsilon} \leq (y/x)^{\alpha-\epsilon}$, and the first inequality in the statement of the lemma holds with $C_1 = 2x_0^{\alpha-\epsilon} h(x_0)^{-1}$ and $B = x_0 + 1$. Similarly, since we have $x/y \geq 1$, we have $(x/y)^{\alpha-\epsilon} \leq (x/y)^{\alpha+\epsilon}$, and hence by (2.4.4) (with x and y exchanged) the second inequality holds with $C_2 = 2h(x_0) x_0^{-(\alpha+\epsilon)}$ and $B = x_0 + 1$. □

In order to show that $\mathcal{C}_4$ occurs with high probability, we prove a lemma about a random walk with the same jump distribution as our N-BRW, but in which jumps larger than a certain size are discarded and count as a jump of size zero. The lemma gives an upper bound on the probability that this random walk moves a large distance x_N in of order ℓ_N steps, if the jumps larger than rx_N are discarded (for some $r \in (0,1)$). For an arbitrarily large $q > 0$, the parameter r can be taken sufficiently small that the above probability is smaller than N^{-q} (for large N). Our lemma is similar to the lemma on page 168 of [21], where the jump distribution is truncated; jumps greater than a threshold value are not allowed at all, instead of being counted as zero. We use ideas from the proof of Theorem 3 in [27], which is a large deviation result for sums of random variables with stretched exponential tails.

Recall that $\mathbb{P}(X > x) = h(x)^{-1}$ for $x \geq 0$, where h is regularly varying with index $\alpha > 0$.

Lemma 2.4.3. *Let $X_1, X_2, \dots$ be i.i.d. random variables with $X_1 \stackrel{d}{=} X$. For any $m \in \mathbb{N}$, $q > 0$, $\lambda > 0$, $0 < r < 1 \wedge \frac{\lambda(1\wedge\alpha)}{8q}$, for N sufficiently large, if $x_N > N^\lambda$ then*

$$\mathbb{P}\left(\sum_{j=1}^{m\ell_N} X_j \mathbb{1}_{\{X_j \leq rx_N\}} \geq x_N\right) \leq N^{-q}.$$

Before proving Lemma 2.4.3, we now state and prove an elementary identity which will be used in the proof. This identity was also used in the proof of Theorem 3 in [27].

Lemma 2.4.4. *Suppose Y is a non-negative random variable. For $v > 0$ and $0 < K_1 < K_2 < \infty$,*

$$\mathbb{E}[\exp(vY\mathbb{1}_{\{Y\leq K_2\}})\mathbb{1}_{\{Y\geq K_1\}}] \\ = \int_{K_1}^{K_2} ve^{vu}\mathbb{P}(Y > u)du + e^{vK_1}\mathbb{P}(Y \geq K_1) - (e^{vK_2} - 1)\mathbb{P}(Y > K_2). \quad (2.4.5)$$

Proof. First note that the random variable in the expectation on the left-hand side of (2.4.5) takes the value 1 if $Y > K_2$. The expectation can be written as

$$\mathbb{E}[\exp(vY\mathbb{1}_{\{Y\leq K_2\}})\mathbb{1}_{\{Y\geq K_1\}}] = \mathbb{E}\left[e^{vY}\mathbb{1}_{\{K_1\leq Y\leq K_2\}}\right] + \mathbb{P}(Y > K_2). \quad (2.4.6)$$

Now we will work on the integral on the right-hand side of (2.4.5). First, by Fubini's theorem we have

$$\int_{K_1}^{K_2} ve^{vu}\mathbb{P}(Y > u)du = \mathbb{E}\left[\int_{K_1}^{K_2} ve^{vu}\mathbb{1}_{\{Y>u\}}du\right] = \mathbb{E}\left[\int_{K_1}^{K_2\wedge Y} ve^{vu}du\mathbb{1}_{\{Y\geq K_1\}}\right].$$

By calculating the integral, it follows that

$$\int_{K_1}^{K_2} ve^{vu}\mathbb{P}(Y > u)du = \mathbb{E}\left[\left(e^{v(K_2\wedge Y)} - e^{vK_1}\right)\mathbb{1}_{\{Y\geq K_1\}}\right] \\ = \mathbb{E}\left[e^{vY}\mathbb{1}_{\{K_1\leq Y\leq K_2\}}\right] + \mathbb{E}\left[e^{vK_2}\mathbb{1}_{\{Y>K_2\}}\right] - \mathbb{E}\left[e^{vK_1}\mathbb{1}_{\{Y\geq K_1\}}\right].$$

The result follows by (2.4.6). □

Proof of Lemma 2.4.3. Let $\tilde{X} := X\mathbb{1}_{\{X\leq rx_N\}}$ and $\tilde{X}_j := X_j\mathbb{1}_{\{X_j\leq rx_N\}}$ for all $j \in \mathbb{N}$. Take N sufficiently large that $\ell_N \leq 2\log_2 N$. Then by Markov's inequality and since $\tilde{X}_1, \tilde{X}_2, \ldots$ are i.i.d. with $\tilde{X}_1 \overset{d}{=} \tilde{X}$, for $c > 0$,

$$\mathbb{P}\left(\sum_{j=1}^{m\ell_N} \tilde{X}_j \geq x_N\right) = \mathbb{P}\left(\exp\left(c\ell_N x_N^{-1}\sum_{j=1}^{m\ell_N}\tilde{X}_j\right) \geq e^{c\ell_N}\right) \\ \leq e^{-c\ell_N}\mathbb{E}\left[e^{c\ell_N x_N^{-1}\tilde{X}}\right]^{m\ell_N} \\ \leq N^{-\frac{c}{\log 2}+\frac{2m}{\log 2}\log\mathbb{E}\left[e^{c\ell_N x_N^{-1}\tilde{X}}\right]}, \quad (2.4.7)$$

since $\log_2 N \leq \ell_N \leq 2\log_2 N$. We will show that with an appropriate choice of $c > 0$, for N sufficiently large, the right-hand side of (2.4.7) is smaller than N^{-q}. First we require

$$c > 2q\log 2. \quad (2.4.8)$$

Second, we will have another condition on c which ensures that $\mathbb{E}[e^{c\ell_N x_N^{-1}\tilde{X}}] \leq 1 + O(N^{-\epsilon})$ as $N \to \infty$ for some $\epsilon > 0$. We now estimate this expectation and determine the choice of

c.

Take $0 < \epsilon < \frac{\lambda(1\wedge\alpha)}{2(\lambda+1)}$, and take $B = B(\epsilon) > 1$ and $C_1 = C_1(\epsilon) > 0$ as in Lemma 2.4.2. Suppose N is sufficiently large that $rx_N > B$. We apply Lemma 2.4.4 with $Y = X$, $v = c\ell_N x_N^{-1}$, $K_1 = B$ and $K_2 = rx_N$, and then use (2.1.3), to obtain

$$\begin{aligned}\mathbb{E}\left[e^{c\ell_N x_N^{-1}\tilde{X}}\right] &\leq \mathbb{E}\left[e^{c\ell_N x_N^{-1}X\mathbb{1}_{\{X\leq rx_N\}}}\mathbb{1}_{\{X\geq B\}}\right] + e^{Bc\ell_N x_N^{-1}}\mathbb{P}(X<B)\\ &\leq \int_B^{rx_N} c\ell_N x_N^{-1} e^{c\ell_N x_N^{-1}u} h(u)^{-1} du + e^{Bc\ell_N x_N^{-1}}. \qquad (2.4.9)\end{aligned}$$

We will choose c such that the first term on the right-hand side of (2.4.9) is close to zero. By Lemma 2.4.2, and then since $r < 1$, we have

$$\begin{aligned}\int_B^{rx_N} c\ell_N x_N^{-1} e^{c\ell_N x_N^{-1}u} h(u)^{-1} du &\leq \int_B^{rx_N} C_1 c\ell_N x_N^{-1} e^{c\ell_N x_N^{-1}u} u^{-\alpha+\epsilon} du\\ &\leq C_1 c\ell_N x_N^{-1}\int_B^{rx_N} e^{c\ell_N x_N^{-1}(rx_N)} x_N^{\epsilon} u^{-\alpha} du.\end{aligned}$$

Integrating the right-hand side, since we took N sufficiently large that $\ell_N \leq 2\log_2 N$, we conclude

$$\int_B^{rx_N} c\ell_N x_N^{-1} e^{c\ell_N x_N^{-1}u} h(u)^{-1} du \leq \begin{cases} \frac{C_1 c}{1-\alpha}\ell_N N^{\frac{2cr}{\log 2}}\left(r^{1-\alpha}x_N^{\epsilon-\alpha} - B^{1-\alpha}x_N^{\epsilon-1}\right), & \text{if } \alpha \neq 1,\\ C_1 c\ell_N x_N^{\epsilon-1} N^{\frac{2cr}{\log 2}}\log x_N, & \text{if } \alpha = 1,\end{cases} \qquad (2.4.10)$$

where in the $\alpha = 1$ case we use that $B > 1$ and that $r < 1$.

Now, since $x_N > N^\lambda$ and $\epsilon < 1\wedge\alpha$, the right-hand side of (2.4.10) is at most of order $N^{-\epsilon}$ if

$$\frac{2cr}{\log 2} + \lambda(\epsilon - (1\wedge\alpha)) < -\epsilon. \qquad (2.4.11)$$

Since $r < \frac{\lambda(1\wedge\alpha)}{8q}$ by the assumptions of the lemma, we can find c such that

$$2q\log 2 < c < \frac{\lambda(1\wedge\alpha)\log 2}{4r}.$$

Then since we chose $\epsilon < \frac{\lambda(1\wedge\alpha)}{2(\lambda+1)}$, c satisfies (2.4.8) and (2.4.11). Note furthermore that since $x_N > N^\lambda$, the second term on the right-hand side of (2.4.9) is close to 1 for N large; for N sufficiently large we have

$$e^{Bc\ell_N x_N^{-1}} \leq e^{Bc\ell_N N^{-\lambda}} \leq 1 + 2Bc\ell_N N^{-\lambda}. \qquad (2.4.12)$$

Hence, (2.4.9), (2.4.10) and the choice of c, and (2.4.12) with the fact that $\epsilon < \lambda$ show that there exists a constant $A > 0$ such that

$$\mathbb{E}\left[e^{c\ell_N x_N^{-1}\tilde{X}}\right] \leq 1 + AN^{-\epsilon}$$

for N sufficiently large and $x_N > N^\lambda$. Therefore, by (2.4.7) and (2.4.8) we have

$$\mathbb{P}\left(\sum_{j=1}^{m\ell_N} \tilde{X}_j \geq x_N\right) \leq N^{-2q+\frac{2m}{\log 2}\log(1+AN^{-\epsilon})} \leq N^{-2q+\frac{2m}{\log 2}AN^{-\epsilon}} < N^{-q},$$

for N sufficiently large, which concludes the proof. $\square$

We now apply Lemma 2.4.3 to the N-BRW, to give us a convenient form of the result which we will use later in this section and also in Section 2.5.

Corollary 2.4.5. *Let $\lambda > 0$ and $0 < r < 1 \wedge \frac{\lambda(1\wedge\alpha)}{48}$. Then there exists $C > 0$ such that for N sufficiently large, if $x_N > N^\lambda$,*

$$\mathbb{P}\left(\begin{array}{l} \exists (k_1, s_1) \in [N] \times [\![t_4, t-1]\!],\ s_2 \in [\![s_1+1, t]\!] \text{ and } k_2 \in \mathcal{N}_{k_1,s_1}(s_2): \\ \sum_{(i,b,s)\in P_{k_1,s_1}^{k_2,s_2}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\leq rx_N\}} \geq x_N \end{array}\right) \leq CN^{-1},$$

where $P_{k_1,s_1}^{k_2,s_2}$ and $\mathcal{N}_{k_1,s_1}(s_2)$ are defined in (2.2.10) *and* (2.2.12) *respectively.*

Proof. Take $(k_1, s_1), (k_2, s_2) \in [N] \times [\![t_4, t-1]\!]$ with $(k_1, s_1) \lesssim (k_2, s_2)$, and let $k' = \zeta_{k_1,s_1}(t_4)$ be the index of the time-t_4 ancestor of (k_1, s_1) (see (2.2.9) for the notation). If the path between particles (k_1, s_1) and (k_2, s_2) moves at least x_N even with discarding jumps greater than rx_N, then the path between (k', t_4) and (k_2, s_2) does the same, because all jumps are non-negative. Therefore we only need to consider paths starting with the N particles of the population at time t_4:

$$\begin{aligned} &\mathbb{P}\left(\begin{array}{l} \exists (k_1, s_1) \in [N] \times [\![t_4, t-1]\!],\ s_2 \in [\![s_1+1, t]\!] \text{ and } k_2 \in \mathcal{N}_{k_1,s_1}(s_2): \\ \sum_{(i,b,s)\in P_{k_1,s_1}^{k_2,s_2}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\leq rx_N\}} \geq x_N \end{array}\right) \\ &\qquad \leq \mathbb{P}\left(\begin{array}{l} \exists k' \in [N],\ s_2 \in [\![t_4+1, t]\!] \text{ and } k_2 \in \mathcal{N}_{k',t_4}(s_2): \\ \sum_{(i,b,s)\in P_{k',t_4}^{k_2,s_2}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\leq rx_N\}} \geq x_N \end{array}\right). \end{aligned} \tag{2.4.13}$$

Now consider the N-BRW constructed from N independent BRWs (see Section 2.4.1). Assume that $k' \in [N]$, $s_2 \in [\![t_4+1, t]\!]$ and $k_2 \in \mathcal{N}_{k',t_4}(s_2)$ are such that

$$\sum_{(i,b,s)\in P_{k',t_4}^{k_2,s_2}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\leq rx_N\}} \geq x_N.$$

Then by Lemma 2.4.1 there exists a path in one of the N independent BRWs that contains the same jumps as the path $P_{k',t_4}^{k_2,s_2}$. Thus Lemma 2.4.1 implies that there exist $(j, u) \in H_{t_4}$

and $(j,v) \in H_{s_2}$ such that $u \prec v$ and

$$\sum_{u \prec w \preceq v} Y_{j,w} \mathbb{1}_{\{Y_{j,w} \le r x_N\}} \ge x_N.$$

That is, there is a path in the N independent BRWs between times t_4 and s_2 which moves at least x_N even with discarding jumps of size greater than $r x_N$. This means that there must be a path with the same property between times t_4 and t as well, because all jumps are non-negative. Therefore

$$\begin{aligned} &\mathbb{P}\left(\begin{array}{l} \exists k' \in [N],\ s_2 \in [\![t_4+1, t]\!] \text{ and } k_2 \in \mathcal{N}_{k',t_4}(s_2) : \\ \sum_{(i,b,s) \in P^{k_2,s_2}_{k',t_4}} X_{i,b,s} \mathbb{1}_{\{X_{i,b,s} \le r x_N\}} \ge x_N \end{array} \right) \\ &\le \mathbb{P}\bigg(\exists (j,u) \in H_{t_4} \text{ and } v \in \{1,2\}^t \text{ with } v \succ u : \sum_{u \prec w \preceq v} Y_{j,w} \mathbb{1}_{\{Y_{j,w} \le r x_N\}} \ge x_N \bigg). \end{aligned} \tag{2.4.14}$$

Let X_i, $i = 1, 2, \ldots$ be i.i.d. with distribution given by (2.1.3), and take $\lambda > 0$, $x_N > N^\lambda$, and $0 < r < 1 \wedge \frac{\lambda(1 \wedge \alpha)}{48}$. Note that the random variables $Y_{j,w}$ are all distributed as the X_i random variables, and that there are $4\ell_N$ terms in the sum on the right-hand side of (2.4.14). We will give a union bound for the probability of the event on the right-hand side of (2.4.14), using that H_{t_4} is a set of N elements and that a particle in the set H_{t_4} has $2^{4\ell_N}$ descendants in a BRW (without selection) at time t, which means $2^{4\ell_N}$ possible labels for v for each $(j,u) \in H_{t_4}$. Then by (2.4.13), (2.4.14) and by conditioning on $\mathcal{F}'_{t_4}$ and using a union bound,

$$\begin{aligned} &\mathbb{P}\left(\begin{array}{l} \exists (k_1, s_1) \in [N] \times [\![t_4, t-1]\!],\ s_2 \in [\![s_1+1, t]\!] \text{ and } k_2 \in \mathcal{N}_{k_1,s_1}(s_2) : \\ \sum_{(i,b,s) \in P^{k_2,s_2}_{k_1,s_1}} X_{i,b,s} \mathbb{1}_{\{X_{i,b,s} \le r x_N\}} \ge x_N \end{array} \right) \\ &\qquad \le N 2^{4\ell_N} \mathbb{P}\left(\sum_{j=1}^{4\ell_N} X_j \mathbb{1}_{\{X_j \le r x_N\}} \ge x_N \right). \end{aligned} \tag{2.4.15}$$

Then by Lemma 2.4.3 with $m = 4$ and $q = 6$, we have that for N sufficiently large,

$$\mathbb{P}\left(\sum_{j=1}^{4\ell_N} X_j \mathbb{1}_{\{X_j \le r x_N\}} \ge x_N \right) \le N^{-6}.$$

The result follows by (2.4.15). □

2.4.3 Simple properties of the regularly varying function h

In order to bound the probabilities of the events $\mathcal{C}_2$ to $\mathcal{C}_7$ and $\mathcal{D}_1$ to $\mathcal{D}_5$, we will need to use several properties of the function h from (2.1.3). Recall that h is regularly varying with index $\alpha > 0$, and that it determines the jump distribution of the N-BRW in the sense that for each jump (i, b, s),

$$\mathbb{P}(X_{i,b,s} > x) = h(x)^{-1} \quad \forall x \geq 0. \tag{2.4.16}$$

Recall that $a_N = h^{-1}(2N\ell_N)$, and note that $a_N \to \infty$ as $N \to \infty$. Indeed, by the definition of h^{-1} in (2.1.6), a_N is non-decreasing, and since h is non-decreasing by (2.1.3), a_N cannot converge to a finite limit $a \in \mathbb{R}$, because this would imply $h(a+1) \geq 2N\ell_N \; \forall N$. Moreover, letting $C_2 = C_2(\alpha)$ as in Lemma 2.4.2, for N sufficiently large that $a_N + 1 \geq B = B(\alpha)$,

$$2N\ell_N < h(a_N + 1) \leq C_2(a_N + 1)^{2\alpha}, \tag{2.4.17}$$

where in the first inequality we use the definition (2.1.6) of h^{-1} and that h is non-decreasing, and the second inequality follows by the second inequality of Lemma 2.4.2.

Since h is regularly varying with index α, we have

$$\frac{2N\ell_N}{h(a_N)} \to 1 \quad \text{as } N \to \infty. \tag{2.4.18}$$

Indeed, since h is non-decreasing, for any $\varepsilon \in (0,1)$, by (2.1.2) and by the definition of a_N we have

$$(1-\varepsilon)^\alpha - \varepsilon \leq \frac{h(a_N(1-\varepsilon))}{h(a_N)} \leq \frac{2N\ell_N}{h(a_N)} \leq \frac{h(a_N(1+\varepsilon))}{h(a_N)} \leq (1+\varepsilon)^\alpha + \varepsilon,$$

for N sufficiently large. Often in our proofs it will be enough to use that (2.4.18) implies

$$\frac{1}{2} < \frac{2N\ell_N}{h(a_N)} < 2, \tag{2.4.19}$$

for N sufficiently large.

For convenience we state a few other simple properties of h, which we will apply several times. Let $r \in (0,1)$ and $\eta < 1/1000$. First, we have

$$\frac{1}{h(ra_N)} < \frac{1}{h(a_N)}(r^{-\alpha} + \eta^4) < \frac{1}{h(a_N)} 2r^{-\alpha}, \tag{2.4.20}$$

for N sufficiently large, by (2.1.2). Second, for N sufficiently large, we also have

$$\frac{2N\ell_N}{h(ra_N)} < \frac{2N\ell_N}{h(a_N)}(r^{-\alpha} + \eta^4) < (1+\eta^4)(r^{-\alpha} + \eta^4) < 2r^{-\alpha}, \tag{2.4.21}$$

by (2.4.20) and (2.4.18). Furthermore, by the same argument as for (2.4.21), for N sufficiently large,

$$\frac{2N\ell_N}{h(ra_N)} > \frac{r^{-\alpha}}{2}. \tag{2.4.22}$$

2.4.4 Probabilities and proof of Proposition 2.2.5

Next we will go through the events $(\mathcal{C}_j)_{j=2}^7$ and $(\mathcal{D}_i)_{i=1}^5$, which we defined in Section 2.3, one by one. We will prove upper bounds on the probabilities of their complement events, which will then allow us to prove Proposition 2.2.5. Recall that the events $(\mathcal{C}_j)_{j=2}^7$ and $(\mathcal{D}_i)_{i=1}^5$ all depend on the constants $\eta, K, \gamma, \delta, \rho, c_1 \dots, c_6$ introduced in (2.3.2)-(2.3.5), and Propositions 2.3.2 and 2.3.11 hold when the constants satisfy the conditions (2.3.2)-(2.3.5). In order to show that the events in question occur with high probability, the constants need to satisfy some extra conditions which are consistent with (2.3.2)-(2.3.5). We now specify these choices.

Recall that $\alpha > 0$. First we assume that $\eta \in (0,1]$ is very small; in particular, that it is small enough to satisfy

$$\eta^2 < \min\left(\left(2^{\alpha+2}\log\left(\frac{1000}{\eta}\right)\right)^{-1/\alpha}, \frac{\eta}{1000\cdot 2^{\alpha}}\right). \tag{2.4.23}$$

Then we choose the remaining constants as follows (we will see shortly that these choices are consistent with (2.3.2)-(2.3.5)):

(a) $c_6 := \eta^2$,

(b) $c_5 := \eta^{6(1\vee\alpha)}$,

(c) $c_4 := c_5^{4/(1\wedge\alpha)}$,

(d) take $c_3 > 0$ small enough to satisfy $c_3 < c_4^{4(1\vee\alpha)}$ and $(1-6c_3/c_4)^{\alpha} \geq 1 - 12\alpha c_3/c_4$,

(e) take $c_2 > 0$ small enough to satisfy $c_2 < c_3^{4(1\vee\alpha)}$ and $(1-4c_2/c_3)^{\alpha} \geq 1 - 8\alpha c_2/c_3$,

(f) $c_1 := c_2^2$,

(g) $\rho := c_1(1\wedge\alpha)^2/(100\alpha)$,

(h) $\delta := \rho^{\alpha+1}$,

(i) $\gamma := \delta/2$,

(j) $K := \rho^{-\alpha-1}$.

Note that the constants with the choices above can be thought of as in (2.3.1). We state a few simple consequences of these choices, which will be useful in proving upper bounds on the probabilities of the complement events of $\mathcal{C}_2$ to $\mathcal{C}_7$ and $\mathcal{D}_1$ to $\mathcal{D}_5$. First, by (2.4.23), we have

$$\eta < \frac{1}{1000\cdot 2^{\alpha}} < \frac{1}{1000}, \tag{2.4.24}$$

and note that all constants $\gamma, \delta, \rho, c_1 \dots, c_6$ and $1/K$ are at most η^2. Thus, from (a)-(f) and (2.4.24), for $j = 1, \dots, 5$, we have

$$c_j \leq c_{j+1}^2 \leq c_{j+1}\eta^2 < \frac{c_{j+1}}{10^6\cdot 2^{2\alpha}}, \tag{2.4.25}$$

which also means

$$c_j < \frac{\eta^2}{10^6} \tag{2.4.26}$$

for $j = 1, \ldots, 5$. In particular we will need that

$$\frac{c_2}{c_3} < \frac{1}{10^6(1 \vee \alpha)} \tag{2.4.27}$$

and

$$\frac{c_3}{c_4} < \frac{1}{10^6(1 \vee \alpha)}, \tag{2.4.28}$$

which both follow by (2.4.25) and by the fact that $2^{2\alpha} \geq e^\alpha \geq 1 \vee \alpha$ for $\alpha > 0$. We will also use that from (e) we have

$$c_3^{-\alpha-1} c_2 < c_3^{-2(1\vee\alpha)+4(1\vee\alpha)} \leq c_3^2 < \frac{c_4}{10^6 \cdot 2^{2\alpha}} \frac{c_4}{10^6 \alpha} < \frac{\eta^4}{16\alpha 2^\alpha}, \tag{2.4.29}$$

where we applied (2.4.25) and that $2^{2\alpha} \geq \alpha$, and then that $c_4 < \eta^2$. Then similarly, from (d) we have

$$c_4^{-\alpha-1} c_3 < \frac{\eta^4}{24\alpha 2^\alpha}. \tag{2.4.30}$$

Finally, from (g) and (2.4.26) we have

$$\rho < c_1 < \frac{\eta}{10^6}. \tag{2.4.31}$$

Considering the choices (a)-(j) together with the consequences (2.4.24) and (2.4.25), and noticing that (g) implies $\rho \leq c_1/100$, we conclude that the constants $\eta, K, \gamma, \delta, \rho, c_1 \ldots, c_6$ satisfy (2.3.2)-(2.3.5), so we will be able to apply Propositions 2.3.2 and 2.3.11 with this choice of constants.

We can now show that the events $\mathcal{C}_2$ to $\mathcal{C}_7$ and $\mathcal{D}_1$ to $\mathcal{D}_5$ occur with high probability.

Lemma 2.4.6. *Suppose the constants η, K, γ, δ, ρ, $c_1, \ldots, c_6 > 0$ satisfy* (2.4.23) *and* (a)-(j)*. Then for N sufficiently large and $t > 4\ell_N$,*

$$\mathbb{P}(\mathcal{C}_j^c) < \frac{\eta}{1000} \quad \text{and} \quad \mathbb{P}(\mathcal{D}_i^c) < \frac{\eta}{1000}$$

for all $j \in \{2, \ldots, 7\}$ and $i \in \{1, \ldots, 5\}$, where the events $(\mathcal{C}_j)_{j=2}^7$ and $(\mathcal{D}_i)_{i=1}^5$ are defined in (2.3.12)-(2.3.17) *and* (2.3.47)-(2.3.52) *respectively.*

Proof. Assume that $\eta > 0$ satisfies (2.4.23). We consider the events $(\mathcal{C}_j)_{j=2}^7$ and $(\mathcal{D}_i)_{i=1}^5$ with the constants K, γ, δ, ρ, $c_1, \ldots, c_6$, and we assume that these constants satisfy (a)-(j). We will upper bound the probabilities of the events $(\mathcal{C}_j^c)_{j=2}^7$ and $(\mathcal{D}_i^c)_{i=1}^5$ using (2.4.25)-(2.4.31) above, and the properties of the regularly varying function h described in Section 2.4.3.

The event $\mathcal{C}_2^c$ (see (2.3.12)) says that there is a time $s \in [t_3, t-1]$ when a particle at distance at least $c_3 a_N$ behind the leader jumps to within distance $2c_2 a_N$ of the leader's position. We use Markov's inequality, and sum over all the jumps happening between times t_3 and $t-1$ to bound the probability of this event. We have

$$\begin{aligned}\mathbb{P}(\mathcal{C}_2^c) &\leq \mathbb{E}\left[\#\left\{\begin{array}{l}(i,b,s)\in[N]\times\{1,2\}\times[\![t_3,t-1]\!]\text{ such that}\\ Z_i(s)\geq c_3a_N\text{ and }X_{i,b,s}\in(Z_i(s)-2c_2a_N,Z_i(s)+2c_2a_N]\end{array}\right\}\right]\\ &= \sum_{(i,b,s)\in[N]\times\{1,2\}\times[\![t_3,t-1]\!]}\mathbb{E}\left[\mathbb{1}_{\{Z_i(s)\geq c_3a_N\}}\mathbb{1}_{\{X_{i,b,s}\in(Z_i(s)-2c_2a_N,Z_i(s)+2c_2a_N]\}}\right].\end{aligned}$$

Recall from Section 2.2.4 that for $s\in\mathbb{N}$ and $i\in[N]$, the distance $Z_i(s)$ of the ith particle from the leader is $\mathcal{F}_s$-measurable, but the jumps performed at time s, $X_{i,1,s}$ and $X_{i,2,s}$, are independent of $\mathcal{F}_s$. Hence by (2.4.16),

$$\begin{aligned}&\mathbb{P}(\mathcal{C}_2^c)\\ &\leq \sum_{(i,b,s)\in[N]\times\{1,2\}\times[\![t_3,t-1]\!]}\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{Z_i(s)\geq c_3a_N\}}\mathbb{1}_{\{X_{i,b,s}\in(Z_i(s)-2c_2a_N,Z_i(s)+2c_2a_N]\}}\,\Big|\,\mathcal{F}_s\right]\right]\\ &= \sum_{(i,b,s)\in[N]\times\{1,2\}\times[\![t_3,t-1]\!]}\mathbb{E}\left[\mathbb{1}_{\{Z_i(s)\geq c_3a_N\}}\left(h(Z_i(s)-2c_2a_N)^{-1}-h(Z_i(s)+2c_2a_N)^{-1}\right)\right].\end{aligned}\tag{2.4.32}$$

Since h is monotone non-decreasing, for any $z\geq c_3a_N$ we have

$$h(z-2c_2a_N)^{-1}-h(z+2c_2a_N)^{-1}\leq h((c_3-2c_2)a_N)^{-1}\left(1-\frac{h(z-2c_2a_N)}{h(z+2c_2a_N)}\right).\tag{2.4.33}$$

Take $\epsilon>0$. For the fraction on the right-hand side of (2.4.33) we have that for N sufficiently large, for $z\geq c_3a_N$,

$$1\geq\frac{h(z-2c_2a_N)}{h(z+2c_2a_N)}\geq\frac{h\left((z+2c_2a_N)\cdot\frac{(c_3-2c_2)a_N}{(c_3+2c_2)a_N}\right)}{h(z+2c_2a_N)}\geq\left(1-\frac{4c_2}{c_3+2c_2}\right)^{\alpha}-\epsilon\geq 1-8\alpha\frac{c_2}{c_3}-\epsilon,\tag{2.4.34}$$

where we first use the monotonicity of h, and in the second inequality we use that $z\geq c_3a_N$, that the function $y\mapsto(y-2c_2a_N)/(y+2c_2a_N)$ is increasing in y, and we again use the monotonicity of h. The third inequality follows by (2.1.2), and the fourth holds by the definition of c_2 in (e). Then, by (2.4.20) and the lower bound in (2.4.34) with $\epsilon=\eta^4(c_3-2c_2)^{\alpha}$, we see from (2.4.33) that for N sufficiently large, for $z\geq c_3a_N$,

$$\begin{aligned}h(z-2c_2a_N)^{-1}-h(z+2c_2a_N)^{-1}&\leq 2(c_3-2c_2)^{-\alpha}h(a_N)^{-1}\left(8\alpha\frac{c_2}{c_3}+\eta^4(c_3-2c_2)^{\alpha}\right)\\ &\leq h(a_N)^{-1}(16\alpha 2^{\alpha}c_3^{-\alpha-1}c_2+2\eta^4),\end{aligned}\tag{2.4.35}$$

where for the first term of the second inequality we used the fact that $(c_3-2c_2)^{-\alpha}<(c_3/2)^{-\alpha}$, because $2c_2<c_3/2$ by (2.4.25).

Now let us return to (2.4.32) and notice that we sum over $6N\ell_N$ jumps. Therefore, by (2.4.35) we conclude that for N sufficiently large,

$$\mathbb{P}(\mathcal{C}_2^c) \leq \frac{6N\ell_N}{h(a_N)}(16\alpha 2^\alpha c_3^{-\alpha-1}c_2 + 2\eta^4) \leq 6(16\alpha 2^\alpha c_3^{-\alpha-1}c_2 + 2\eta^4) < 18\eta^4 < \frac{\eta}{1000},$$

where we used (2.4.19) in the second inequality, (2.4.29) in the third, and (2.4.24) in the fourth.

The event $\mathcal{C}_3^c$ (see (2.3.13)) says that there exists a big jump in the time interval $[t_4, t-1]$ such that a descendant also performs a big jump during the time interval $[t_4, t-1]$, within time ℓ_N of the first big jump.

Consider the N-BRW constructed from N independent BRWs (see Section 2.4.1). If $\mathcal{C}_3^c$ occurs then there must be two big jumps in the N-BRW as above; that is, we must have $(i_1, s_1) \lesssim_{b_1} (i_2, s_2)$ with $s_1 \in [\![t_4, t-2]\!]$ and $s_2 \in [\![s_1+1, \min\{s_1+\ell_N, t-1\}]\!]$, and $(i_1, b_1, s_1), (i_2, b_2, s_2) \in B_N$, where B_N is the set of big jumps defined in (2.2.15). Then by Lemma 2.4.1 there are two big jumps with the same properties in the N independent BRWs; that is, there exist $j \in [N]$, $u_1, u_2 \in \mathcal{U}_0$ such that $(j, u_1) \in H_{s_1}$, $(j, u_2) \in H_{s_2}$, $u_1 b_1 \preceq u_2$, $X_{i_1,b_1,s_1} = Y_{j,u_1b_1}$ and $X_{i_2,b_2,s_2} = Y_{j,u_2b_2}$. Therefore, since $s_2 \in [\![s_1+1, \min\{s_1+\ell_N, t-1\}]\!] \subseteq [\![s_1+1, s_1+\ell_N]\!]$ and $H_{s_2} \subseteq [N] \times \{1,2\}^{s_2}$, we have

$$\mathbb{P}(\mathcal{C}_3^c) \leq \mathbb{P}\begin{pmatrix} \exists s_1 \in [\![t_4, t-2]\!], (j, u_1) \in H_{s_1}, b_1 \in \{1,2\} \text{ and} \\ s_2 \in [\![s_1+1, s_1+\ell_N]\!], u_2 \in \{1,2\}^{s_2}, u_2 \succ u_1, b_2 \in \{1,2\} : \\ Y_{j,u_1b_1} > \rho a_N \text{ and } Y_{j,u_2b_2} > \rho a_N \end{pmatrix}. \tag{2.4.36}$$

Recall the definition of $\mathcal{F}'_n$ in Section 2.4.1. By a union bound over the possible values of s_1, s_2, b_1 and b_2, and then conditioning on $\mathcal{F}'_{s_1}$ and applying another union bound over the possible values of (j, u_1) and u_2,

$$\mathbb{P}(\mathcal{C}_3^c) \leq \sum_{\substack{s_1 \in [\![t_4, t-2]\!], \\ s_2 \in [\![s_1+1, s_1+\ell_N]\!], \\ b_1, b_2 \in \{1,2\}}} \mathbb{E}\left[\sum_{(j,u_1)\in H_{s_1}, u_2 \in \{1,2\}^{s_2}, u_2 \succ u_1} \mathbb{P}\left(Y_{j,u_1b_1} > \rho a_N, Y_{j,u_2b_2} > \rho a_N \middle| \mathcal{F}'_{s_1}\right)\right].$$

Then since $(Y_{j,u})_{j\in[N], u\in\cup_{m>s_1}\{1,2\}^m}$ are independent of $\mathcal{F}'_{s_1}$, for $(j, u_1) \in H_{s_1}$ and $u_2 \in \{1,2\}^{s_2}$ we have

$$\mathbb{P}\left(Y_{j,u_1b_1} > \rho a_N, Y_{j,u_2b_2} > \rho a_N \middle| \mathcal{F}'_{s_1}\right) = h(\rho a_N)^{-2}.$$

Hence by summing over the $4\ell_N - 1$ possible values for s_1, and the two possible values for b_1 and b_2, and since $|H_{s_1}| = N$, and for $u_1 \in \{1,2\}^{s_1}$ there are $2^{s_2-s_1}$ possible values of

$u_2 \in \{1,2\}^{s_2}$ with $u_2 \succ u_1$, for N sufficiently large we have

$$\mathbb{P}(\mathcal{C}_3^c) \le 4\ell_N \cdot 4 \sum_{s_2 \in [\![s_1+1, s_1+\ell_N]\!]} N 2^{s_2 - s_1} h(\rho a_N)^{-2}$$
$$\le 16 N \ell_N \cdot 2 \cdot 2^{\log_2 N + 1} h(\rho a_N)^{-2} = \left(\frac{2N\ell_N}{h(\rho a_N)} \right)^2 16 \ell_N^{-1} \le 4\rho^{-2\alpha} \cdot 16 \ell_N^{-1} < \frac{\eta}{1000}, \tag{2.4.37}$$

where in the third inequality we used (2.4.21).

The event $\mathcal{C}_4^c$ (see (2.3.14)) can be bounded using Corollary 2.4.5. We apply the corollary with $x_N = c_1 a_N$, $r = \rho / c_1$ and $\lambda = 1/(2\alpha)$. We can make this choice for λ, because we have

$$c_1 a_N > N^{1/(2\alpha)} \tag{2.4.38}$$

for all N sufficiently large by (2.4.17). By our choice of ρ in (g), we have $r < 1 \wedge \frac{\lambda(1\wedge\alpha)}{48}$, and so Corollary 2.4.5 tells us that for some constant $C > 0$, for N sufficiently large,

$$\mathbb{P}(\mathcal{C}_4^c) \le C N^{-1} < \frac{\eta}{1000}. \tag{2.4.39}$$

The event $\mathcal{C}_5^c$ (see (2.3.15)) says that two big jumps occur at the same time, that is

$$\mathcal{C}_5^c = \big\{ \exists s \in [\![t_4, t-1]\!],\ (k_1, b_1) \ne (k_2, b_2) \in [N] \times \{1,2\}$$
$$: X_{k_1, b_1, s} > \rho a_N \text{ and } X_{k_2, b_2, s} > \rho a_N \big\}.$$

By a union bound over the $4\ell_N$ time steps and the possible pairs of jumps at each time step,

$$\mathbb{P}(\mathcal{C}_5^c) \le 4\ell_N \binom{2N}{2} h(\rho a_N)^{-2} \le \left(\frac{2N\ell_N}{h(\rho a_N)} \right)^2 2\ell_N^{-1} \le 4\rho^{-2\alpha} \cdot 2\ell_N^{-1} < \frac{\eta}{1000} \tag{2.4.40}$$

for N sufficiently large, where the third inequality follows by (2.4.21).

The event $\mathcal{C}_6^c$ (see (2.3.16)) says that a big jump happens in (at least) one of two very short time intervals, $[t_2, t_2 + \lceil \delta\ell_N \rceil]$ and $[t_1 - \lceil \delta\ell_N \rceil, t_1 + \lceil \delta\ell_N \rceil]$. In total there are $2N \cdot (3\lceil \delta\ell_N \rceil + 2)$ jumps performed during these two time intervals. By a union bound over these jumps, we get

$$\mathbb{P}(\mathcal{C}_6^c)$$
$$= \mathbb{P}(\exists (i,b,s) \in [N] \times \{1,2\} \times ([\![t_2, t_2 + \lceil \delta\ell_N \rceil]\!] \cup [\![t_1 - \lceil \delta\ell_N \rceil, t_1 + \lceil \delta\ell_N \rceil]\!]) :\ X_{i,b,s} > \rho a_N)$$
$$\le 2N(3\delta\ell_N + 5) h(\rho a_N)^{-1} \le 6\delta\rho^{-\alpha}(1 + 2\delta^{-1}\ell_N^{-1}) < \frac{\eta}{1000}, \tag{2.4.41}$$

for N sufficiently large, where in the second inequality we used (2.4.21), and the last inequality follows by the choice of δ in (h) and by (2.4.31).

The event $\mathcal{C}_7$ gives an upper bound on the number of big jumps (see (2.3.17)). There are $8N\ell_N$ jumps performed in the time interval $[t_4, t-1]$; by Markov's inequality and then by (2.4.21), we have

$$\begin{aligned}\mathbb{P}(\mathcal{C}_7^c) &= \mathbb{P}(\#\{(i,b,s) \in [N] \times \{1,2\} \times [\![t_4, t-1]\!] : X_{i,b,s} > \rho a_N\} > K) \\ &\leq \frac{8N\ell_N h(\rho a_N)^{-1}}{K} \leq \frac{8}{K}\rho^{-\alpha} < \frac{\eta}{1000}\end{aligned} \tag{2.4.42}$$

for N sufficiently large, where the last inequality follows by the choice of K in (j) and by (2.4.31).

The event $\mathcal{D}_1$ (see (2.3.47)) has the same definition as that of $\mathcal{C}_2$ (see (2.3.12)), except with different constants. By the same argument as for (2.4.35), using the definition of c_3 in (d), for N sufficiently large we have

$$h(z - 3c_3 a_N)^{-1} - h(z + 3c_3 a_N)^{-1} \leq h(a_N)^{-1}(24\alpha \cdot 2^\alpha c_4^{-\alpha-1} c_3 + 2\eta^4) \quad \forall z \geq c_4 a_N. \tag{2.4.43}$$

Then continuing in the same way as after (2.4.35) we obtain

$$\mathbb{P}(\mathcal{D}_1^c) \leq 6(24\alpha 2^\alpha c_4^{-\alpha-1} c_3 + 2\eta^4) < 18\eta^4 < \frac{\eta}{1000}, \tag{2.4.44}$$

for N sufficiently large, by (2.4.30) and (2.4.24).

The event $\mathcal{D}_2$ in (2.3.48) says that in every interval of length $c_5\ell_N$ in $[t_3, t_1]$ there is a particle which performs a jump of size greater than $2c_4 a_N$. We introduce a slightly different event to show that $\mathcal{D}_2$ happens with high probability. Let us divide the interval $[t_3, t_1]$ into subintervals of length $\frac{1}{2}c_5\ell_N$, to get $\lceil 4c_5^{-1} \rceil$ subintervals (the last subinterval may end after time t_1; we also note that $\frac{1}{2}c_5\ell_N$ is not necessarily an integer, but we will intersect the intervals with $\mathbb{N}$). If a jump of size greater than $2c_4 a_N$ happens in each of these subintervals then $\mathcal{D}_2$ occurs. We describe this formally by the following event:

$$\tilde{\mathcal{D}}_2 := \left\{ \begin{array}{l} \forall m \in \{1, 2, \dots, \lceil 4c_5^{-1} \rceil\}, \\ \exists (k,b,s) \in [N] \times \{1,2\} \times [\![t_3 + (m-1)\frac{1}{2}c_5\ell_N, t_3 + m\frac{1}{2}c_5\ell_N]\!] : \\ X_{k,b,s} > 2c_4 a_N \end{array} \right\};$$

as mentioned above, if $\tilde{\mathcal{D}}_2$ occurs then $\mathcal{D}_2$ occurs. The complement event of $\tilde{\mathcal{D}}_2$ is that there is a subinterval in which every jump made by a particle has size at most $2c_4 a_N$. Note that in each subinterval $[\![t_3 + (m-1)\frac{1}{2}c_5\ell_N, t_3 + m\frac{1}{2}c_5\ell_N]\!]$, there are at least $2N \cdot \frac{1}{2}c_5\ell_N$ jumps.

Therefore, by a union bound, we have

$$\begin{aligned}\mathbb{P}(\mathcal{D}_2^c) \le \mathbb{P}(\tilde{\mathcal{D}}_2^c) \le \lceil 4c_5^{-1} \rceil \left(1 - \frac{1}{h(2c_4 a_N)}\right)^{c_5 \ell_N N} &\le (4c_5^{-1}+1) \exp\left(-\frac{c_5 N \ell_N}{h(2c_4 a_N)}\right) \\ &\le 5c_5^{-1} \exp\left(-\frac{c_5 (2c_4)^{-\alpha}}{4}\right), \qquad (2.4.45)\end{aligned}$$

where in the third inequality we use that $1-x \le e^{-x}$ for $x \ge 0$, and the fourth inequality follows by (2.4.22) for N sufficiently large and since $c_5 < 1$. Now note that by (c),

$$c_5 c_4^{-\alpha} = c_5^{1-4\alpha/(1\wedge\alpha)} \ge c_5^{-3} > 2^{2+\alpha} \log\left(\frac{5000}{c_5 \eta}\right),$$

where the last inequality holds because $c_5^{-1} > 2^{2+\alpha}$ by (2.4.25), $0 < \log x < x$ for $x > 1$, and $c_5^{-1} > \frac{5000}{\eta}$ by (2.4.26). Substituting this into (2.4.45) shows that $\mathbb{P}(\mathcal{D}_2^c) < \eta/1000$.

The event $\mathcal{D}_3^c$ defined in (2.3.49) says that every jump in the time interval $[t_2, t_2 + \lceil \ell_N/2 \rceil]$ has size at most $2c_6 a_N$. There are at least $N\ell_N$ jumps in this time interval, and so for N sufficiently large, since $e^{-x} \ge 1-x$ for $x \ge 0$, and then by (2.4.22),

$$\mathbb{P}(\mathcal{D}_3^c) \le \left(1 - \frac{1}{h(2c_6 a_N)}\right)^{N\ell_N} \le \exp\left(-\frac{N\ell_N}{h(2c_6 a_N)}\right) \le \exp\left(-\frac{(2c_6)^{-\alpha}}{4}\right). \qquad (2.4.46)$$

Now (a) and (2.4.23) tell us that $c_6^{-\alpha} = \eta^{-2\alpha} > 2^{\alpha+2} \log(\frac{1000}{\eta})$, and substituting this into (2.4.46) shows that $\mathbb{P}(\mathcal{D}_3^c) < \eta/1000$.

The event $\mathcal{D}_4^c$ (see (2.3.50)) says that in the time interval $[t_2 - \lceil c_5 \ell_N \rceil, t_2]$, a particle performs a jump of size greater than $c_6 a_N$ (recall from (a) and (b) that $c_5 \ll c_6$). Since there are at most $2N(\lceil c_5 \ell_N \rceil + 1) \le 2N(c_5\ell_N + 2)$ jumps in the time interval $[t_2 - \lceil c_5 \ell_N \rceil, t_2]$, by a union bound,

$$\begin{aligned}\mathbb{P}(\mathcal{D}_4^c) &= \mathbb{P}(\exists (i,b,s) \in [N] \times \{1,2\} \times [\![t_2 - \lceil c_5 \ell_N \rceil, t_2]\!] : X_{i,b,s} > c_6 a_N) \\ &\le \frac{2N(c_5 \ell_N + 2)}{h(c_6 a_N)} \le 2c_5 c_6^{-\alpha}(1 + 2c_5^{-1}\ell_N^{-1}) \le 4\eta^{6(1\vee\alpha)}\eta^{-2\alpha} < \frac{\eta}{1000}, \qquad (2.4.47)\end{aligned}$$

for N sufficiently large, where in the second inequality we use (2.4.21), the third inequality holds by the choices in (b) and (a) for N sufficiently large, and the fourth follows by (2.4.24).

The event $\mathcal{D}_5^c$ (see (2.3.52)) says that in a short time interval after time τ_2 (defined in (2.3.51)) a jump is performed whose size falls into a small interval, $(2c_4 a_N, (2c_4 + 3c_3) a_N]$. We can see from the definition of τ_2 as the first time after t_2 when the diameter is at most $\frac{3}{2} c_4 a_N$, that τ_2 is a stopping time. Therefore we can condition on $\mathcal{F}_{\tau_2}$, and apply the strong

Markov property. By Markov's inequality we have

$$\begin{aligned}
&\mathbb{P}(\mathcal{D}_5^c) \\
&= \mathbb{P}(\exists (k,b,s) \in [N] \times \{1,2\} \times [\![\tau_2, \tau_2 + c_5 \ell_N]\!] : \; X_{k,b,s} \in (2c_4 a_N, (2c_4 + 3c_3) a_N]) \\
&\le \mathbb{E}\left[\mathbb{E}[\# \{(k,b,s) \in [N] \times \{1,2\} \times [\![\tau_2, \tau_2 + c_5 \ell_N]\!] : X_{k,b,s} \in (2c_4 a_N, (2c_4 + 3c_3) a_N]\} \,|\mathcal{F}_{\tau_2}]\right].
\end{aligned}$$

Note that if $\tau_2 < \infty$ then during the time interval $[\tau_2, \tau_2 + c_5 \ell_N]$ there are at most $2N(c_5 \ell_N + 1)$ jumps; it follows that

$$\begin{aligned}
\mathbb{P}(\mathcal{D}_5^c) &\le \mathbb{E}\Bigg[\sum_{(k,b,s) \in [N] \times \{1,2\} \times [\![\tau_2, \tau_2 + c_5 \ell_N]\!]} \mathbb{P}\left(X_{k,b,s} \in (2c_4 a_N, (2c_4 + 3c_3) a_N] \,|\, \mathcal{F}_{\tau_2}\right) \mathbb{1}_{\{\tau_2 < \infty\}} \Bigg] \\
&\le 2N(c_5 \ell_N + 1) \left(h(2c_4 a_N)^{-1} - h((2c_4 + 3c_3) a_N)^{-1} \right) \qquad (2.4.48)
\end{aligned}$$

by the strong Markov property. Now we can use the monotonicity of h and then the upper bound (2.4.43) to get

$$\begin{aligned}
h(2c_4 a_N)^{-1} - h((2c_4 + 3c_3) a_N)^{-1} &\le h((2c_4 - 3c_3) a_N)^{-1} - h((2c_4 + 3c_3) a_N)^{-1} \\
&\le h(a_N)^{-1} (24\alpha \cdot 2^\alpha c_4^{-\alpha - 1} c_3 + 2\eta^4) \qquad (2.4.49)
\end{aligned}$$

for N sufficiently large. Therefore, by (2.4.48), (2.4.49), and (2.4.18), we have that for N sufficiently large,

$$\mathbb{P}(\mathcal{D}_5^c) \le (1 + c_5^{-1} \ell_N^{-1}) c_5 (1 + \eta^4)(24\alpha 2^\alpha c_4^{-\alpha-1} c_3 + 2\eta^4) < 4c_5 \cdot 3\eta^4 < \frac{\eta}{1000}, \qquad (2.4.50)$$

where in the second inequality we use (2.4.30) and that $(1 + c_5^{-1} \ell_N^{-1})(1 + \eta^4) < 4$ for N sufficiently large, and the last inequality follows by (2.4.26) and (2.4.24). This concludes the proof of Lemma 2.4.6. □

We have seen in Lemma 2.4.6 above that with an appropriate choice of constants, the probabilities of the events $\mathcal{C}_2$ to $\mathcal{C}_7$ and $\mathcal{D}_1$ to $\mathcal{D}_5$ which imply $\mathcal{A}_1$ and $\mathcal{A}_3$ are close to 1. We can now use this to prove Proposition 2.2.5.

Proof of Proposition 2.2.5. Take $\eta \in (0,1]$. Without loss of generality, we can assume that η is sufficiently small that it satisfies (2.4.23). Then choose K, γ, δ, ρ, $c_1, \ldots, c_6$ as in (a)-(j) (at the beginning of Section 2.4.4). Note that before stating Lemma 2.4.6 we checked that these constants also satisfy (2.3.2)-(2.3.5). Therefore by Proposition 2.3.2 and Proposition 2.3.11, for N sufficiently large and $t > 4\ell_N$,

$$\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i \subseteq \mathcal{A}_1 \cap \mathcal{A}_3.$$

Therefore, for N sufficiently large and $t > 4\ell_N$, by a union bound,

$$\mathbb{P}((\mathcal{A}_1 \cap \mathcal{A}_3)^c) \leq \mathbb{P}\left(\left(\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i\right)^c\right) \leq \sum_{j=2}^{7} \mathbb{P}(\mathcal{C}_j^c) + \sum_{i=1}^{5} \mathbb{P}(\mathcal{D}_i^c) < \eta$$

by Lemma 2.4.6, which completes the proof. □

2.5 Proof of Proposition 2.2.6: star-shaped coalescence

We will prove Proposition 2.2.6 in this section. So far we have proved Proposition 2.2.5, which says that with high probability the common ancestor of the majority of the population at time t is particle (N, T), where T is given by (2.2.17); in particular, T is between times t_2 and t_1. Now recall the notation introduced in (2.2.19)-(2.2.23). Proposition 2.2.6 says that for $\nu > 0$, with high probability, every particle in the set $\mathcal{N}_{N,T}(T + \varepsilon_N \ell_N)$ has at most νN surviving descendants at time t, where we may assume that $(\varepsilon_N)_{N \in \mathbb{N}_0}$ satisfies

$$\varepsilon_N \ell_N \in \mathbb{N}_0 \;\; \forall N \geq 1, \qquad \varepsilon_N \ell_N \to \infty \text{ as } N \to \infty \quad \text{ and } \quad \varepsilon_N \leq \frac{1}{4} \frac{\log_2 \ell_N}{\ell_N} \;\; \forall N \geq 1. \tag{2.5.1}$$

The first two of these assumptions on ε_N are from (2.2.2). The third can be made without loss of generality, because if $\varepsilon'_N > \varepsilon_N$, and every particle in $\mathcal{N}_{N,T}(T + \varepsilon_N \ell_N)$ has at most νN surviving descendants at time t, then certainly every particle in $\mathcal{N}_{N,T}(T + \varepsilon'_N \ell_N)$ has at most νN surviving descendants at time t.

Fix $\eta \in (0, 1]$ sufficiently small that it satisfies (2.4.23). Then choose K, γ, δ, ρ, $c_1, \ldots, c_6$ as in (a)-(j). Then take N sufficiently large that Proposition 2.2.5 and Lemma 2.4.6 hold for our chosen constants, and take $t > 4\ell_N$. Let $\nu > 0$ be fixed and let us write $\mathcal{A}_4 := \mathcal{A}_4(\nu)$ from now on.

2.5.1 Strategy

Our strategy for showing Proposition 2.2.6 is to give a lower bound on the position of the leftmost particle at time t with high probability, and then bound the number of time-t descendants of each particle in $\mathcal{N}_{N,T}(T^{\varepsilon_N})$ which can reach that lower bound by time t. We will be able to control the number of such descendants because of Corollary 2.4.5. Assume that we know $\mathcal{X}_1(t) \geq \mathcal{X}_N(T) + \hat{a}_{T,N}$, where $\hat{a}_{T,N} > N^{\lambda}$ for some $\lambda > 0$, but $\hat{a}_{T,N} \ll a_N$. Then Corollary 2.4.5 implies that with high probability all surviving particles at time t must have an ancestor which made a jump of size greater than $r\hat{a}_{T,N}$ for an appropriate choice of $r \in (0, 1)$. So given a particle $i \in \mathcal{N}_{N,T}(T^{\varepsilon_N})$, we can find an upper bound for the number of its time-t descendants with high probability, by considering the number of its descendants which made a jump of size greater than $r\hat{a}_{T,N}$ before time t. Thus, we should choose $\hat{a}_{T,N}$ such that we have $\mathcal{X}_1(t) \geq \mathcal{X}_N(T) + \hat{a}_{T,N}$ with high probability, and also such that we can get a good enough upper bound for each D_i (see (2.2.21)) from Corollary 2.4.5 to conclude Proposition 2.2.6.

We now give a sketch argument to motivate our choice of lower bound on $\mathcal{X}_1(t)$. Assume that $T \in [t_2 + \lceil \delta\ell_N \rceil, t_1 - \lceil \delta\ell_N \rceil]$. We also assume that the record set at time T is not broken by a big jump before time $t_1 + \delta\ell_N$, and so almost all the descendants of particle (N, T) survive between times T and $T + \ell_N$. This all happens with high probability, as we saw in Section 2.4; in particular recall the event $\mathcal{C}_6$ from (2.3.16). Set $\theta_{T,N} := (t_1 - T)/\ell_N$.

Note that if a descendant of particle (N, T) makes a jump of size greater than $\hat{a}_{T,N}$ at time $T + k$ for some $k \in [(1-\delta)\ell_N, \ell_N]$, then it can have $2^{(1+\theta_{T,N})\ell_N - k}$ descendants at time t, and all of these descendants are to the right of $\mathcal{X}_N(T) + \hat{a}_{T,N}$. Also, there are approximately 2^k particles in the leading tribe descending from (N, T) at time $T + k$. Therefore, we expect that jumps of size greater than $\hat{a}_{T,N}$, performed by the descendants of (N, T) in the time interval $[T + (1-\delta)\ell_N, T + \ell_N]$, contribute to the number of particles to the right of $\mathcal{X}_N(T) + \hat{a}_{T,N}$ at time t by roughly

$$\sum_{k \in [\![(1-\delta)\ell_N, \ell_N]\!]} 2^k \cdot 2^{(1+\theta_{T,N})\ell_N - k} \frac{1}{h(\hat{a}_{T,N})} \approx \delta\ell_N 2^{(1+\theta_{T,N})\ell_N} \frac{1}{h(\hat{a}_{T,N})}.$$

If we want to make sure that all the N particles are to the right of $\mathcal{X}_N(T) + \hat{a}_{T,N}$ at time t, then the above should be approximately N, and so $\hat{a}_{T,N}$ should be roughly $h^{-1}(\delta\ell_N N^{\theta_{T,N}})$.

There are several potential inaccuracies in this argument. For example, the descendants of a particle making a jump of size greater than $\hat{a}_{T,N}$ do not necessarily all survive until time t. We will use a reasoning similar to Lemma 2.2.3 to clarify this issue. Another problem might occur if a particle $(i, T + k)$ makes a jump of size greater than $\hat{a}_{T,N}$, and then at time $T + k + 1$, its offspring does the same. In this case our sketch argument double counts the time-t descendants of particle $(i, T + k)$. We will therefore make some adjustments in the rigorous proof to avoid double counting.

In Sections 2.5.2 to 2.5.5 below, we will make the sketch argument precise, then use Corollary 2.4.5 to see that with high probability, particles must have at least one jump greater than a certain size (roughly but not exactly $h^{-1}(\delta\ell_N N^{\theta_{T,N}})$) in their ancestry to survive until time t. Finally, for each particle (i, T^{ε_N}), we upper bound the number of particles at time t which descend from particle (i, T^{ε_N}) and have a jump greater than this certain size in their ancestry between times T^{ε_N} and t.

2.5.2 Sequence of stopping times

In the strategy above we suggested that $h^{-1}(\delta\ell_N N^{\theta_{T,N}})$ should be a good lower bound for $\mathcal{X}_1(t) - \mathcal{X}_N(T)$. A problem with this lower bound is that it depends on T, and conditioning on T would change the distribution of the process, as T is not a stopping time; see the definition in (2.2.17).

Note however, that the first, second, $\ldots$, nth times after time t_2 at which a jump of size greater than ρa_N breaks the record between times t_2 and t_1, are stopping times, and

T is equal to one of these times with high probability. Furthermore, the number of such times is at most K with high probability, by Lemma 2.4.6 and the definition of the event $\mathcal{C}_7$. Therefore, we can define a finite set of stopping times in such a way that T is in the set with high probability. Then we can prove a similar statement to Proposition 2.2.6 for each stopping time in the finite set with the strategy described in the previous section. This will be enough to prove Proposition 2.2.6.

Recall the definition of $\mathbf{S}_N$ in (2.2.16). Define a sequence of stopping times by setting $T_0 := t_2 + \lceil \delta \ell_N \rceil - 1$, and

$$T_n := 1 + \inf \{\mathbf{S}_N(\rho) \cap [T_{n-1}, t_1 - \lceil \delta \ell_N \rceil - 1]\}, \tag{2.5.2}$$

for $n \in \mathbb{N}$; let $T_n := t_1$ if the intersection above is empty.

For all $n \in \mathbb{N}$, we introduce some new notation which will be frequently used in the course of the proof. First we let

$$T_n^{\varepsilon_N} := T_n + \varepsilon_N \ell_N. \tag{2.5.3}$$

The set and number of time-t descendants of the ith particle at time $T_n^{\varepsilon_N}$ will be denoted by

$$\mathcal{N}_{i,n} := \mathcal{N}_{i,T_n^{\varepsilon_N}}(t) \quad \text{and} \quad D_{i,n} := |\mathcal{N}_{i,n}|. \tag{2.5.4}$$

We also introduce

$$\theta_{n,N} := \frac{(t_1 - T_n)}{\ell_N} \geq 0. \tag{2.5.5}$$

Take $0 < \delta_1 < \delta/8$ and set

$$\hat{a}_{n,N} := h^{-1}(\delta_1 N^{\theta_{n,N}} \ell_N), \tag{2.5.6}$$

where h^{-1}, defined in (2.1.6), is the generalised inverse of h from (2.1.3). We explained the motivation for this definition of $\hat{a}_{n,N}$ in Section 2.5.1. By the same argument as for (2.4.18), for any $\epsilon > 0$ we can choose N_0 sufficiently large (and deterministic, since all quantities involved are deterministic) such that

$$\frac{x}{h(h^{-1}(x))} \in [1-\epsilon, 1+\epsilon] \quad \text{for all } N \geq N_0 \text{ and } x \geq \delta_1 \ell_N,$$

and then since $\delta_1 N^{\theta_{n,N}} \ell_N \geq \delta_1 \ell_N$, we also have for each $n \in \mathbb{N}$,

$$\frac{\delta_1 N^{\theta_{n,N}} \ell_N}{h(\hat{a}_{n,N})} \in [1-\epsilon, 1+\epsilon] \quad \text{for all } N \geq N_0. \tag{2.5.7}$$

We note that $\hat{a}_{n,N}$ is roughly $N^{\theta_{n,N}/\alpha}$; in particular, if $h(x) = x^\alpha$ for $x \geq 1$ then $\hat{a}_{n,N} = (\delta_1 N^{\theta_{n,N}} \ell_N)^{1/\alpha}$.

Take $0 < \delta_2 < \delta^2$. Throughout Section 2.5 we will use the term 'medium jump' for jumps of size greater than $\delta_2 \hat{a}_{n,N}$, as the relevant space scale in this section is $\hat{a}_{n,N}$. We denote the set of medium jumps on a time interval $[s_1, s_2] \subseteq [t_2, t-1]$ by

$$\mathcal{M}_{n,N}^{[s_1,s_2]} := \{(k,b,s) \in [N] \times \{1,2\} \times [\![s_1, s_2]\!] : \; X_{k,b,s} > \delta_2 \hat{a}_{n,N}\}, \tag{2.5.8}$$

and we let

$$\mathcal{M}_{n,N} := \mathcal{M}_{n,N}^{[t_2,t-1]}. \tag{2.5.9}$$

The stopping times $(T_n)_{n\in\mathbb{N}}$ allow us to give an upper bound on the probability of $\mathcal{A}_4^c$. Suppose $|B_N^{[t_2,t_1]}| \le K$ and $T \in [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]$. Then $|\mathbf{S}_N(\rho) \cap [t_2, t_1]| \le K$ by the definition of $\mathbf{S}_N$ in (2.2.16), and so by the definition of T in (2.2.17) and the definition of T_n in (2.5.2), it follows that $T = T_n$ for some $n \in [K]$. Hence, by the definition of $\mathcal{A}_4$ in (2.2.23) and then by a union bound,

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_4^c) &= \mathbb{P}\bigg(\max_{i \in \mathcal{N}_{N,T}(T^{\varepsilon_N})} D_i > \nu N \bigg) \\
&\le \mathbb{P}\bigg(\exists n \in [K] : \; T_n \le t_1 - \lceil \delta \ell_N \rceil \text{ and } \max_{i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})} D_{i,n} > \nu N \bigg) \\
&\qquad\qquad + \mathbb{P}\big(|B_N^{[t_2,t_1]}| > K\big) + \mathbb{P}\big(T \notin [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]\big).
\end{aligned} \tag{2.5.10}$$

By the definition of the event $\mathcal{C}_7$ in (2.3.17) and by Lemma 2.4.6,

$$\mathbb{P}(|B_N^{[t_2,t_1]}| > K) \le \mathbb{P}(\mathcal{C}_7^c) < \frac{\eta}{1000}.$$

Then by the definition of the event $\mathcal{A}_3$ in (2.2.22) and by Proposition 2.2.5,

$$\mathbb{P}(T \notin [t_2 + \lceil \delta \ell_N \rceil, t_1 - \lceil \delta \ell_N \rceil]) \le \mathbb{P}(\mathcal{A}_3^c) < \eta.$$

Therefore, applying a union bound for the first term on the right-hand side of (2.5.10), we obtain

$$\mathbb{P}(\mathcal{A}_4^c) \le \mathbb{E}\left[\sum_{n=1}^{K} \mathbb{1}_{\{T_n \le t_1 - \lceil \delta \ell_N \rceil\}} \mathbb{P}\bigg(\max_{i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})} D_{i,n} > \nu N \;\bigg|\; \mathcal{F}_{T_n} \bigg) \right] + \frac{1001}{1000}\eta. \tag{2.5.11}$$

From now on we aim to show that each term of the sum inside the expectation is small. For all $n \in \mathbb{N}$, we let $\mathbb{P}_{T_n}$ denote the law of the N-BRW conditioned on $\mathcal{F}_{T_n}$:

$$\mathbb{P}_{T_n}(\cdot) := \mathbb{P}(\cdot \mid \mathcal{F}_{T_n}) \quad \text{and} \quad \mathbb{E}_{T_n}[\cdot] := \mathbb{E}[\cdot \mid \mathcal{F}_{T_n}]. \tag{2.5.12}$$

2.5.3 Proof of Proposition 2.2.6

We now state the most important intermediate results in the proof of Proposition 2.2.6, and show that they imply the result. We then prove these intermediate results in Sections 2.5.4 and 2.5.5.

Our first main intermediate result says the probability that a particle in $\mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$ has a descendant at time t such that there is no medium jump on the path between the particle and the descendant is small. We prove this result in Section 2.5.4.

Lemma 2.5.1. *For all N sufficiently large, $t > 4\ell_N$, and $n \in \mathbb{N}$ with $T_n < t_1$,*

$$\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} : \; P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right) < \frac{\eta}{100K},$$

where T_n, $T_n^{\varepsilon_N}$ and $\mathbb{P}_{T_n}$ are given by (2.5.2), (2.5.3) *and* (2.5.12), *$\mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$ and $\mathcal{N}_{i,n}$ are defined in* (2.2.12) *and* (2.5.4), *$P^{k,t}_{i,T_n^{\varepsilon_N}}$ in* (2.2.10), *and $\mathcal{M}_{n,N}$ in* (2.5.9).

Our second main intermediate result says that with high probability, for each $i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$, there cannot be more than νN time-t descendants of particle $(i, T_n^{\varepsilon_N})$ if each descendant has a medium jump on their path. We prove this result in Section 2.5.5.

Lemma 2.5.2. *For all N sufficiently large, $t > 4\ell_N$, and $n \in \mathbb{N}$ with $T_n < t_1$,*

$$\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : \; D_{i,n} > \nu N \text{ and } P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} \neq \emptyset \;\; \forall k \in \mathcal{N}_{i,n}\right) < \frac{\eta}{100K},$$

where T_n, $T_n^{\varepsilon_N}$ and $\mathbb{P}_{T_n}$ are given by (2.5.2), (2.5.3) *and* (2.5.12), *$\mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$, $\mathcal{N}_{i,n}$ and $D_{i,n}$ are defined in* (2.2.12) *and* (2.5.4), *$P^{k,t}_{i,T_n^{\varepsilon_N}}$ in* (2.2.10), *and $\mathcal{M}_{n,N}$ in* (2.5.9).

Proof of Proposition 2.2.6. Suppose N is sufficiently large that Lemmas 2.5.1 and 2.5.2 hold. Take $n \in \mathbb{N}$ and suppose $T_n < t_1$ (which also implies $T_n \leq t_1 - \lceil \delta \ell_N \rceil$ by the definition (2.5.2) of T_n). Suppose a particle in $\mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$ has more than νN surviving descendants at time t. Then either all the descendants have an ancestor which performed a medium jump between times $T_n^{\varepsilon_N}$ and t, or there is at least one particle which survives without a medium jump in its ancestry. Therefore we have

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\max_{i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})} D_{i,n} > \nu N\right) \\
&\leq \mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} : \; P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right) \\
&\quad + \mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : \; D_{i,n} > \nu N \text{ and } P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} \neq \emptyset \;\; \forall k \in \mathcal{N}_{i,n}\right) \\
&< \frac{\eta}{50K}
\end{aligned} \tag{2.5.13}$$

by Lemmas 2.5.1 and 2.5.2. Then by (2.5.11), it follows that

$$\mathbb{P}(\mathcal{A}_4^c) < K \cdot \frac{\eta}{50K} + \frac{1001}{1000}\eta < 2\eta,$$

which completes the proof. □

2.5.4 Leaders must take medium jumps to survive: proof of Lemma 2.5.1

There are two key ideas in the proof. First we show that for a fixed $n \in \mathbb{N}$ with $T_n < t_1$, the whole population is to the right of position $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ at time t, with high probability.

Second, we prove that with high probability paths cannot reach position $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ without having a medium jump on the path.

Lemma 2.5.3. *For all N sufficiently large, $t > 4\ell_N$, and $n \in \mathbb{N}$ with $T_n < t_1$,*

$$\mathbb{P}_{T_n}(\mathcal{X}_1(t) < \mathcal{X}_N(T_n) + \hat{a}_{n,N}) < \frac{\eta}{200K},$$

where T_n and $\hat{a}_{n,N}$ are given by (2.5.2) *and* (2.5.6) *respectively.*

Proof. Recall the definition of $G_x(n)$ in (2.2.7). Let $G := G_{\mathcal{X}_N(T_n)+\hat{a}_{n,N}}(t)$; then, to prove the statement of the lemma, we aim to show that for N sufficiently large and $t > 4\ell_N$,

$$\mathbb{P}_{T_n}(|G| < N) < \frac{\eta}{200K}. \tag{2.5.14}$$

Recall the definition of $\delta_1 > 0$ in (2.5.6); fix $\delta' \in (8\delta_1, \delta)$ and then take $\delta_3 \in (8\delta_1, \delta')$ such that $\delta_3 \ell_N$ is an integer (this is possible for N sufficiently large). Let $S_k := T_n + \ell_N - k$ for $k \in [\![1, \delta_3\ell_N]\!]$. Then for each $k \in [\![1, \delta_3\ell_N]\!]$, at time S_k there are at least $2^{\ell_N - k}$ particles to the right of (or at) position $\mathcal{X}_N(T_n)$, by Lemma 2.2.3. These particles are either in the interval $[\mathcal{X}_N(T_n), \mathcal{X}_N(T_n) + \hat{a}_{n,N})$ or to the right of this interval. Let us denote the set of particles in $[\mathcal{X}_N(T_n), \mathcal{X}_N(T_n) + \hat{a}_{n,N})$ at time S_k by A_k, i.e. for $k \in [\![1, \delta_3\ell_N]\!]$ let

$$A_k := \{i \in [N] : \mathcal{X}_i(S_k) \in [\mathcal{X}_N(T_n), \mathcal{X}_N(T_n) + \hat{a}_{n,N})\}.$$

We will handle the following two cases separately:

(a) the event $\mathcal{E} := \left\{|A_k| \geq \frac{1}{2}2^{\ell_N - k} \ \forall k \in [\![1, \delta_3\ell_N]\!]\right\}$ occurs,

(b) the event $\mathcal{E}^c = \left\{\exists k \in [\![1, \delta_3\ell_N]\!] : |G_{\mathcal{X}_N(T_n)+\hat{a}_{n,N}}(S_k)| > \frac{1}{2}2^{\ell_N - k}\right\}$ occurs.

First we deal with case (a). We give a lower bound on $|G|$ using a similar argument to the proof of Lemma 2.2.3. First note that jumps of size greater than $\hat{a}_{n,N}$ from particles in A_k arrive to the right of position $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ for all $k \in [\![1, \delta_3\ell_N]\!]$. Thus all time-$t$ descendants of a particle that makes such a jump will be in the set G. For $k \in [\![1, \delta_3\ell_N]\!]$, let $\mathcal{M}'_k$ denote the set of such jumps:

$$\mathcal{M}'_k := \{(i, b, S_k) : X_{i,b,S_k} > \hat{a}_{n,N} \text{ and } i \in A_k\}.$$

Suppose for all $k \in [\![1, \delta_3\ell_N]\!]$, all particles descending from the set $\mathcal{M}'_k$ survive until time t. Then the total number of such descendants will be

$$\left| \bigcup_{k \in [\![1, \delta_3\ell_N]\!]} \bigcup_{(i,b,S_k) \in \mathcal{M}'_k} \mathcal{N}^b_{i,S_k}(t) \right| = \sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N - 1} \sum_{i \in A_k, b \in \{1,2\}} \mathbb{1}_{\{X_{i,b,S_k} > \hat{a}_{n,N}\}}. \tag{2.5.15}$$

The first term in the sum is the number of time-t descendants of a particle at time $S_k + 1 = T_n + \ell_N - k + 1$, and the second sum gives the number of jumps of size greater than $\hat{a}_{n,N}$ from particles in A_k.

If instead there exists $k \in [\![1, \delta_3\ell_N]\!]$ such that not every particle descending from a jump in $\mathcal{M}_k'$ survives until time t, then there must be N particles to the right of (or at) $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ at some time $s \le t$ (and therefore at time t, by monotonicity). We conclude the following lower bound:

$$|G| \ge \min\left(N, \sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N-1} \sum_{i\in A_k, b\in\{1,2\}} \mathbb{1}_{\{X_{i,b,S_k}>\hat{a}_{n,N}\}}\right). \tag{2.5.16}$$

Let $\xi_{j,k} \sim \mathrm{Ber}(h(\hat{a}_{n,N})^{-1})$ be i.i.d. random variables, by which we mean that

$$\mathbb{P}_{T_n}(\xi_{j,k}=1) = \frac{1}{h(\hat{a}_{n,N})} = 1 - \mathbb{P}_{T_n}(\xi_{j,k}=0) \quad \text{for all } k, j \in \mathbb{N}.$$

The indicator random variables in (2.5.16) all have this distribution. Thus by (2.5.16),

$$\begin{aligned}\mathbb{P}_{T_n}(\{|G|<N\}\cap\mathcal{E}) &\le \mathbb{P}_{T_n}\left(\left\{\sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N-1} \sum_{i\in A_k, b\in\{1,2\}} \mathbb{1}_{\{X_{i,b,S_k}>\hat{a}_{n,N}\}} < N\right\}\cap\mathcal{E}\right)\\ &\le \mathbb{P}_{T_n}\left(\sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N-1} \sum_{j=1}^{2^{\ell_N-k}} \xi_{j,k} < N\right),\end{aligned} \tag{2.5.17}$$

since on the event $\mathcal{E}$ there are at least 2^{ℓ_N-k} jumps from the set A_k for each $k \in [\![1, \delta_3\ell_N]\!]$.

We will use the concentration inequality from [35, Theorem 2.3(c)] to estimate the right-hand side of (2.5.17). As the inequality applies for independent random variables taking values in $[0,1]$, we consider the random variables $2^{-\delta_3\ell_N+k}\xi_{j,k} \in [0,1]$ for $k \in [\![1, \delta_3\ell_N]\!]$ and $j \in [2^{\ell_N-k}]$. Let μ denote the expectation of the sum of these random variables over k and j:

$$\mu := \mathbb{E}_{T_n}\left[\sum_{k=1}^{\delta_3\ell_N} 2^{-\delta_3\ell_N+k} \sum_{j=1}^{2^{\ell_N-k}} \xi_{j,k}\right] = \sum_{k=1}^{\delta_3\ell_N} 2^{-\delta_3\ell_N+k} \frac{2^{\ell_N-k}}{h(\hat{a}_{n,N})} \ge \frac{\delta_3\ell_N N^{1-\delta_3}}{h(\hat{a}_{n,N})} \ge 4N^{1-\delta_3-\theta_{n,N}} \tag{2.5.18}$$

for N sufficiently large, where the last inequality holds because $h(\hat{a}_{n,N}) \le 2\delta_1 N^{\theta_{n,N}}\ell_N$ by (2.5.7) for N sufficiently large, and because we chose $\delta_3/\delta_1 \ge 8$. Thus

$$\begin{aligned}\mathbb{P}_{T_n}\left(\sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N-1} \sum_{j=1}^{2^{\ell_N-k}} \xi_{j,k} < N\right) &\le \mathbb{P}_{T_n}\left(\sum_{k=1}^{\delta_3\ell_N} 2^{-\delta_3\ell_N+k} \sum_{j=1}^{2^{\ell_N-k}} \xi_{j,k} < 2N^{1-\delta_3-\theta_{n,N}}\right)\\ &\le \mathbb{P}_{T_n}\left(\sum_{k=1}^{\delta_3\ell_N} 2^{-\delta_3\ell_N+k} \sum_{j=1}^{2^{\ell_N-k}} \xi_{j,k} < \frac{1}{2}\mu\right)\end{aligned}$$

for N sufficiently large, where in the first inequality we multiply by $2^{1-(\delta_3+\theta_{n,N})\ell_N}$ to get terms in $[0,1]$ in the sum and notice that $2^{-\ell_N} \le N^{-1}$, and the second inequality holds by (2.5.18). We now apply the concentration inequality from [35, Theorem 2.3(c)] to the independent random variables $2^{-\delta_3\ell_N+k}\xi_{j,k} \in [0,1]$ on the right-hand side above, giving that

$$\mathbb{P}_{T_n}\left(\sum_{k=1}^{\delta_3\ell_N} 2^{k+\theta_{n,N}\ell_N-1}\sum_{j=1}^{2^{\ell_N-k}}\xi_{j,k} < N\right) \le e^{-\mu/8} \le e^{-\frac{1}{2}N^{\delta-\delta_3}}, \tag{2.5.19}$$

where in the second inequality we use (2.5.18) again and that $\theta_{n,N} \le 1-\delta$ by (2.5.5) and since $T_n \ge t_2 + \delta\ell_N$ by (2.5.2). Now putting (2.5.17) and (2.5.19) together, since $\delta - \delta_3 > \delta - \delta' > 0$ we conclude that

$$\mathbb{P}_{T_n}(\{|G| < N\} \cap \mathcal{E}) < \frac{\eta}{200K} \tag{2.5.20}$$

for N sufficiently large.

In case (b), $\mathcal{E}^c$ deterministically implies that $|G| = N$. Indeed, if $\mathcal{E}^c$ occurs then it follows that there exists $k_0 \in [\![1, \delta_3\ell_N]\!]$ such that $|G_{\mathcal{X}_N(T_n)+\hat{a}_{n,N}}(S_{k_0})| > \frac{1}{2}2^{\ell_N-k_0}$. Recall that $S_{k_0} = T_n + \ell_N - k_0$. Then by Lemma 2.2.3 we have

$$|G| \ge \min\left(N, \tfrac{1}{2}2^{\ell_N-k_0}2^{k_0+\theta_{n,N}\ell_N}\right) = N \tag{2.5.21}$$

for N sufficiently large, because $\theta_{n,N} \ge \delta$ by (2.5.5) and (2.5.2), and since we are assuming $T_n < t_1$. Thus for N sufficiently large,

$$\mathbb{P}_{T_n}(\{|G| < N\} \cap \mathcal{E}^c) = 0,$$

which together with (2.5.20) and (2.5.14) concludes the proof. □

Now we are ready to prove Lemma 2.5.1. Corollary 2.4.5 tells us that paths cannot move a large distance without having jumps which have size at least the order of magnitude of that large distance. So Lemma 2.5.3 and Corollary 2.4.5 together will show that paths without medium jumps cannot survive until time t with high probability.

Proof of Lemma 2.5.1. We partition the event in Lemma 2.5.1 based on the position of the leftmost particle:

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} :\ P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right) \\
&= \mathbb{P}_{T_n}\left(\left\{\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} : P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right\} \cap \{\mathcal{X}_1(t) < \mathcal{X}_N(T_n) + \hat{a}_{n,N}\}\right) \\
&\quad + \mathbb{P}_{T_n}\left(\left\{\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} : P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right\} \cap \{\mathcal{X}_1(t) \ge \mathcal{X}_N(T_n) + \hat{a}_{n,N}\}\right).
\end{aligned} \tag{2.5.22}$$

This will be useful, because from Lemma 2.5.3 we know that the leftmost particle at time t is to the right of (or at) $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ with high probability. Hence it is enough to focus on the second term on the right-hand side of (2.5.22), and show that with high probability, paths cannot move beyond $\mathcal{X}_N(T_n) + \hat{a}_{n,N}$ without medium jumps.

Assume that the event in the second term on the right-hand side of (2.5.22) occurs with $i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$ and $k \in \mathcal{N}_{i,n}$, and so we have $P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset$ and $\mathcal{X}_k(t) \geq \mathcal{X}_1(t) \geq \mathcal{X}_N(T_n) + \hat{a}_{n,N}$. Note that particle (k,t) is a descendant of particle (N,T_n) as well. The path between these two particles has to move distance at least $\hat{a}_{n,N}$. Thus one of the following must happen. Either the path between particles (N,T_n) and (k,t) moves $\hat{a}_{n,N}$ even without medium jumps, or there must be a medium jump on this path. In the latter case the medium jump must be in the time interval $[T_n, T_n^{\varepsilon_N} - 1]$, because we assumed $P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset$. This leads to the following upper bound:

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\left\{\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} : \, P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right\} \cap \{\mathcal{X}_1(t) \geq \mathcal{X}_N(T_n) + \hat{a}_{n,N}\}\right) \\
&\leq \mathbb{P}_{T_n}\left(\exists k \in \mathcal{N}_{N,T_n}(t) : \sum_{(i,b,s)\in P^{k,t}_{N,T_n}} X_{i,b,s}\mathbb{1}_{\{X_{i,b,s}\leq \delta_2 \hat{a}_{n,N}\}} \geq \hat{a}_{n,N}\right) \\
&\qquad + \mathbb{P}_{T_n}\left(\exists s \in [\![T_n, T_n^{\varepsilon_N} - 1]\!],\ i \in \mathcal{N}_{N,T_n}(s) \text{ and } b \in \{1,2\} : \, X_{i,b,s} > \delta_2 \hat{a}_{n,N}\right) \\
&\leq CN^{-1} + \mathbb{P}_{T_n}\left(\exists s \in [\![T_n, T_n^{\varepsilon_N} - 1]\!],\ i \in \mathcal{N}_{N,T_n}(s) \text{ and } b \in \{1,2\} : \, X_{i,b,s} > \delta_2 \hat{a}_{n,N}\right)
\end{aligned} \tag{2.5.23}$$

for N sufficiently large, where the second inequality holds for some constant $C > 0$ because of Corollary 2.4.5 applied with $x_N = \hat{a}_{n,N}$, $r = \delta_2$ and $\lambda = \delta/(2\alpha)$. To check the conditions of Corollary 2.4.5 we first notice that we chose $\delta_2 < \delta^2$, and claim that $\delta^2 < 1 \wedge \frac{\delta(1\wedge\alpha)}{96\alpha}$. Indeed, at the beginning of Section 2.5 we chose δ together with the other constants η, K, γ, ρ, $c_1, \ldots, c_6$ satisfying (a)-(j). From (h) and (g) we have $\delta \leq \rho \leq \frac{c_1(1\wedge\alpha)}{100\alpha}$, and since c_1 is certainly smaller than 1 (for example by (2.4.26) and (2.4.24)) the claim follows. Regarding the condition that $x_N > N^{\lambda}$, we have $\hat{a}_{n,N} > N^{\theta_{n,N}/2\alpha} \geq N^{\delta/2\alpha}$ for N sufficiently large, where the first inequality follows by (2.5.7) and Lemma 2.4.2 by the same argument as for (2.4.17) and (2.4.38), and the second inequality holds because $\theta_{n,N} \geq \delta$ by (2.5.5), (2.5.2) and since we are assuming $T_n < t_1$.

Next we use a union bound to control the second term on the right-hand side of (2.5.23), using that there are at most $2 \cdot 2^k$ jumps descending from particle (N,T_n) at time $T_n + k$. We have

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\exists s \in [\![T_n, T_n^{\varepsilon_N} - 1]\!],\ i \in \mathcal{N}_{N,T_n}(s) \text{ and } b \in \{1,2\} : \, X_{i,b,s} > \delta_2 \hat{a}_{n,N}\right) \\
&\qquad \leq \sum_{k=0}^{\varepsilon_N \ell_N - 1} \frac{2\cdot 2^k}{h(\delta_2 \hat{a}_{n,N})} < \frac{2^{1+\varepsilon_N \ell_N}}{h(\delta_2 \hat{a}_{n,N})} \leq \frac{8N^{\varepsilon_N}}{\delta_2^{\alpha}\delta_1 N^{\theta_{n,N}}\ell_N} \leq \frac{8}{\delta_2^{\alpha}\delta_1} N^{\varepsilon_N - \delta}
\end{aligned} \tag{2.5.24}$$

for N sufficiently large, where in the third inequality we use that $2^{\varepsilon_N \ell_N} \leq 2N^{\varepsilon_N}$ for N sufficiently large, and that $h(\delta_2 \hat{a}_{n,N}) \geq \delta_2^{\alpha}\delta_1 N^{\theta_{n,N}}\ell_N/2$ for N sufficiently large because of (2.1.2) and (2.5.7), and in the fourth inequality we use that $\theta_{n,N} \geq \delta$ by (2.5.5), (2.5.2) and since we are assuming $T_n < t_1$.

Note that we have $\varepsilon_N < \delta/2$ for N sufficiently large by our assumptions in (2.5.1). Therefore, by (2.5.22), Lemma 2.5.3, (2.5.23), and (2.5.24) we conclude that

$$\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), k \in \mathcal{N}_{i,n} :\ P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} = \emptyset\right) < \frac{\eta}{100K}$$

for N sufficiently large. □

2.5.5 The number of descendants of medium jumps: proof of Lemma 2.5.2

Proof of Lemma 2.5.2. We partition the time interval $[T_n^{\varepsilon_N}, t-1]$ into two subintervals, and look at the number of medium jumps and the number of time-t descendants of the medium jumps. Let

$$I_1 := [T_n^{\varepsilon_N}, t_1 + 2\varepsilon_N \ell_N - 1] \quad \text{and} \quad I_2 := [t_1 + 2\varepsilon_N \ell_N, t-1]$$

be the two intervals, and let A^i_j denote the set of particles in $\mathcal{N}_{i,n}$ which have a medium jump in their ancestral lines which happened in the time interval I_j:

$$A^i_j := \left\{k \in \mathcal{N}_{i,n} :\ P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}^{I_j}_{n,N} \neq \emptyset\right\}, \quad i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}), \quad j \in \{1,2\}. \tag{2.5.25}$$

If there is a medium jump in I_1, then there may be many, possibly of order N, particles at time t descending from this medium jump. However, we will see that with high probability there are no medium jumps at all in I_1: particle (N, T_n) does not have enough descendants by the end of I_1 for any to have made a medium jump. In contrast, in the second interval there are many particles to make medium jumps (although not more than N at any one time), but there is less time to produce many descendants by time t. Indeed, for each $i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$ the expected number of time-t descendants of $(i, T_n^{\varepsilon_N})$ whose path has a medium jump in I_2 is of order $N^{1-\varepsilon_N}$. Using a concentration result from [35], we will see that the number of descendants itself (rather than the expected number) is of order $N^{1-\varepsilon_N}$ with high probability, and therefore for each i, the total contribution of A^i_1 and A^i_2 is $o(N)$ with high probability. With the above strategy in mind, we give the following upper bound on the probability in the statement of Lemma 2.5.2, using (2.5.4):

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) :\ D_{i,n} > \nu N \text{ and } P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} \neq \emptyset\ \forall k \in \mathcal{N}_{i,n}\right) \\
&\quad \leq \mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) :\ \#\left\{k \in \mathcal{N}_{i,n} :\ P^{k,t}_{i,T_n^{\varepsilon_N}} \cap \mathcal{M}_{n,N} \neq \emptyset\right\} > \nu N\right) \\
&\quad = \mathbb{P}_{T_n}(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) :\ |A^i_1 \cup A^i_2| > \nu N) \\
&\quad \leq \mathbb{P}_{T_n}(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) :\ A^i_1 \neq \emptyset) + \mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) :\ |A^i_2| > CN^{1-\varepsilon_N}\right)
\end{aligned} \tag{2.5.26}$$

for N sufficiently large and any constant C, since $\varepsilon_N \ell_N \to \infty$ as $N \to \infty$ by our choice of ε_N in (2.5.1).

We let $\tilde{I}_1 := [T_n, t_1 + 2\varepsilon_N \ell_N - 1] \supset I_1$. It is enough to bound the first term on the right-hand side of (2.5.26) by the probability that any of the descendants of particle (N, T_n) makes a medium jump by time $t_1 + 2\varepsilon_N \ell_N - 1$:

$$\mathbb{P}_{T_n}(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : A_1^i \neq \emptyset) \leq \mathbb{P}_{T_n}\left(\exists (j,b,s) \in \mathcal{M}_{n,N}^{\tilde{I}_1} : (N, T_n) \lesssim (j,s)\right). \quad (2.5.27)$$

This probability will be very small, as the total number of descendants of (N, T_n) in the time interval $\tilde{I}_1$ is not large enough to see jumps of order $\hat{a}_{n,N}$. Indeed, applying a union bound over the jumps made by descendants of (N, T_n) at times $T_n + k$ shows that the right-hand side of (2.5.27) is at most

$$\sum_{k=0}^{(\theta_{n,N}+2\varepsilon_N)\ell_N - 1} \frac{2 \cdot 2^k}{h(\delta_2 \hat{a}_{n,N})} \leq 2 \cdot 2^{(\theta_{n,N}+2\varepsilon_N)\ell_N} \frac{2}{\delta_2^\alpha \delta_1 N^{\theta_{n,N}} \ell_N} \leq \frac{8}{\delta_2^\alpha \delta_1} \ell_N^{-1/2} \quad (2.5.28)$$

for N sufficiently large, where in the first inequality we use the fact that $h(\delta_2 \hat{a}_{n,N}) \geq \delta_2^\alpha \delta_1 N^{\theta_{n,N}} \ell_N / 2$ for N sufficiently large by (2.1.2) and (2.5.7), and in the second inequality we use the assumption on ε_N in (2.5.1), and that $2^{\theta_{n,N}\ell_N} \leq 2N^{\theta_{n,N}}$.

For the second term on the right-hand side of (2.5.26) we will give an upper bound using the concentration inequality from [35, Theorem 2.3(b)]. First we bound $|A_2^i|$ for any $i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N})$:

$$|A_2^i| \leq \sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N - 1} \sum_{j \in \mathcal{N}_{i,T_n^{\varepsilon_N}}(T_n+k), b \in \{1,2\}} \mathbb{1}_{\{X_{j,b,T_n+k} > \delta_2 \hat{a}_{n,N}\}} |\mathcal{N}_{j,T_n+k}^b(t)|, \quad (2.5.29)$$

where we sum up the number of time-t descendants of every particle descended from $(i, T_n^{\varepsilon_N})$ which made a jump of size greater than $\delta_2 \hat{a}_{n,N}$ at a time $T_n + k$ in the time interval I_2. Now let $\xi_{j,k}^i \sim \mathrm{Ber}(h(\delta_2 \hat{a}_{n,N})^{-1})$ be i.i.d. random variables, by which we mean that

$$\mathbb{P}_{T_n}(\xi_{j,k}^i = 1) = h(\delta_2 \hat{a}_{n,N})^{-1} = 1 - \mathbb{P}_{T_n}(\xi_{j,k}^i = 0),$$

for all $i, j, k \in \mathbb{N}$. The indicator random variables in (2.5.29) all have this distribution. Considering that we have $|\mathcal{N}_{i,T_n^{\varepsilon_N}}(T_n + k)| \leq \min(N, 2^{k - \varepsilon_N \ell_N})$ and $|\mathcal{N}_{j,T_n+k}^b(t)| \leq 2^{(1+\theta_{n,N})\ell_N - k - 1} \leq 2^{(1+\theta_{n,N})\ell_N - k}$ for all $k \in [\![(\theta_{n,N} + 2\varepsilon_N)\ell_N, (1 + \theta_{n,N})\ell_N - 1]\!]$, and since $\mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) \leq 2^{\varepsilon_N \ell_N}$, we obtain the following upper bound from (2.5.29):

$$\begin{aligned}
&\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : |A_2^i| > CN^{1-\varepsilon_N}\right) \\
&\leq \mathbb{P}_{T_n}\left(\exists i \in [2^{\varepsilon_N \ell_N}] : \sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N} 2^{(1+\theta_{n,N})\ell_N - k} \sum_{j=1}^{2\min(N, 2^{k-\varepsilon_N \ell_N})} \xi_{j,k}^i > CN^{1-\varepsilon_N}\right) \\
&\leq 2^{\varepsilon_N \ell_N} \mathbb{P}_{T_n}\left(\sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N} 2^{(1+\theta_{n,N})\ell_N - k} \sum_{j=1}^{2\min(N, 2^{k-\varepsilon_N \ell_N})} \xi_{j,k}^1 > CN^{1-\varepsilon_N}\right) \quad (2.5.30)
\end{aligned}$$

by a union bound.

Now [35, Theorem 2.3(b)] applies for independent random variables taking values in $[0,1]$, so we consider the random variables $2^{(2\varepsilon_N+\theta_{n,N})\ell_N-k}\xi^1_{j,k} \in [0,1]$ for each k and j in the sum. Let μ denote the expectation of the sum of these random variables over k and j:

$$\begin{aligned}\mu &:= \mathbb{E}_{T_n}\left[\sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N} 2^{(2\varepsilon_N+\theta_{n,N})\ell_N-k} \sum_{j=1}^{2\min(N,2^{k-\varepsilon_N\ell_N})} \xi^1_{j,k}\right] \\ &= \mathbb{E}_{T_n}\left[\sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\varepsilon_N)\ell_N-1} 2^{(2\varepsilon_N+\theta_{n,N})\ell_N-k}\frac{2^{k-\varepsilon_N\ell_N+1}}{h(\delta_2\hat{a}_{n,N})}\right] \\ &\qquad + \mathbb{E}_{T_n}\left[\sum_{k=(1+\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N} 2^{(2\varepsilon_N+\theta_{n,N})\ell_N-k}\frac{2N}{h(\delta_2\hat{a}_{n,N})}\right]. \end{aligned} \tag{2.5.31}$$

Now considering that for N sufficiently large, $\delta_2^\alpha\delta_1 N^{\theta_{n,N}}\ell_N/2 \le h(\delta_2\hat{a}_{n,N}) \le 2\delta_2^\alpha\delta_1 N^{\theta_{n,N}}\ell_N$ by (2.1.2) and (2.5.7), that $N^{2\varepsilon_N+\theta_{n,N}} \le 2^{(2\varepsilon_N+\theta_{n,N})\ell_N} \le 4N^{2\varepsilon_N+\theta_{n,N}}$, that $\delta \le \theta_{n,N} \le 1-\delta$ and that $\varepsilon_N < \delta/4$ for N sufficiently large, it can be seen that we have

$$K_1 N^{\varepsilon_N} \le \mu \le K_2 N^{\varepsilon_N}, \tag{2.5.32}$$

for some constants $K_1, K_2 > 0$. Then, if we multiply both sides of the sum in (2.5.30) by $2^{(2\varepsilon_N-1)\ell_N}$ and use that $2^{(2\varepsilon_N-1)\ell_N} \ge N^{2\varepsilon_N-1}/2$, we get

$$\begin{aligned}&\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : |A_2^i| > CN^{1-\varepsilon_N}\right) \\ &\qquad \le 2^{\varepsilon_N\ell_N}\mathbb{P}_{T_n}\left(\sum_{k=(\theta_{n,N}+2\varepsilon_N)\ell_N}^{(1+\theta_{n,N})\ell_N} 2^{(2\varepsilon_N+\theta_{n,N})\ell_N-k}\sum_{j=1}^{2\min(N,2^{k-\varepsilon_N\ell_N})}\xi^1_{j,k} > \frac{1}{2}CN^{\varepsilon_N}\right).\end{aligned}$$

By (2.5.32) we have $\mu \ge K_1 N^{\varepsilon_N}$, and we can choose $C > 3K_2$ so that $\frac{1}{2}CN^{\varepsilon_N} \ge \frac{3}{2}\mu$ for N sufficiently large. Then by [35, Theorem 2.3(b)] we have for N sufficiently large,

$$\mathbb{P}_{T_n}\left(\exists i \in \mathcal{N}_{N,T_n}(T_n^{\varepsilon_N}) : |A_2^i| > CN^{1-\varepsilon_N}\right) \le 2N^{\varepsilon_N}\exp\left(-\frac{\frac{1}{4}K_1N^{\varepsilon_N}}{2(1+\frac{1}{6})}\right), \tag{2.5.33}$$

which is small if N is large, by our choice of ε_N in (2.5.1). Then by (2.5.26), (2.5.27), (2.5.28) and (2.5.33) we conclude Lemma 2.5.2. □

2.6 Proofs of Propositions 2.2.1 and 2.2.2

In Proposition 2.2.1 we need to prove that for any interval of the form $[t_2 + \lceil s_1\ell_N\rceil, t_2 + \lceil s_2\ell_N\rceil]$ with $0 < s_1 < s_2 < 1$, the probability that the time of the common ancestor T is in this interval is bounded away from 0 for large N. The main idea of the proof is that if there is a big jump in the time interval $[t_2 + \lceil s_1\ell_N\rceil, t_2 + \lceil s_2\ell_N\rceil]$ which is much larger than any other jump in the time interval $[t_3, t_1]$, then that big jump will break the record, and we will have $T \in [t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil]$.

More precisely, let $r > 0$ be as in Proposition 2.2.1. We will ask that a particle performs a jump larger than $(r+3)a_N$ at some time $s^* \in [t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil)$, and all the other jumps in the time interval $[t_3, t_1]$ are smaller than a_N. We will show that this happens with a probability bounded below by a positive constant (independent of N).

Suppose the above event occurs, and also the events $\mathcal{C}_3$ and $\mathcal{C}_4$ occur. Then we will also see that $d(\mathcal{X}(s^*)) \leq (1 + c_1)a_N$. This will imply that the particle which makes the jump larger than $(r+3)a_N$ at time s^* breaks the record, and it will lead by more than roughly $(r+2)a_N$ at time s^*+1. As a result, the tribe of this particle will lead between times s^*+1 and t_1, because we assumed that all jumps in $[s^*+1, t_1)$ are smaller than a_N. Moreover, particles not in the leading tribe cannot get closer than ra_N to the leading tribe by time t_1; therefore, we will conclude $d(\mathcal{X}(t_1)) \geq ra_N$ as well.

The following lemma will be useful for proving the above statements.

Lemma 2.6.1. *Take $\rho, c_1 > 0$. Then for $N \geq 2$ and $t > 4\ell_N$, for all $s_0 \in [t_4, t_1]$ and $r_0 > 0$, on the event $\mathcal{C}_3 \cap \mathcal{C}_4$,*

$$\{X_{i,b,s} \leq r_0 a_N \ \forall (i,b,s) \in [N] \times \{1,2\} \times [\![s_0, s_0 + \ell_N - 1]\!]\} \subseteq \{d(\mathcal{X}(s_0 + \ell_N)) \leq (r_0 + c_1)a_N\},$$

where the events $\mathcal{C}_3$ and $\mathcal{C}_4$ are defined in (2.3.13) *and* (2.3.14) *respectively.*

Proof. Let $\mathcal{G}_1$ denote the event on the left-hand side in the statement of the lemma:

$$\mathcal{G}_1 := \{X_{i,b,s} \leq r_0 a_N \ \forall (i,b,s) \in [N] \times \{1,2\} \times [\![s_0, s_0 + \ell_N - 1]\!]\}.$$

Let $j \in [N]$ be arbitrary, and let $i = \zeta_{j, s_0+\ell_N}(s_0)$. Then, on the event $\mathcal{C}_3$, we have $|B_N \cap P_{i,s_0}^{j,s_0+\ell_N}| \leq 1$, and on the event $\mathcal{C}_4$, no particle moves further than $c_1 a_N$ once big jumps have been removed from its path. Thus, on the event $\mathcal{C}_3 \cap \mathcal{C}_4 \cap \mathcal{G}_1$,

$$\mathcal{X}_j(s_0 + \ell_N) \leq \mathcal{X}_i(s_0) + c_1 a_N + \sum_{(i',b',s') \in B_N \cap P_{i,s_0}^{j,s_0+\ell_N}} X_{i',b',s'} \leq \mathcal{X}_N(s_0) + (r_0 + c_1)a_N.$$

But by Lemma 2.2.3, we have $\mathcal{X}_N(s_0) \leq \mathcal{X}_1(s_0 + \ell_N)$, and the result follows. □

Proof of Proposition 2.2.1. Recall the definition of $\mathcal{A}_2'$ from (2.2.5), and consider a uniform sample of M particles at time t with indices $\mathcal{P}_1, \ldots, \mathcal{P}_M$. Also recall the definitions of $T(\rho)$ in (2.2.17) and $T^{\varepsilon_N}(\rho)$ in (2.2.19). For any $\rho > 0$ we have

$$\{T(\rho) \in [t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil]\} \cap \left\{\zeta_{\mathcal{P}_j,t}(T(\rho)) = N \ \forall j \in [M]\right\} \cap \left\{\zeta_{\mathcal{P}_j,t}(T^{\varepsilon_N}(\rho)) \neq \zeta_{\mathcal{P}_l,t}(T^{\varepsilon_N}(\rho)) \ \forall j,l \in [M],\ j \neq l\right\} \subseteq \mathcal{A}_2'. \quad (2.6.1)$$

For $r > 0$, we define $\mathcal{A}_3'$ as a modification of the event $\mathcal{A}_3$ from (2.2.22):

$$\begin{aligned}\mathcal{A}_3' = \mathcal{A}_3'(t, N, \rho, \gamma, r, s_1, s_2) := \{T(\rho) \in [t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil]\} \\ \cap \left\{|\mathcal{N}_{N,T(\rho)}(t)| \geq N - N^{1-\gamma}\right\} \cap \{d(\mathcal{X}(t_1)) \geq r a_N\}.\end{aligned} \tag{2.6.2}$$

We also define the set of jumps in the time interval $[t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil)$ which are larger than $(r+3)a_N$:

$$B_N'(t, r, s_1, s_2) := \left\{ \begin{array}{l} (i, b, s) \in [N] \times \{1, 2\} \times [\![t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil - 1]\!] : \\ X_{i,b,s} > (r+3) a_N \end{array} \right\}, \tag{2.6.3}$$

and the event $\mathcal{G}$, which says that there is only one jump in the set B_N', and every other jump is smaller than a_N during the time interval $[t_3, t_1 - 1]$:

$$\mathcal{G} = \mathcal{G}(t, N, r, s_1, s_2) := \left\{ \begin{array}{l} |B_N'| = 1 \text{ and } X_{i,b,s} \leq a_N, \\ \forall (i, b, s) \in ([N] \times \{1, 2\} \times [t_3, t_1 - 1]) \setminus B_N' \end{array} \right\}. \tag{2.6.4}$$

Fix $0 < s_1 < s_2 < 1$, $M \in \mathbb{N}$ and $r > 0$. Choose $\pi_{r, s_2 - s_1} > 0$ such that

$$\pi_{r, s_2 - s_1} < \frac{s_2 - s_1}{8(r+3)^\alpha} \cdot e^{-8}, \tag{2.6.5}$$

and then $\eta > 0$ sufficiently small that it satisfies (2.4.23) and

$$5\eta < \frac{s_2 - s_1}{8(r+3)^\alpha} \cdot e^{-8} - \pi_{r, s_2 - s_1}. \tag{2.6.6}$$

Then choose the constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ such that they satisfy (a)-(j). Recall from Section 2.4.4 that this implies the properties in (2.3.2)-(2.3.5) and (2.4.24)-(2.4.31) also hold for η and $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$. Let $0 < \nu < \eta / M^2$.

In the course of the proof we will use the events $\mathcal{A}_3$ and $\mathcal{A}_4$ from (2.2.22) and (2.2.23), and we will show the following for N sufficiently large and $t > 4\ell_N$:

1. $\mathbb{P}\left((\mathcal{A}_2')^c \cup \{d(\mathcal{X}(t_1)) < r a_N\}\right) \leq \mathbb{P}((\mathcal{A}_3')^c) + \mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta$
2. $\bigcap_{j=2}^7 \mathcal{C}_j \cap \bigcap_{i=1}^5 \mathcal{D}_i \cap \mathcal{G} \subseteq \mathcal{A}_3'$
3. $\mathbb{P}(\mathcal{G}) \geq \frac{s_2 - s_1}{8(r+3)^\alpha} \cdot e^{-8}$
4. $\mathbb{P}\left((\mathcal{A}_2')^c \cup \{d(\mathcal{X}(t_1)) < r a_N\}\right) \leq 1 - \pi_{r, s_2 - s_1}$.

We start by proving step 1. Notice that with our choices of constants, the conditions of Lemma 2.2.4 hold. Therefore, we know

$$\mathbb{P}(\exists j, l \in [M],\ j \neq l :\ \zeta_{\mathcal{P}_j, t}(T^{\varepsilon_N}) = \zeta_{\mathcal{P}_l, t}(T^{\varepsilon_N})) \leq \mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta/2, \tag{2.6.7}$$

for N sufficiently large. Hence, because of (2.6.1), in order to prove step 1 it remains to show that

$$\begin{aligned}&\mathbb{P}\left(\{T(\rho)\notin[t_2+\lceil s_1\ell_N\rceil,t_2+\lceil s_2\ell_N\rceil]\}\cup\left\{\exists j\in[M]:\,\zeta_{\mathcal{P}_j,t}(T)\neq N\right\}\cup\{d(\mathcal{X}(t_1))<ra_N\}\right)\\&\leq\mathbb{P}((\mathcal{A}_3')^c)+\eta/2,\end{aligned}\tag{2.6.8}$$

for N sufficiently large. This follows similarly to the proof of (2.2.25). Partitioning the event on the left-hand side of (2.6.8) using the event $\mathcal{A}_3'$, and then conditioning on $\mathcal{F}_t$, we obtain

$$\begin{aligned}&\mathbb{P}\left(\{T(\rho)\notin[t_2+\lceil s_1\ell_N\rceil,t_2+\lceil s_2\ell_N\rceil]\}\cup\left\{\exists j\in[M]:\,\zeta_{\mathcal{P}_j,t}(T)\neq N\right\}\cup\{d(\mathcal{X}(t_1))<ra_N\}\right)\\&\leq\;\mathbb{E}\left[\mathbb{1}_{\mathcal{A}_3'}\mathbb{P}(\exists j\in[M]:\,\zeta_{\mathcal{P}_j,t}(T)\neq N\mid\mathcal{F}_t)\right]+\mathbb{P}\left((\mathcal{A}_3')^c\right)\end{aligned}\tag{2.6.9}$$

where we use that if $\mathcal{A}_3'$ occurs then $T(\rho)\in[t_2+\lceil s_1\ell_N\rceil,t_2+\lceil s_2\ell_N\rceil]$ and $d(\mathcal{X}(t_1))\geq ra_N$, and that $\mathcal{A}_3'$ is $\mathcal{F}_t$-measurable. Now, on the event $\mathcal{A}_3'$, at most $N^{1-\gamma}$ time-t particles are not descended from (N,T), and therefore a union bound on the uniformly chosen sample (which is not $\mathcal{F}_t$-measurable) shows that the right-hand side of (2.6.9) is at most $MN^{1-\gamma}/N+\mathbb{P}\left((\mathcal{A}_3')^c\right)$. This implies (2.6.8) for N sufficiently large, and by (2.6.7) and (2.6.8) we are done with step 1.

We next prove step 2. Assume the event $\bigcap_{j=2}^{7}\mathcal{C}_j\cap\mathcal{G}$ occurs. Then there exists $(i^*,b^*,s^*)\in B_N'$ with $s^*\in[\![t_2+\lceil s_1\ell_N\rceil,t_2+\lceil s_2\ell_N\rceil-1]\!]$. We notice that every jump in the time interval $[t_3,s^*-1]$ has size at most a_N on the event $\mathcal{G}$. Thus, we can apply Lemma 2.6.1 with $s_0=s^*-\ell_N>t_3$, ρ and c_1 as chosen at the beginning of the proof, and with $r_0=1$. We then obtain

$$d(\mathcal{X}(s^*))\leq(1+c_1)a_N.\tag{2.6.10}$$

This means that a particle that makes a jump larger than $(r+3)a_N$ at time s^* must take the lead at time s^*+1. Indeed,

$$\mathcal{X}_{i^*}(s^*)+X_{i^*,b^*,s^*}>\mathcal{X}_1(s^*)+(r+3)a_N\geq\mathcal{X}_N(s^*)+(r+2-c_1)a_N,\tag{2.6.11}$$

where in the first inequality we use that $\mathcal{X}_{i^*}(s^*)\geq\mathcal{X}_1(s^*)$ and that $(i^*,b^*,s^*)\in B_N'$, and the second inequality follows by (2.6.10). Note that our choice of constants means that $\rho<r+2-c_1<r+3$ holds (it is enough that $\rho<1$ and $c_1\in(0,1)$, which certainly follow from (2.4.24) and (2.4.31)); thus we have $B_N'\subseteq B_N$, and Lemma 2.3.5(b) applies. Therefore, by Lemma 2.3.5(b), we have $(i^*,s^*)\lesssim_{b^*}(N,s^*+1)$ and

$$\mathcal{X}_{i^*}(s^*)+X_{i^*,b^*,s^*}=\mathcal{X}_N(s^*+1)>\mathcal{X}_{N-1}(s^*+1)+(r+2-c_1-\rho)a_N,\tag{2.6.12}$$

which also shows that $s^*\in\mathbf{S}_N(\rho)$, where $\mathbf{S}_N(\rho)$ is the set of times when the record is broken by a big jump (see (2.2.16)).

Now we prove that $s^*+1 = T(\rho)$ and $d(\mathcal{X}(t_1)) \geq ra_N$. Let $\hat{s} \in [\![s^*+1, t_1-1]\!]$ be arbitrary (and note that $[\![s^*+1, t_1-1]\!]$ is not empty for N sufficiently large). We will see that $\hat{s} \notin \mathbf{S}_N(\rho)$, and therefore $T(\rho) \notin [\![s^*+2, t_1]\!]$, i.e. $T(\rho) = s^*+1$.

Take $k \in [N-1]$, and assume that $j \in \mathcal{N}_{k,s^*+1}(\hat{s}+1)$. Note that $|B_N \cap P^{j,\hat{s}+1}_{k,s^*+1}| \leq 1$ by the definition of the event $\mathcal{C}_3$, and that every jump in the time interval $[s^*+1, t_1-1]$ is at most of size a_N by the definition of the event $\mathcal{G}$. Hence, by the definition of the event $\mathcal{C}_4$ we have

$$\begin{aligned}
\mathcal{X}_j(\hat{s}+1) &\leq \mathcal{X}_k(s^*+1) + c_1 a_N + \sum_{(i,b,s)\in B_N \cap P^{j,\hat{s}+1}_{k,s^*+1}} X_{i,b,s} \\
&\leq \mathcal{X}_{N-1}(s^*+1) + (c_1+1)a_N \\
&< \mathcal{X}_N(s^*+1) - (r+1-2c_1-\rho)a_N \\
&\leq \mathcal{X}_N(\hat{s}+1) - (r+1-2c_1-\rho)a_N,
\end{aligned} \tag{2.6.13}$$

where in the second inequality we also use that $k \leq N-1$, the third inequality follows by (2.6.12), and the fourth by monotonicity.

Then (2.6.13) has two consequences. First, it shows that $\mathcal{X}_j(\hat{s}+1) < \mathcal{X}_N(\hat{s}+1)$ (see e.g. (2.4.24) and (2.4.31)); thus the leader at time $\hat{s}+1$ must descend from particle (N, s^*+1); that is, $\zeta_{N,\hat{s}+1}(s^*+1) = N$. Note that we also have $X_{i,b,\hat{s}} \leq \rho a_N$ for all $i \in \mathcal{N}_{N,s^*+1}(\hat{s})$ and $b \in \{1,2\}$ by the definition of the event $\mathcal{C}_3$. We conclude that the record is not broken by a big jump at time $\hat{s}+1$, which means that $\hat{s} \notin \mathbf{S}_N(\rho)$. Since $\hat{s} \in [\![s^*+1, t_1-1]\!]$ was arbitrary, and $s^* \in \mathbf{S}_N(\rho)$, we must have $T(\rho) = s^*+1$, by the definition (2.2.17) of $T(\rho)$. Hence,

$$\bigcap_{i=2}^{7} \mathcal{C}_i \cap \mathcal{G} \subseteq \{T(\rho) \in [t_2 + \lceil s_1 \ell_N \rceil, t_2 + \lceil s_2 \ell_N \rceil]\}\,. \tag{2.6.14}$$

The second consequence of (2.6.13) is that $d(\mathcal{X}(\hat{s}+1)) > ra_N$, since $2c_1 + \rho < 1$. Indeed, we notice that since $s^*+1 > t_2$ and $\hat{s}+1 \leq t_1$, the number of descendants of particle (N, s^*+1) is strictly less than N at time $\hat{s}+1$. Thus, there exists $k \in [N-1]$ such that $\mathcal{N}_{k,s^*+1}(\hat{s}+1) \neq \emptyset$, and for such a k and for some $j \in \mathcal{N}_{k,s^*+1}(\hat{s}+1)$ the bound in (2.6.13) holds, and shows that $d(\mathcal{X}(\hat{s}+1)) > ra_N$. Since $\hat{s} \in [\![s^*+1, t_1-1]\!]$ was arbitrary we conclude

$$\bigcap_{i=2}^{7} \mathcal{C}_i \cap \mathcal{G} \subseteq \{d(\mathcal{X}(t_1)) \geq ra_N\}\,. \tag{2.6.15}$$

Propositions 2.3.11 and 2.3.2 (and the definition of $\mathcal{A}_3$ in (2.2.22)) imply for N sufficiently large that

$$\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i \subseteq \bigcap_{i=1}^{7} \mathcal{C}_i \subseteq \mathcal{A}_3 \subseteq \left\{ |\mathcal{N}_{N,T(\rho)}(t)| \geq N - N^{1-\gamma} \right\}.$$

The same statement obviously remains true if we also intersect with $\mathcal{G}$ on the left-hand side, and therefore step 2 follows by (2.6.14) and (2.6.15) and the definition of $\mathcal{A}_3'$ from (2.6.2).

For step 3, the event $\mathcal{G}$ says that out of the $4N\ell_N$ jumps occurring in the time interval $[t_3, t_1 - 1]$, there are $4N\ell_N - 1$ jumps of size at most a_N, and there is one larger than $(r+3)a_N$, which can happen any time during the time interval $[t_2 + \lceil s_1\ell_N \rceil, t_2 + \lceil s_2\ell_N \rceil)$. Using that $\lceil s_2\ell_N \rceil - 1 - \lceil s_1\ell_N \rceil \geq (s_2 - s_1)\ell_N/2$ for large N, we have for N sufficiently large,

$$\begin{aligned}
\mathbb{P}(\mathcal{G}) &\geq 2N\frac{(s_2 - s_1)}{2}\ell_N \cdot h((r+3)a_N)^{-1}\left(1 - h(a_N)^{-1}\right)^{4N\ell_N - 1} \\
&\geq \frac{(s_2 - s_1)}{2}\frac{h(a_N)}{h((r+3)a_N)} \cdot \frac{2N\ell_N}{h(a_N)} \cdot e^{-2\frac{4N\ell_N}{h(a_N)}} \\
&\geq \frac{s_2 - s_1}{8(r+3)^\alpha} \cdot e^{-8},
\end{aligned}$$

where the second inequality holds if N is sufficiently large that $1 - h(a_N)^{-1} > e^{-2h(a_N)^{-1}}$, which is possible because $h(a_N) \to \infty$ as $N \to \infty$ by (2.4.18). In the third inequality we use that $h(a_N)/h((r+3)a_N) \geq (r+3)^{-\alpha}/2$ for N large enough by (2.1.2) and (2.4.17), and that $1/2 \leq 2N\ell_N/h(a_N) \leq 2$ for N large enough by (2.4.19). This completes step 3.

For step 4, we note that we chose the constants η, γ, δ, ρ, $c_1, c_2, \ldots, c_6$, K and ν in such a way that the probability bounds in Propositions 2.2.5 and 2.2.6 and Lemma 2.4.6 hold for N sufficiently large and $t > 4\ell_N$. Hence, putting steps 1 to 3 together we conclude

$$\begin{aligned}
\mathbb{P}\left((\mathcal{A}_2')^c \cup \{d(\mathcal{X}(t_1)) < ra_N\}\right) &\leq \sum_{j=2}^{7}\mathbb{P}(\mathcal{C}_j^c) + \sum_{i=1}^{5}\mathbb{P}(\mathcal{D}_i^c) + \mathbb{P}(\mathcal{G}^c) + \mathbb{P}(\mathcal{A}_3^c) + \mathbb{P}(\mathcal{A}_4(\nu)^c) + \eta \\
&\leq 1 - \frac{s_2 - s_1}{8(r+3)^\alpha} \cdot e^{-8} + 5\eta \\
&< 1 - \pi_{r, s_2 - s_1},
\end{aligned}$$

where in the last inequality we used (2.6.6). This finishes the proof of Proposition 2.2.1. □

The proof of Proposition 2.2.2 involves some of our previous results. We will use the statement of Proposition 2.2.1 about the diameter to prove that for any fixed $r > 0$, $\mathbb{P}(d(\mathcal{X}(n)) \geq ra_N)$ can be lower bounded by a positive constant. Then the statement of Proposition 2.3.2 about the diameter shows that on the events $\mathcal{C}_1$ to $\mathcal{C}_7$ the diameter at time t_1 is greater than $c_3 a_N$, so, considering Lemma 2.4.6, we will see that the diameter is at least of order a_N at a typical time with high probability. Finally, we will conclude that the diameter is at most of order a_N with high probability using Lemma 2.6.1, and also using that jumps of size ra_N are unlikely to happen in ℓ_N time if r is very large.

Proof of Proposition 2.2.2. Take $\eta, \gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ such that they satisfy (2.4.23), (a)-(j), and therefore also (2.3.2)-(2.3.5) and (2.4.24)-(2.4.31) (and η may be arbitrarily small). Let $r > 0$ be arbitrary. Let $s_1 = 1/4$, $s_2 = 1/2$, $M = 3$. Then we take $\pi_{r,s_2-s_1} > 0$ and $N \in \mathbb{N}$ sufficiently large that the bounds in Proposition 2.2.1 and Lemma 2.4.6 and the inclusions in Propositions 2.3.2 and 2.3.11 and in Lemma 2.6.1 hold with the above constants and for all $t > 4\ell_N$. Furthermore, we assume that N is sufficiently large that

$$e^{-2h(ra_N/2)^{-1}} < 1 - h(ra_N/2)^{-1}, \tag{2.6.16}$$

$$\frac{h(a_N)}{h(ra_N/2)} \leq 2(r/2)^{-\alpha}, \tag{2.6.17}$$

and

$$\frac{2N\ell_N}{h(a_N)} \leq 2. \tag{2.6.18}$$

We can take N sufficiently large that (2.6.16), (2.6.17) and (2.6.18) hold because of (2.4.18), (2.4.17) (i.e. $a_N \to \infty$ as $N \to \infty$), (2.1.2) and (2.4.19). Having fixed N with these properties, take $n > 3\ell_N$.

First we apply Proposition 2.2.1 in the above setting with $t = n + \ell_N$ (and $t_1 = n$). The proposition implies that

$$0 < \pi_{r,s_2-s_1} < \mathbb{P}\left(d(\mathcal{X}(n)) \geq ra_N\right). \tag{2.6.19}$$

Now we prove that if r is sufficiently small then we have

$$\mathbb{P}\left(d(\mathcal{X}(n)) < ra_N\right) < \eta. \tag{2.6.20}$$

Assume that $r < c_3$, where c_3 was specified at the beginning of this proof.

Consider the events $(\mathcal{C}_j)_{j=2}^7$ and $(\mathcal{D}_i)_{i=1}^5$ with the constants $\gamma, \delta, \rho, c_1, c_2, \ldots, c_6, K$ and with $t = n + \ell_N$. By Propositions 2.3.11 and 2.3.2 we have

$$\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i \subseteq \bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \left\{d(\mathcal{X}(n)) \geq \tfrac{3}{2} c_3 a_N\right\}.$$

Therefore, since $r < c_3$, and then by Lemma 2.4.6, we have

$$\mathbb{P}(d(\mathcal{X}(n)) < ra_N) \leq \mathbb{P}(d(\mathcal{X}(n)) < \tfrac{3}{2} c_3 a_N) \leq \sum_{j=2}^{7} \mathbb{P}(\mathcal{C}_j^c) + \sum_{i=1}^{5} \mathbb{P}(\mathcal{D}_i^c) < \eta,$$

which establishes (2.6.20).

Next we prove that if r is sufficiently large then

$$\mathbb{P}\left(d(\mathcal{X}(n)) \geq r a_N\right) < \eta. \tag{2.6.21}$$

Assume $r > 1$. We apply Lemma 2.6.1 with $t = n + \ell_N$, $s_0 = n - \ell_N$ and $r_0 = r/2$. Note that by (2.4.24) and (2.4.26) we have $r_0 + c_1 < r$. Then Lemma 2.6.1 implies

$$\begin{aligned}\mathbb{P}(d(\mathcal{X}(n)) \geq r a_N) &\leq \mathbb{P}(\exists (i,b,s) \in [N] \times \{1,2\} \times [\![n-\ell_N, n-1]\!] : \ X_{i,b,s} > \tfrac{r}{2} a_N) \\ &= 1 - (1 - h(r a_N/2)^{-1})^{2N\ell_N} \\ &\leq 1 - \exp\left(-2 \frac{2N\ell_N}{h(r a_N/2)}\right) \\ &= 1 - \exp\left(-2 \frac{2N\ell_N}{h(a_N)} \frac{h(a_N)}{h(r a_N/2)}\right) \\ &\leq 1 - \exp\left(-8 (r/2)^{-\alpha}\right), \end{aligned} \tag{2.6.22}$$

where in the equality we use the tail distribution (2.1.3) for the $2N\ell_N$ jumps in the time interval $[\![n-\ell_N, n-1]\!]$, the second inequality holds by (2.6.16), and in the third we use (2.6.17) and (2.6.18). Then (2.6.22) shows that (2.6.21) holds for r sufficiently large.

Since $\eta > 0$ was arbitrarily small, (2.6.19) and (2.6.20) show the existence of p_r and (2.6.21) proves the existence of q_r as in the statement of Proposition 2.2.2, and therefore we have finished the proof of this result. □

2.7 Glossary of notation

Below we list the most frequently used notation of this paper. In the second column of the table we give a brief description, and in the third column we refer to the section or equation where the notation is defined or first appears.

Notation	**Meaning**	**Def./Sect.**
N	number of particles	Sect. 2.1.1
(i, n)	refers to the ith particle from the left at time n	Sect. 2.1.1
$\mathcal{X}_i(n)$	location of the ith particle from the left at time n	Sect. 2.1.1
h	the function $1/h$ defines the tail of the jump distribution	(2.1.3)
α	h is regularly varying with index $\alpha > 0$	(2.1.2), (2.1.3)
ℓ_N	time scale: $\ell_N = \lceil \log_2 N \rceil$	(2.1.4)
a_N	space scale: $a_N = h^{-1}(2N\ell_N)$, $h(a_N) \sim 2N\ell_N$	(2.1.5)
t	$t \in \mathbb{N}$ is an arbitrary time, we assume $t > 4\ell_N$	Sect. 2.1.3

t_i	$t_i = t - i\ell_N$, we use t_1, t_2, t_3, t_4	(2.1.7)
$X_{i,b,n}$	jump size of the bth offspring of particle (i,n)	Sect. 2.2.1
(i,b,n)	refers to the jump $X_{i,b,n}$ of the bth offspring of particle (i,n)	Sect. 2.2.4
$d(\mathcal{X}(n))$	diameter of the particle cloud at time n	(2.2.6)
$(i,n) \lesssim (j,n+k)$	particle (i,n) is the time-n ancestor of particle $(j,n+k)$	(2.2.8)
$(i,n) \lesssim_b (j,n+k)$	the bth offspring of particle (i,n) is the time-$(n+1)$ ancestor of particle $(j,n+k)$	Sect. 2.2.4
$\zeta_{i,n+k}(n)$	$\zeta_{i,n+k}(n) \in [N]$ is the index of the time-n ancestor of the particle $(i,n+k)$	(2.2.9)
$P_{i_0,n}^{i_k,n+k}$	path (sequence of jumps) between particles (i_0,n) and $(i_k,n+k)$, if $(i_0,n) \lesssim (i_k,n+k)$	(2.2.10)
$\mathcal{N}_{i,n}(n+k)$	$\mathcal{N}_{i,n}(n+k) \subseteq [N]$ is the set of time-$(n+k)$ descendants of particle (i,n)	(2.2.12)
$\mathcal{N}_{i,n}^b(n+k)$	$\mathcal{N}_{i,n}^b(n+k) \subseteq [N]$ is the set of time-$(n+k)$ descendants of the bth offspring of particle (i,n)	(2.2.13)
ρa_N	jumps of size greater than ρa_N are called big jumps	Sect. 2.2.5
B_N	set of big jumps	(2.2.14), (2.2.15)
$\mathbf{S}_N$	set of times when the record is broken by a big jump	(2.2.16)
$\hat{\mathbf{S}}_N$	times when the leader is surpassed by a big jump	(2.2.18)
T	time of the common ancestor of almost every particle at time t	Sect. 2.1.3
$T = T(\rho)$	the last time before t_1 when a particle breaks the record with a big jump	(2.2.17)
(N,T)	the leader (rightmost) particle at time T	Sect. 2.1.4
$Z_i(s)$	distance between the ith and the rightmost particle	(2.3.11)

Next, we list the events which appear throughout our main argument. We give a brief explanation of each event and refer to the equation where the event is defined. We also include short descriptions of the main results involving these events to give a summary of the major steps of the proof of Theorem 2.2.1. We write “whp” as shorthand for “with high probability”.

Event	**Meaning**	**Def./Sect.**

$\mathcal{A}_1$	Almost the whole population is close to the leftmost particle at time t.	(2.2.3)
$\mathcal{A}_2$	The genealogy of the population at time t is given by a star-shaped coalescent; there is a common ancestor at time $T \in [t_2, t_1]$.	(2.2.4)
	$\mathcal{A}_1$ and $\mathcal{A}_2$ occur whp (Theorem 2.2.1)	
$\mathcal{A}_3$	Almost every particle at time t descends from the leader at time $T \in [t_2, t_1]$.	(2.2.22)
$\mathcal{A}_4$	Shortly after time T no particle has a positive proportion of the population as descendants at time t.	(2.2.23)
	If $\mathcal{A}_3$ and $\mathcal{A}_4$ occur whp then $\mathcal{A}_2$ occurs whp (Lemma 2.2.4)	
	The event $\mathcal{A}_4$ occurs whp (Proposition 2.2.6)	
	The event $\mathcal{A}_1 \cap \mathcal{A}_3$ occurs whp (Proposition 2.2.5). This is shown using the events below.	
$\mathcal{B}_1$	There is a leading tribe, descended from the leader at time $T \in [t_2, t_1]$, which is a significant distance from the other particles at time t_1.	(2.3.7)
$\mathcal{B}_2$	Particles which are not in the leading tribe at time t_1 have $o(N)$ descendants in total at time t.	(2.3.8)
	$\mathcal{B}_1 \cap \mathcal{B}_2 \subseteq \mathcal{A}_3$ (Lemma 2.3.1)	
$\mathcal{C}_1$	A particle leads by a large distance compared to the second rightmost particle at some point in $[t_2 + 1, t_1]$.	(2.3.10)
$\mathcal{C}_2$	Particles far from the leader stay far behind or beat the leader by a lot.	(2.3.12)
$\mathcal{C}_3$	There is at most one big jump on a path of length ℓ_N.	(2.3.13)
$\mathcal{C}_4$	Paths without big jumps move very little on the a_N space scale.	(2.3.14)
$\mathcal{C}_5$	Two big jumps cannot happen at the same time.	(2.3.15)
$\mathcal{C}_6$	No big jumps happen at times very close to t_2 or t_1.	(2.3.16)
$\mathcal{C}_7$	The number of big jumps performed in $[t_4, t]$ is bounded above by a constant independent of N.	(2.3.17)
	$\bigcap_{j=1}^{7} \mathcal{C}_j \subseteq \mathcal{B}_1 \cap \mathcal{B}_2 \cap \mathcal{A}_1 \subseteq \mathcal{A}_1 \cap \mathcal{A}_3$ (Proposition 2.3.2)	
$\mathcal{D}_1$	Same as $\mathcal{C}_2$ with different constants.	(2.3.47)
$\mathcal{D}_2$	In every short interval on the ℓ_N time scale, at least one big jump larger than a certain size occurs.	(2.3.48)

$\mathcal{D}_3$	In the first half of $[t_2, t_1]$ a big jump larger than a certain size occurs.	(2.3.49)
$\mathcal{D}_4$	Shortly before time t_2, only jumps smaller than a certain size occur.	(2.3.50)
$\mathcal{D}_5$	During a short time interval, jumps of size in a certain small range do not happen.	(2.3.52)
	$\bigcap_{j=2}^{7} \mathcal{C}_j \cap \bigcap_{i=1}^{5} \mathcal{D}_i \subseteq \mathcal{C}_1$ (Proposition 2.3.11)	
	The events $\mathcal{C}_2 - \mathcal{C}_7$, $\mathcal{D}_1 - \mathcal{D}_5$ all occur whp (Lemma 2.4.6)	

Chapter 3

Speed of the particle cloud in the N-BRW with stretched exponential jump distribution

In this chapter we study the speed of the particle cloud in the N-BRW when the jump distribution has stretched exponential tails. First we state and recall the proof of the result of Bérard and Gouéré about the existence of an asymptotic speed v_N for any fixed N as time goes to infinity. Then we state and prove our theorem about the behaviour of v_N as $N \to \infty$, in the stretched exponential case. This chapter is based on joint work with Sarah Penington.

3.1 Reminder of notation

Consider the N-BRW (as defined in Chapter 2 in Section 2.2.1, but note that we will work with a different jump distribution) and recall the following notation from Chapter 2. We refer to the glossary of notation in Section 2.7 for references to the section or equation where the notation is defined or first appears.

- $[n] = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$, and $[\![a, b]\!] = [a, b] \cap \mathbb{N}_0$ for $0 \le a \le b$
- $\mathcal{X}(n) = \{\mathcal{X}_1(n) \le \cdots \le \mathcal{X}_N(n)\}$: ordered positions of the N particles at time n
- $X_{i,b,n}$: jump size of the ith particle's bth offspring at time n
- $\ell_N = \lceil \log_2 N \rceil$
- (i, n): ith particle from the left at time n
- $(i, n) \lesssim (j, n + k)$: particle (i, n) is the time-n ancestor of particle $(j, n + k)$
- $(i, n) \lesssim_b (j, n + k)$: the bth offspring of particle (i, n) is the time-$(n + 1)$ ancestor of particle $(j, n + k)$

- $P_{i,n}^{j,n+k}$: path (sequence of jumps) between particles (i,n) and $(j,n+k)$ if $(i,n) \lesssim (j,n+k)$
- $G_x(n)$: set of particles to the right of position x at time n (see (2.2.7))

We will also use the following lemmas which were already stated and proved in Chapter 2.

Lemma 3.1.1. *Let $x \in \mathbb{R}$ and $n,k \in \mathbb{N}_0$. Then*

$$|G_x(n+k)| \geq \min\Big(N,\ 2^k|G_x(n)|\Big).$$

This statement is Lemma 2.2.3 in Chapter 2, where we prove this lemma.

Lemma 3.1.2. *Suppose Y is a non-negative random variable. For $v > 0$ and $0 < K_1 < K_2 < \infty$,*

$$\begin{aligned}\mathbb{E}[\exp(vY\mathbb{1}_{\{Y\leq K_2\}})\mathbb{1}_{\{Y\geq K_1\}}]&\\ = \int_{K_1}^{K_2} ve^{vu}\mathbb{P}(Y>u)du + e^{vK_1}\mathbb{P}(Y\geq K_1) - (e^{vK_2}-1)\mathbb{P}(Y>K_2).&\end{aligned} \tag{3.1.1}$$

We proved this identity as Lemma 2.4.4.

We say that a function f is regularly varying with index $\beta \in \mathbb{R}$ if for all $y > 0$,

$$\frac{f(xy)}{f(x)} \to y^\beta \text{ as } x \to \infty, \tag{3.1.2}$$

and slowly varying if for all $y > 0$,

$$\frac{f(xy)}{f(x)} \to 1 \text{ as } x \to \infty. \tag{3.1.3}$$

Lemma 3.1.3. *Let f be a regularly varying function with index $\beta > 0$. For $\epsilon > 0$, there exist $B(\epsilon) > 1$ and $C_1(\epsilon), C_2(\epsilon) > 0$ such that*

$$\frac{1}{f(x)} \leq C_1 x^{\epsilon-\beta} \quad \textit{and} \quad f(x) \leq C_2 x^{\beta+\epsilon} \quad \forall x \geq B.$$

We proved this lemma as Lemma 2.4.2.

3.2 The main result

First we state a result of Bérard and Gouéré, which says that if the jump distribution of the N-BRW has finite mean, then for any fixed N the particle cloud has a deterministic asymptotic speed as time goes to infinity, which depends on N and on the jump distribution.

Proposition 3.2.1. *[2, Proposition 2] Consider an N-BRW with arbitrary initial configuration and with a jump distribution given by the non-negative random variable X. Assume* *that* $\mathbb{E}[X] < \infty$. *Then for any fixed* $N \in \mathbb{N}$, *there exists* $v_N \in \mathbb{R}$ *such that*

$$\lim_{n\to\infty} \frac{\mathcal{X}_1(n)}{n} = \lim_{n\to\infty} \frac{\mathcal{X}_N(n)}{n} = v_N, \tag{3.2.1}$$

almost surely and in L^1, *where* v_N *depends on the jump distribution.*

We will discuss the proof of this result in Section 3.3 following the steps in [2]. Now we turn to the main result of this chapter, which describes the behaviour of the asymptotic speed v_N as $N \to \infty$, when the jump distribution of the N-BRW has stretched exponential tails.

Let X be a random variable, and let us define the function g by letting

$$\mathbb{P}(X > x) = e^{-g(x)}, \text{ for } x \geq 0. \tag{3.2.2}$$

We assume throughout that $\mathbb{P}(X \geq 0) = 1$ and that g is regularly varying with index $\beta \in (0,1)$. We assume furthermore that the function $x \mapsto g(x)/x$ is non-increasing for x sufficiently large; that is, there exists $K > 0$ such that

$$\frac{g(x_1)}{x_1} \geq \frac{g(x_2)}{x_2} \text{ for all } x_2 > x_1 > K. \tag{3.2.3}$$

One of our arguments will be similar to the proof of Theorem 3 in [27], where the same (fairly weak) assumption is made.

Let us also define

$$L_N := g^{-1}(\log N), \tag{3.2.4}$$

for all $N \in \mathbb{N}$, where g^{-1} denotes the generalised inverse of g defined by

$$g^{-1}(x) := \inf\{y \geq 0 : \; g(y) > x\}. \tag{3.2.5}$$

It is worth keeping in mind the particular case when $g(x) = x^\beta$ for $x \geq 0$, and $L_N = (\log N)^{1/\beta}$.

We remark here that the function g is non-decreasing because of (3.2.2), and the function g^{-1} is non-decreasing as well. To see this, take $0 \leq x_1 < x_2$, and let $y_i := g^{-1}(x_i)$, $i = 1, 2$. Then by the definition of g^{-1} and since g is non-decreasing, for all $\varepsilon > 0$,

$$g(y_2 + \varepsilon) \geq x_2 > x_1.$$

By the definition of g^{-1}, this implies that $y_1 \leq y_2 + \varepsilon$, and since $\varepsilon > 0$ was arbitrary, we conclude that $g^{-1}(x_1) \leq g^{-1}(x_2)$.

Now we state our main result. Let v_N denote the asymptotic speed defined by (3.2.1). For two positive sequences a_N and b_N we say that $a_N \sim b_N$ as $N \to \infty$, if $a_N/b_N \to 1$ as $N \to \infty$.

Theorem 3.2.2. *Consider an N-BRW with arbitrary initial configuration and with a jump distribution given by the random variable X. Assume that the jump distribution satisfies* (3.2.2) *and* (3.2.3) *with a function g which is regularly varying with index $\beta \in (0,1)$. Then*

$$v_N \sim \frac{L_N \log 2}{\log N} \quad \text{as } N \to \infty,$$

where L_N is defined in (3.2.4)*.*

Note that in the case when $g(x) = x^\beta$ for $x \geq 0$, the theorem says that $v_N \sim (\log 2)(\log N)^{1/\beta - 1}$ as $N \to \infty$, as stated in Theorem 1.3.2 in Chapter 1.

3.2.1 Strategy and intuition for the proof of Theorem 3.2.2

In Section 3.3 we will prove a simple monotonicity property of the N-BRW (Lemma 3.3.2 and Lemma 1 in [2]), which has some helpful consequences also used in [2]. In our case these consequences will be that in the proof of Theorem 3.2.2

- it will be enough to consider the initial condition when all particles start from position zero;
- to prove a good lower bound on v_N it will be enough to give a lower bound on $\mathcal{X}_1(\ell_N + 1)$ with high probability;
- to prove a good upper bound on v_N it will be enough to show an upper bound on $\mathcal{X}_N(\lceil A \log N \rceil)$ with high probability for some large constant $A > 0$.

We now discuss the intuition behind the lower and upper bounds on v_N that we will establish to prove Theorem 3.2.2. In order to get an idea of how the speed should behave, let us assume in this section that for $x \geq 0$,

$$g(x) = x^\beta,$$

and that therefore

$$L_N = g^{-1}(\log N) = (\log N)^{1/\beta}.$$

Then we also have

$$\mathbb{P}\left(X > L_N\right) = e^{-g(L_N)} = N^{-1}.$$

Thus, the probability that at time 0 there is at least one jump of size larger than L_N is $1 - (1 - 1/N)^{2N}$, which can be bounded below by a positive constant (close to $1 - e^{-2}$) for large values of N. If there is such a jump at time 0, then by Lemma 3.1.1, at time $\ell_N + 1$ there must be N particles to the right of position L_N, that is, $\mathcal{X}_1(\ell_N + 1) \geq L_N$ with probability bounded away from zero if we assume that all particles have initial position 0.

If we instead fix $\varepsilon \in (0,1)$ and consider jumps larger than $(1-\varepsilon)L_N$, we can show that at time 0, with high probability at least one such jump occurs, and therefore we will have

$\mathcal{X}_1(\ell_N+1) \geq (1-\varepsilon)L_N$ for any $\varepsilon \in (0,1)$ with high probability. This suggests that the speed should be at least

$$v_N \approx \frac{\mathcal{X}_1(\ell_N+1)}{\ell_N+1} \geq \frac{(1-\varepsilon)L_N}{\lceil \log_2 N\rceil + 1} \approx \frac{(1-\varepsilon)L_N \log 2}{\log N} \tag{3.2.6}$$

for large values of N, where we use $\approx$ as an informal notation to indicate that the two sides are close to each other in some sense. The details of this argument will be discussed in Section 3.4.

The more difficult part of our proof is to give an upper bound on v_N. One of the key ideas to do this is to use the strategy of the proof of Theorem 3 in [27], which is a large deviation result of Gantert. Assume we have a random walk $(S_n)_{n\in\mathbb{N}}$ with stretched exponential jump distribution given by (3.2.2), assuming $g(x) = x^\beta$. The main message of Gantert's result is that for fixed $x > 0$ and large $n \in \mathbb{N}$,

$$\mathbb{P}\left(S_n > xn^{1/\beta}\right) \approx ne^{-x^\beta n}.$$

That is, the probability that the random walk goes further than $xn^{1/\beta}$ is roughly the expected number of jumps greater than $xn^{1/\beta}$ by time n; and the proof in [27] also shows that the most likely way for a random walk path to reach $xn^{1/\beta}$ is to have a single big jump of this size on the path.

If we apply this result directly to our problem, we will get an upper bound for v_N, but it will be weaker than what we are aiming for. The following short calculation shows this. We consider a $\lceil \log N\rceil$ time scale similarly to the lower bound. Recall the construction of the N-BRW from N independent branching random walks (BRWs) in Section 2.4.1 in Chapter 2 (see also Figure 1-4). The total number of paths in the N independent BRWs without selection up to time $\lceil \log N\rceil$ is $N2^{\lceil \log N\rceil}$. Fix $x > 0$. If in the N-BRW the rightmost particle at time $\lceil \log N\rceil$ is to the right of position $x\lceil \log N\rceil^{1/\beta} \approx xL_N$, then there must be a path in at least one of the N independent BRWs which moves more than $x\lceil \log N\rceil^{1/\beta}$ in $\lceil \log N\rceil$ time (again assuming that all particles started initially from position 0). Therefore, fixing $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\mathcal{X}_N(\lceil \log N\rceil) > x\lceil \log N\rceil^{1/\beta}\right) &\leq N2^{\lceil \log N\rceil}\mathbb{P}\left(S_{\lceil \log N\rceil} > x\lceil \log N\rceil^{1/\beta}\right)\\ &\leq N2^{\lceil \log N\rceil}e^{-\lceil \log N\rceil(x^\beta-\varepsilon)}\\ &\leq 2N^{1+\log 2 - x^\beta+\varepsilon},\end{aligned} \tag{3.2.7}$$

where the second inequality holds for N sufficiently large by Theorem 3 in [27]. Now in order for the right-hand side to tend to zero as $N \to \infty$, we need $x > (1+\log 2)^{1/\beta}$. Using the upper bound in (3.2.7), it is possible (but not immediate) to show that for $x > (1+\log 2)^{1/\beta}$, we have $v_N \leq xL_N/\log N$ for N sufficiently large. Hence, this argument leads to an upper bound which already shows the order of v_N but the constant $(1+\log 2)^{1/\beta}$ does not match with the constant $\log 2$ in the lower bound in (3.2.6).

The idea to improve this upper bound is the following. Fix $A > 0$ a large constant. We claim that in the N-BRW in $A \log N$ time (let us assume $A \log N \in \mathbb{N}$ for now), jumps of size significantly larger than L_N do not occur with high probability for large N. Indeed, since the number of jumps up to time $A \log N$ is $2NA \log N$, for any $\varepsilon > 0$, the probability that there exists a jump larger than $(1+\varepsilon)L_N$ in $A \log N$ time is at most

$$2NA \log N e^{-(1+\varepsilon)^\beta \log N} \to 0$$

as $N \to \infty$.

The key statement to show in our proof for the upper bound on v_N will be that, for any $\varepsilon > 0$, there exists $A > 0$ such that $\mathcal{X}_N(A \log N) \leq (1+\varepsilon)A(\log 2)L_N$ with high probability for large N. To do this, we again use the construction of the N-BRW from N independent BRWs, and it will be enough to count paths in the BRWs up to time $A \log N$ with jumps of size at most $(1+\varepsilon/100)L_N$, since larger jumps do not occur in the N-BRW with high probability. Then the minimal number of jumps needed to cover the distance $(1+\varepsilon)A(\log 2)L_N$ is roughly $(1+\varepsilon)A(\log 2)$ with jumps of size roughly L_N. The idea of 'one big jump does the job' in [27] suggests that this should be the optimal (most likely) way for a path to cover the above distance, if the jumps are restricted to being at most roughly L_N.

The probability that on a given path of length $A \log N$ in a BRW there are $(1+\varepsilon)A(\log 2)$ jumps of size L_N is less than

$$\binom{A \log N}{(1+\varepsilon)(\log 2)A} N^{-(1+\varepsilon)(\log 2)A}.$$

The number of paths in N independent BRWs up to time $A \log N$ is $N 2^{A \log N} = N^{1+A\log 2}$. So for any $\varepsilon > 0$ we can choose A large enough that the probability that we see a path with $(1+\varepsilon)A(\log 2)$ jumps of size L_N in at least one of N independent BRWs by time $A \log N$ goes to zero as $N \to \infty$.

In our rigorous proof we will give an upper bound for the probability that a random walk path of length $A \log N$ goes further than $(1+\varepsilon)A(\log 2)L_N$ without making any big jumps of size larger than $(1+\varepsilon/100)L_N$. In our argument we will consider an exponential moment of the jump distribution restricted to jumps of size at most $(1+\varepsilon/100)L_N$, and we will follow similar steps to the ones in the proof of Theorem 3 in [27].

We will see that the upper bound on this probability implies that $\mathcal{X}_N(A \log N) \leq (1+\varepsilon)A(\log 2)L_N$ with high probability, which then gives the desired upper bound on v_N.

3.3 Asymptotic speed and proof of Theorem 3.2.2

In this section we discuss the proof of Proposition 3.2.1 following the steps in [2], then state the lower and upper bounds on v_N which lead to the proof of Theorem 3.2.2. The proofs of the lower and upper bounds are in Section 3.4 and 3.5 respectively.

Up to Section 3.3.4 we only assume a general jump distribution given by a non-negative random variable X and with $\mathbb{E}[X] < \infty$.

3.3.1 Diameter

Proposition 3.2.1 states that the leftmost and rightmost particles have the same asymptotic speed as time goes to infinity. Proposition 3.3.1 below is a step towards proving this result: it says that the distance between the rightmost and leftmost particles at time n is $o(n)$ as $n \to \infty$; hence, if (say) the rightmost particle has an asymptotic speed then the leftmost particle has an asymptotic speed as well, and both will have the same value. For $n \in \mathbb{N}_0$, let

$$d(\mathcal{X}(n)) := \mathcal{X}_N(n) - \mathcal{X}_1(n)$$

denote the diameter of the particle cloud at time n.

Proposition 3.3.1. *[2, Corollary 1] For all $N \in \mathbb{N}$,*

$$\lim_{n\to\infty} \frac{d(\mathcal{X}(n))}{n} = 0 \textit{ almost surely and in } L^1.$$

Proof. We follow the proof given in [2], and start with an estimate for the diameter at any fixed time greater than $\log_2 N$, which is stated in Proposition 1 in [2]. Let $n \in \mathbb{N}_0$ be arbitrary. Assume that the rightmost particle is at position $y := \mathcal{X}_N(n)$ at time n. Consider the process at time $n + \ell_N$.

Note that there are $2N\ell_N$ jumps between time n and $n + \ell_N$ in the N-BRW. Let us denote the maximum of these jumps by $M_N(n)$. Then the position of the rightmost particle increases at most by $M_N(n)$ at each step between times n and $n + \ell_N$, giving the upper bound

$$\mathcal{X}_N(n + \ell_N) \leq y + \ell_N M_N(n).$$

Furthermore, since we have a particle at position y at time n, Lemma 3.1.1 implies

$$\mathcal{X}_1(n + \ell_N) \geq y.$$

Therefore, using the above upper and lower bounds, we conclude that for any $n \in \mathbb{N}$, we have

$$d(\mathcal{X}(n + \ell_N)) \leq \ell_N M_N(n). \tag{3.3.1}$$

Let $X_1, \ldots, X_{2N\ell_N}$ be i.i.d. jumps distributed as X, and let $M_N := \max_{j=1,\ldots,2N\ell_N} X_j$. Then we have

$$M_N(n) \stackrel{d}{=} M_N \tag{3.3.2}$$

for all $n \in \mathbb{N}_0$. Since the jumps are non-negative, we can bound M_N, the maximum of

$2N\ell_N$ jumps, by the sum of these jumps. Therefore, for any fixed $N \in \mathbb{N}$,

$$\mathbb{E}\left[M_N(n)\right] = \mathbb{E}\left[M_N\right] \leq \mathbb{E}\left[\sum_{j=1}^{2N\ell_N} X_j\right] = 2N\ell_N\mathbb{E}\left[X\right] =: C_N < \infty. \qquad (3.3.3)$$

Now (3.3.1) immediately implies the L^1 convergence; indeed, for all fixed $N \in \mathbb{N}$ there exists $C_N > 0$ such that

$$0 \leq \liminf_{n\to\infty} \frac{\mathbb{E}\left[d(\mathcal{X}(n))\right]}{n} \leq \limsup_{n\to\infty} \frac{\mathbb{E}\left[d(\mathcal{X}(n))\right]}{n} = \limsup_{n\to\infty} \frac{\mathbb{E}\left[d(\mathcal{X}(n+\ell_N))\right]}{n+\ell_N} \leq \limsup_{n\to\infty} \frac{\ell_N C_N}{n+\ell_N} = 0.$$

To see the almost sure convergence, we observe that for any fixed $N \in \mathbb{N}$ and $\varepsilon > 0$ we have

$$\begin{aligned}\sum_{n=\ell_N}^{\infty} \mathbb{P}\left(\frac{d(\mathcal{X}(n))}{n} > \varepsilon\right) &= \sum_{n\geq 0} \mathbb{P}\left(\frac{d(\mathcal{X}(n+\ell_N))}{n+\ell_N} > \varepsilon\right)\\ &\leq \sum_{n=0}^{\infty} \mathbb{P}\left(\frac{\ell_N M_N(n)}{\varepsilon} > n+\ell_N\right)\\ &\leq \sum_{n=0}^{\infty} \mathbb{P}\left(\frac{\ell_N M_N}{\varepsilon} > n\right)\\ &= \frac{\ell_N\mathbb{E}\left[M_N\right]}{\varepsilon} < \infty,\end{aligned}$$

where in the first inequality we use (3.3.1), the second inequality follows by (3.3.2), and the third by (3.3.3). Therefore, the Borel-Cantelli lemma implies that for all $N \in \mathbb{N}$ and $\varepsilon > 0$, almost surely there exists a random number $n_0 \in \mathbb{N}_0$ such that $d(\mathcal{X}(n))/n \leq \varepsilon$ for all $n > n_0$. Since ε can be taken arbitrarily small, the result follows. □

3.3.2 Monotonicity properties

In [2] it is shown that due to a monotonicity property, to prove Proposition 3.2.1, it is enough to consider the N-BRW with the initial configuration in which all particles are at zero. We now recall this property (in a slightly less general form than is written in [2]), and we include its proof as well for completeness.

The monotonicity property says the following. Assume that we have a pair of N-BRWs, determined by the same sequence of jumps but starting from different initial configurations. Assume furthermore that at some time k, the positions of the first N-BRW are pairwise smaller than those of the second N-BRW. Then we claim that the first N-BRW will have pairwise smaller positions than the second at all times after time k.

Recall the notation $X_{i,b,n}$ from Section 3.1.

Lemma 3.3.2. *[2, Lemma 1] Consider a pair of N-BRWs $(\mathcal{X}(n), \mathcal{X}^*(n))_{n\in\mathbb{N}_0}$, where $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$ is determined by the jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and the initial configuration $\mathcal{X}(0)$, and $(\mathcal{X}^*(n))_{n\in\mathbb{N}_0}$ is determined by the same jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and the* *initial configuration* $\mathcal{X}^*(0)$. *Suppose that there exists a time* $k \in \mathbb{N}_0$ *such that* $\mathcal{X}_i(k) \leq \mathcal{X}_i^*(k)$ *for all* $i \in [N]$. *Then*

$$\mathcal{X}_i(n) \leq \mathcal{X}_i^*(n),$$

for all $n \geq k$ *and* $i \in [N]$.

Proof. Assume that we have

$$\mathcal{X}_i(k) \leq \mathcal{X}_i^*(k) \tag{3.3.4}$$

for all $i \in [N]$ and for some $k \in \mathbb{N}_0$. We will show that

$$\mathcal{X}_i(k+1) \leq \mathcal{X}_i^*(k+1) \tag{3.3.5}$$

for all $i \in [N]$; by induction on k, this implies the result.

For $i \in [N]$ and $n \in \mathbb{N}_0$, let (i, n) denote the ith particle from the left at time n in $\mathcal{X}(n)$, at position $\mathcal{X}_i(n)$. Recall the notation $(i, n) \lesssim_b (j, n+k)$ from Section 3.1. Let also (j, b, k) refer to the jump of the bth offspring of the jth particle at time k (in both $\mathcal{X}$ and $\mathcal{X}^*$). The size of the jump (j, b, k) is $X_{j,b,k}$ (in both $\mathcal{X}$ and $\mathcal{X}^*$).

Take $i' \in [N]$. Let $A_{i'}$ denote the set of jumps at time k for which the particles performing these jumps are in the set $(i', k+1), (i'+1, k+1), \ldots, (N, k+1)$ in $\mathcal{X}(k+1)$:

$$A_{i'} := \bigcup_{i=i'}^{N} \{(j, b, k) : (j, k) \lesssim_b (i, k+1)\}.$$

Then $|A_{i'}| = N - i' + 1$, and we claim that all positions in the collection $(\mathcal{X}_j^*(k) + X_{j,b,k})_{(j,b,k) \in A_{i'}}$ are to the right of or at position $\mathcal{X}_{i'}(k+1)$. The reason for this is that for each $(j, b, k) \in A_{i'}$ there exists $i \geq i'$ such that $(j, k) \lesssim_b (i, k+1)$, and thus, by the assumption in (3.3.4) and since $i \geq i'$, we have

$$\mathcal{X}_j^*(k) + X_{j,b,k} \geq \mathcal{X}_j(k) + X_{j,b,k} = \mathcal{X}_i(k+1) \geq \mathcal{X}_{i'}(k+1).$$

This establishes the claim. It follows that if all the $N - i' + 1$ particles that performed a jump from the set $A_{i'}$ survive the selection step in $\mathcal{X}^*(k+1)$, then there are at least $N - i' + 1$ particles to the right of or at position $\mathcal{X}_{i'}(k+1)$ in $\mathcal{X}^*(k+1)$, and therefore (3.3.5) holds for $i = i'$. If instead there is a particle which performed a jump from the set $A_{i'}$, but it does not survive the selection step in $\mathcal{X}^*(k+1)$, then we must have some j and b with $(j, b, k) \in A_{i'}$ such that

$$\mathcal{X}_{i'}(k+1) \leq \mathcal{X}_j^*(k) + X_{j,b,k} \leq \mathcal{X}_1^*(k+1) \leq \mathcal{X}_{i'}^*(k+1),$$

which again implies that (3.3.5) holds for $i = i'$. Since i' was arbitrary the result follows. □

The next corollary makes sure that in the proof of Proposition 3.2.1 and later in the proof of Theorem 3.2.2 we can assume that the initial configuration is $\mathcal{X}_i(0) = 0$ for all

$i \in [N]$. In [2], the authors argue that we can make this assumption because of Lemma 3.3.2 and by translation invariance (the dynamics does not change if we shift each particle by a translation on $\mathbb{R}$). In Corollary 3.3.3 we expand on this argument.

Corollary 3.3.3. *Consider a pair of N-BRWs $(\mathcal{X}(n), \mathcal{X}^*(n))_{n\in\mathbb{N}_0}$, where $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$ is determined by the initial configuration $\mathcal{X}(0)$ and the jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$, and $(\mathcal{X}^*(n))_{n\in\mathbb{N}_0}$ is determined by the same jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and by the initial configuration $\mathcal{X}_i^*(0) = 0$ for all $i \in [N]$. Assume that*

$$v_N := \lim_{n\to\infty} \frac{\mathcal{X}_N^*(n)}{n}$$

exists almost surely and in L^1. Then

$$\lim_{n\to\infty} \frac{\mathcal{X}_N(n)}{n} = v_N$$

almost surely and in L^1.

Proof. Consider the process $(\mathcal{X}^*(n) + \mathcal{X}_1(0))_{n\in\mathbb{N}_0}$, by which we mean that the particle positions of this process at time $n \in \mathbb{N}_0$ are given by $\mathcal{X}_i^*(n) + \mathcal{X}_1(0)$ for all $i \in [N]$. Then the process $(\mathcal{X}^*(n) + \mathcal{X}_1(0))_{n\in\mathbb{N}_0}$ is an N-BRW given by the same jumps as $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$, started with all particles at $\mathcal{X}_1(0)$. For this process it certainly holds that $\mathcal{X}_i^*(0) + \mathcal{X}_1(0) = \mathcal{X}_1(0) \leq \mathcal{X}_i(0)$ for all $i \in [N]$. Therefore, applying Lemma 3.3.2 with $k = 0$, we have

$$\mathcal{X}_N^*(n) + \mathcal{X}_1(0) \leq \mathcal{X}_N(n) \tag{3.3.6}$$

for all $n \in \mathbb{N}_0$. Similarly, we can also consider the process $(\mathcal{X}^*(n) + \mathcal{X}_N(0))_{n\in\mathbb{N}_0}$. For this process it holds that $\mathcal{X}_i^*(0) + \mathcal{X}_N(0) = \mathcal{X}_N(0) \geq \mathcal{X}_i(0)$ for all $i \in [N]$. Thus, by Lemma 3.3.2 we have

$$\mathcal{X}_N^*(n) + \mathcal{X}_N(0) \geq \mathcal{X}_N(n) \tag{3.3.7}$$

for all $n \in \mathbb{N}_0$. We then conclude by (3.3.6) and (3.3.7) that

$$\frac{\mathcal{X}_N^*(n) + \mathcal{X}_1(0)}{n} \leq \frac{\mathcal{X}_N(n)}{n} \leq \frac{\mathcal{X}_N^*(n) + \mathcal{X}_N(0)}{n},$$

which implies the result. □

3.3.3 Proof of Proposition 3.2.1 – existence of asymptotic speed

In this section, we are following the steps of [2] to complete the proof of Proposition 3.2.1. Let $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$ be an N-BRW given by the jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and by the initial condition

$$\mathcal{X}_i(0) = 0 \text{ for all } i \in [N]. \tag{3.3.8}$$

By Corollary 3.3.3 it is enough to prove the statement of Proposition 3.2.1 for $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$.

The proof of Proposition 3.2.1 relies on Kingman's subadditive ergodic theorem which we now state.

Theorem 3.3.4. *(Kingman's subadditive ergodic theorem – Ligget's version [22, Theorem 7.4.1])*
Suppose that the process $(Y_{m,n})_{0\le m<n}$ satisfies the following conditions:

(i) $Y_{0,m} + Y_{m,n} \ge Y_{0,n}$ *(subadditivity).*

(ii) $\left\{Y_{nk,(n+1)k},\ n \ge 1\right\}$ *is a stationary sequence for all k.*

(iii) The distribution of $\{Y_{m,m+k},\ k \ge 1\}$ *does not depend on m.*

(iv) $\mathbb{E}\left[\max(Y_{0,n}, 0)\right] < \infty$ *for all n, and there exists $\gamma_0 > -\infty$ such that $\mathbb{E}[Y_{0,n}] \ge \gamma_0 n$ for all n.*

Then

(a) $\lim_{n\to\infty} \mathbb{E}\left[\frac{Y_{0,n}}{n}\right] = \inf_m \frac{\mathbb{E}[Y_{0,m}]}{m} \equiv \gamma.$

(b) $Y := \lim_{n\to\infty} \frac{Y_{0,n}}{n}$ *exists almost surely and in L^1, and hence $\mathbb{E}[Y] = \gamma$.*

(c) If $\left\{Y_{nk,(n+1)k},\ n \ge 1\right\}$ *is ergodic for all k, then $Y = \gamma$ almost surely.*

We will apply Kingman's theorem for the so-called shifted processes of the N-BRW. Before introducing these, we will consider a modified N-BRW, which will be related to the shifted processes. Take $m, n \in \mathbb{N}_0$. Let $\mathcal{X}^*$ be an identical copy of $\mathcal{X}$ up to time m, i.e. $\mathcal{X}(l) = \mathcal{X}^*(l)$ for all $l \le m$. At time m we make a change in $\mathcal{X}^*(m)$: we push every particle in $\mathcal{X}^*(m)$ to the rightmost position, so that $\mathcal{X}_i^*(m) = \mathcal{X}_N(m)$ for all $i \in [N]$. Then after time m, we let $\mathcal{X}^*$ evolve as an N-BRW with the same sequence of jumps as $\mathcal{X}$. Now by Lemma 3.3.2, $\mathcal{X}^*$ will dominate $\mathcal{X}$ at all times after time m; in particular,

$$\mathcal{X}_N^*(m+n) \ge \mathcal{X}_N(m+n). \tag{3.3.9}$$

Similarly, we also let $\mathcal{X}^{**}$ be an identical copy of $\mathcal{X}$ up to time m, and the change we make at time m is that we pull back every particle in $\mathcal{X}^{**}(m)$ to the leftmost position, so that $\mathcal{X}_i^{**}(m) = \mathcal{X}_1(m)$ for all $i \in [N]$. Then after time m, we let $\mathcal{X}^{**}$ evolve as an N-BRW with the same sequence of jumps as $\mathcal{X}$. Then $\mathcal{X}$ will dominate $\mathcal{X}^{**}$ at all times after time m; in particular

$$\mathcal{X}_1^{**}(m+n) \le \mathcal{X}_1(m+n). \tag{3.3.10}$$

In order to use the above properties conveniently, we introduce the process $(\mathcal{X}^m(m+k))_{k\in\mathbb{N}_0}$, which will be the shifted process of the N-BRW $(\mathcal{X}(k))_{k\in\mathbb{N}_0}$ shifted by m. We define this shifted process as follows: $(\mathcal{X}^m(m+k))_{k\in\mathbb{N}_0}$ is an N-BRW which starts at time

m, and it is given by the initial configuration $\mathcal{X}_i^m(m) = \mathcal{X}_i(0) = 0$ for all $i \in [N]$, and by the jumps $(X_{i,b,k})_{i\in[N],b\in\{1,2\},k\geq m}$. Since the jumps are i.i.d. and $\mathcal{X}^m(m) = \mathcal{X}(0)$, we have

$$\mathcal{X}_i^m(m+n) \stackrel{d}{=} \mathcal{X}_i(n), \tag{3.3.11}$$

for all $i \in [N]$.

Furthermore, because of the construction of $\mathcal{X}^*$, the rightmost position in $\mathcal{X}^*$ can be written as

$$\mathcal{X}_N^*(m+n) = \mathcal{X}_N(m) + \mathcal{X}_N^m(m+n). \tag{3.3.12}$$

Similarly, we also have

$$\mathcal{X}_1^{**}(m+n) = \mathcal{X}_1(m) + \mathcal{X}_1^m(m+n). \tag{3.3.13}$$

A useful consequence of the above construction and the inequalities (3.3.9) and (3.3.10), which we will use in our main proofs later on, is stated in the following lemma.

Lemma 3.3.5. *For all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$, we have*

$$\mathbb{E}\left[\mathcal{X}_N(m+n)\right] \leq \mathbb{E}\left[\mathcal{X}_N(m)\right] + \mathbb{E}\left[\mathcal{X}_N(n)\right]$$

and

$$\mathbb{E}\left[\mathcal{X}_1(m+n)\right] \geq \mathbb{E}\left[\mathcal{X}_1(m)\right] + \mathbb{E}\left[\mathcal{X}_1(n)\right].$$

Proof. The first inequality of the lemma follows from (3.3.9), (3.3.11) and (3.3.12); and we see the second inequality by (3.3.10), (3.3.11) and (3.3.13). □

Now we can prove Proposition 3.2.1 by applying Kingman's subadditive ergodic theorem to the shifted process we introduced earlier.

Proof of Proposition 3.2.1. Proposition 3.3.1 implies that if either of the two limits in Proposition 3.2.1 exists, the other exists as well and has the same value. We will prove the result by considering the evolution of the rightmost position $\mathcal{X}_N(n)$. We will apply Kingman's subadditive ergodic theorem to the shifted process $(\mathcal{X}_N^m(n))_{m,n\in\mathbb{N}_0,n\geq m}$.

Now we check the conditions of Theorem 3.3.4. Let $Y_{m,n} := \mathcal{X}_N^m(n)$ for all $0 \leq m \leq n$, and then fix m and n. Condition (i) requires

$$\mathcal{X}_N^0(m) + \mathcal{X}_N^m(n) \geq \mathcal{X}_N^0(n).$$

Notice that $\mathcal{X}_N^0(k) = \mathcal{X}_N(k)$ for all $k \in \mathbb{N}_0$ by the definition of the shifted process. Therefore, (3.3.9) and (3.3.12) imply condition (i).

Regarding condition (ii), notice that by (3.3.11) we have $\mathcal{X}_N^{nk}((n+1)k) \stackrel{d}{=} \mathcal{X}_N(k)$ for all $n \in \mathbb{N}_0$. Thus, $(\mathcal{X}_N^{nk}((n+1)k))_{n\in\mathbb{N}_0}$ is an i.i.d. sequence and therefore it is stationary and ergodic, which shows both condition (ii) and the condition in statement (c) of Theorem 3.3.4.

For condition (iii) we have $\mathcal{X}_N^m(m+k) \overset{d}{=} \mathcal{X}_N(k)$, and therefore the distribution of $(\mathcal{X}_N^m(m+k))_{k\geq 1}$ indeed does not depend on m. For the first condition in (iv), note that because of (3.3.8) and since the jumps are non-negative, we have $\mathcal{X}_N^0(n) \geq 0$ for all $n \in \mathbb{N}_0$, and we can bound $\mathcal{X}_N^0(n)$ above by the sum of all jumps that occurred in $\mathcal{X}$ before time n. Hence, for all $n \in \mathbb{N}_0$ we have

$$\mathbb{E}\left[\mathcal{X}_N^0(n)\right] \leq \mathbb{E}\left[\sum_{i\in[N],b\in\{1,2\},k\in[0,n-1]} X_{i,b,k}\right] = 2Nn\mathbb{E}\left[X\right] < \infty.$$

The second condition in (iv) also holds, since $\mathbb{E}\left[\mathcal{X}_N^0(n)\right] \geq 0$ for all $n \in \mathbb{N}_0$.

Therefore Theorem 3.3.4 applies, and since $\mathcal{X}_N^0(n) = \mathcal{X}_N(n)$, it says that

$$\lim_{n\to\infty} \frac{\mathcal{X}_N(n)}{n} = \lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{X}_N(n)\right]}{n} = \inf_k \frac{\mathbb{E}\left[\mathcal{X}_N(k)\right]}{k}$$

almost surely and in L^1, which concludes the proof of Proposition 3.2.1. □

3.3.4 Proof of Theorem 3.2.2

We now move on to proving the main result of this chapter, Theorem 3.2.2. Consider an N-BRW $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$ with arbitrary initial configuration $\mathcal{X}(0)$ and with a jump distribution given by the random variable X in (3.2.2). Recall that X is non-negative and that g in (3.2.2) is regularly varying with index $\beta \in (0,1)$.

Since X has finite moments (for example by applying Lemma 3.1.3 to $f = g$), Proposition 3.2.1 applies, and the L^1 convergence implies

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{X}_1(n)\right]}{n} = \lim_{n\to\infty} \frac{\mathbb{E}\left[\mathcal{X}_N(n)\right]}{n} = v_N. \tag{3.3.14}$$

Our main argument will consist of proving the lower and upper bounds stated below.

Proposition 3.3.6. *Assume $\mathcal{X}_i(0) = 0$ for all $i \in [N]$. Then for all $c_1 < \log 2$, for N sufficiently large,*

$$c_1 \frac{L_N}{\log N} \leq \lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{X}_1(n)]}{n} = v_N.$$

Proposition 3.3.7. *Assume $\mathcal{X}_i(0) = 0$ for all $i \in [N]$. Then for all $c_2 > \log 2$, for N sufficiently large,*

$$v_N = \lim_{n\to\infty} \frac{\mathbb{E}[\mathcal{X}_N(n)]}{n} \leq c_2 \frac{L_N}{\log N}.$$

Proof of Theorem 3.2.2. Corollary 3.3.3, Propositions 3.3.6 and 3.3.7, together with (3.3.14) immediately imply the theorem. □

We will prove Proposition 3.3.6 in Section 3.4 and the proof of Proposition 3.3.7 is in Section 3.5.

3.3.5 Simple properties of the regularly varying function g

Recall the definitions of g and g^{-1} from (3.2.2) and (3.2.5). In the proofs of Propositions 3.3.6 and 3.3.7, the following properties will be helpful.

Lemma 3.3.8. *For all $r > 0$,*

(a) $g(g^{-1}(r\log N)) \sim r\log N$ as $N\to\infty$; in particular $g(L_N)\sim \log N$ as $N\to\infty$

(b) $g(rL_N)\sim r^\beta \log N$ as $N\to\infty$

(c) $g^{-1}(r\log N)\sim r^{1/\beta}L_N$ as $N\to\infty$.

Proof. Take $\varepsilon\in(0,1)$.

(a) By the definition of a regularly varying function in (3.1.2), the definition of g^{-1} in (3.2.5), and since g is non-decreasing, for N sufficiently large,

$$(1-\varepsilon)^\beta-\varepsilon \le \frac{g((1-\varepsilon)g^{-1}(r\log N))}{g(g^{-1}(r\log N))} \le \frac{r\log N}{g(g^{-1}(r\log N))} \le \frac{g((1+\varepsilon)g^{-1}(r\log N))}{g(g^{-1}(r\log N))} \le (1+\varepsilon)^\beta+\varepsilon,$$

which shows (a), since $\varepsilon\in(0,1)$ was arbitrary.

(b) Recall the definition of L_N in (3.2.4). By part (a) (with $r=1$) and by (3.1.2), for N sufficiently large we have

$$(1-\varepsilon)^2r^\beta\log N\le(1-\varepsilon)r^\beta g(L_N)\le g(rL_N)\le(1+\varepsilon)r^\beta g(L_N)\le(1+\varepsilon)^2r^\beta\log N,$$

which shows (b), since $\varepsilon\in(0,1)$ was arbitrary.

(c) For N sufficiently large,

$$\begin{aligned} g\left((1-\varepsilon)^{2/\beta}r^{1/\beta}L_N\right) &\le \frac{1}{1-\varepsilon}(1-\varepsilon)^2 r\log N < g(g^{-1}(r\log N)) \\ &< \frac{1}{1+\varepsilon}(1+\varepsilon)^2 r\log N \\ &\le g\left((1+\varepsilon)^{2/\beta}r^{1/\beta}L_N\right), \end{aligned}$$

where in the first inequality we use part (b) with $(1-\varepsilon)^{2/\beta}r^{1/\beta}$ in the role of r, the second follows by part (a), and the third and fourth again by (a) and (b) respectively. Since the function g is non-decreasing because of (3.2.2), the above implies

$$(1-\varepsilon)^{2/\beta}r^{1/\beta}L_N<g^{-1}(r\log N)<(1+\varepsilon)^{2/\beta}r^{1/\beta}L_N,$$

and since we can take ε arbitrarily small, we conclude (c).

□

We will also need a lower bound on L_N, which we state in the corollary below. This lower bound follows from Lemma 3.1.3 and Lemma 3.3.8.

Corollary 3.3.9. *For all $\varepsilon > 0$, for N sufficiently large,*

$$L_N \geq (\log N)^{1/\beta - \varepsilon}.$$

Proof. Without loss of generality, assume $\varepsilon \in (0,1)$. Let $\epsilon_1 \in (0, \varepsilon\beta^2)$, and $C_2(\epsilon_1)$ as in Lemma 3.1.3 with $f = g$. Since $\varepsilon < 1/\beta$, we can take N sufficiently large that $(\log N)^{1/\beta-\varepsilon} > B(\epsilon_1)$ from Lemma 3.1.3. Now first by Lemma 3.1.3, then by the choice of ϵ_1, and finally by Lemma 3.3.8(a) we have

$$g\left((\log N)^{1/\beta-\varepsilon}\right) \leq C_2(\epsilon_1)(\log N)^{(1/\beta-\varepsilon)(\beta+\epsilon_1)} \leq (1-\varepsilon)\log N < g(L_N),$$

for N sufficiently large. Since the function g is non-decreasing, the result follows. □

3.4 Lower bound: proof of Proposition 3.3.6

Proof of Proposition 3.3.6. Take $a \in (0,1)$ and assume that $c_1 < a \log 2$. Also, choose $\eta > 0$ such that $a^\beta(1+\eta) < 1$. Recall the definition of L_N from (3.2.4). First we claim that for large N, the leftmost particle is to the right of position aL_N at time $\ell_N + 1$ with high probability. Indeed, for the probability that a single jump is greater than aL_N we have

$$\mathbb{P}(X > aL_N) = e^{-g(aL_N)} \geq e^{-(1+\eta)a^\beta \log N} = N^{-(1+\eta)a^\beta},$$

where the inequality follows for N sufficiently large by Lemma 3.3.8(b).

The probability that at time one there is at least one particle to the right of aL_N (i.e. $\mathcal{X}_N(1) > aL_N$) is equal to the probability that we see a jump of size greater than aL_N among the $2N$ jumps happening at time 0 (since we assume that all particles start from 0 at time 0). Since the jumps are i.i.d., we have for N sufficiently large,

$$\mathbb{P}(\mathcal{X}_N(1) > aL_N) = 1 - (1 - \mathbb{P}(X > aL_N))^{2N} \geq 1 - (1 - N^{-(1+\eta)a^\beta})^{2N}.$$

By Lemma 3.1.1 it follows that for N sufficiently large,

$$\mathbb{P}(\mathcal{X}_1(\ell_N + 1) > aL_N) \geq \mathbb{P}(\mathcal{X}_N(1) > aL_N) \geq 1 - (1 - N^{-(1+\eta)a^\beta})^{2N}, \tag{3.4.1}$$

which tends to 1 as N goes to infinity, because $(1+\eta)a^\beta \in (0,1)$ by our assumption at the beginning of the proof. This establishes the claim above.

In order to establish a lower bound on $\mathbb{E}[\mathcal{X}_1(n)]$ for any large n, we will split the time interval $[0,n]$ into subintervals of length $\ell_N + 1$, and use Lemma 3.3.5 to gain a lower bound on $\mathbb{E}[\mathcal{X}_1(n)]$ in terms of $\mathbb{E}[\mathcal{X}_1(\ell_N+1)]$. Let us fix $N \in \mathbb{N}$ and $n \in \mathbb{N}_0$, and introduce the notation

$$M := \left\lfloor \frac{n}{\ell_N + 1} \right\rfloor .$$

Then by the definition of M and because $\mathcal{X}_1(s)$ is a non-decreasing function of s, and then writing $\mathcal{X}_1(M(\ell_N+1))$ as a telescoping sum, we have

$$\begin{aligned}\frac{1}{n}\mathbb{E}[\mathcal{X}_1(n)] \geq \frac{1}{n}\mathbb{E}[\mathcal{X}_1(M(\ell_N+1))] &= \frac{1}{n}\sum_{k=0}^{M-1}\mathbb{E}[\mathcal{X}_1((k+1)(\ell_N+1)) - \mathcal{X}_1(k(\ell_N+1))] \\ &\geq \frac{1}{n}M\mathbb{E}[\mathcal{X}_1(\ell_N+1)] \\ &\geq \frac{1}{n}\left(\frac{n}{\ell_N+1} - 1\right)\mathbb{E}[\mathcal{X}_1(\ell_N+1)], \end{aligned} \tag{3.4.2}$$

where in the second inequality we use Lemma 3.3.5, and the third follows since $M+1 \geq n/(\ell_N+1)$. Now letting n go to infinity and applying Proposition 3.2.1, we conclude that for any fixed N, we have

$$v_N = \lim_{n\to\infty}\frac{\mathbb{E}[\mathcal{X}_1(n)]}{n} \geq \frac{\mathbb{E}\left[\mathcal{X}_1(\ell_N+1)\right]}{\ell_N+1}.$$

This implies that for N sufficiently large,

$$\begin{aligned} v_N \geq \frac{\mathbb{E}[\mathcal{X}_1(\ell_N+1)]}{aL_N}\frac{aL_N}{\ell_N+1} &\geq \mathbb{P}(\mathcal{X}_1(\ell_N+1) > aL_N)\frac{aL_N}{\ell_N+1} \\ &\geq (1-(1-N^{-(1+\eta)a^\beta})^{2N})\frac{aL_N}{\frac{\log N}{\log 2}+2}, \end{aligned}$$

where the second inequality comes from Markov's inequality, and the third holds for N sufficiently large by (3.4.1) and because $\ell_N \leq \log N/\log 2 + 1$. Now the right-hand side converges to $a(\log 2)L_N/\log N$ as $N \to \infty$, and note that we assumed $c_1 < a\log 2$ at the start of the proof. Therefore, for N sufficiently large, the right-hand side is greater than $c_1 L_N/\log N$, which shows the result. □

3.5 Upper bound: proof of Proposition 3.3.7

In this section, we first prove Proposition 3.3.7 using the lemma below, then we prove the lemma. For all $A > 0$ and $N \in \mathbb{N}$ we introduce the notation

$$t_N = t_N(A) := \lceil A\log N \rceil . \tag{3.5.1}$$

Recall the definition of L_N from (3.2.4). The first lemma below says that for a suitable constant A, the probability that there is a particle to the right of position $(\log 2)(1+\varepsilon)AL_N$ at time t_N is very small for large N. We will see in Section 3.5.1 that this lemma implies Proposition 3.3.7.

Lemma 3.5.1. *Assume $\mathcal{X}_i(0) = 0$ for all $i \in [N]$. Then for all $\varepsilon > 0$, there exist $\delta > 0$ and $A > 0$ such that for N sufficiently large,*

$$\mathbb{P}(\mathcal{X}_N(t_N(A)) > (\log 2)(1+\varepsilon)AL_N) \le N^{-\delta}.$$

Before showing how Lemma 3.5.1 can be used to prove Proposition 3.3.7, we now state a second lemma. Our second lemma is about a path in which jumps of size significantly larger than L_N are discarded and count as a jump of size zero. The lemma says that the probability that such a path of length t_N goes further than $(\log 2)(1+\varepsilon)AL_N$ is very small for large N. The result is strong enough to say that if we consider all paths of length t_N in N independent branching random walks without selection, then it is still very unlikely that one of them makes it further than $(\log 2)(1+\varepsilon)AL_N$ if jumps much larger than L_N are discarded. We can also check that jumps of size much greater than L_N are unlikely to occur during a time interval of length t_N. Hence, using the lemma below, we will be able to prove Lemma 3.5.1 in Section 3.5.2.

Lemma 3.5.2. *For all $\varepsilon > 0$ and $\delta_1 > 0$, there exist $A > 0$ and $\tilde{\varepsilon} > 0$ such that for N sufficiently large,*

$$\mathbb{P}\left(\tilde{S}_{t_N} > (\log 2)(1+\varepsilon)AL_N\right) \le N^{-A\log 2 - 1 - \delta_1},$$

where $\tilde{S}_n = \sum_{j=1}^n Y_j \mathbb{1}_{\{Y_j \le (1+\tilde{\varepsilon})L_N\}}$ for $n \ge 1$, and Y_j, $j = 1, \ldots, n$ are non-negative and i.i.d. and distributed as (3.2.2)*, where g is regularly varying with index $\beta \in (0,1)$ and satisfies* (3.2.3)*.*

We will prove Lemma 3.5.2 in Section 3.5.3.

3.5.1 Proof of the upper bound

Proof of Proposition 3.3.7. Similarly to the proof of the lower bound in Proposition 3.3.6, we can show that it is enough to consider the first t_N steps, using Lemma 3.3.5. Indeed, let us fix $N \in \mathbb{N}$, $n \in \mathbb{N}_0$, $A > 0$, set $t_N = t_N(A)$ and introduce $M := \lceil n/t_N \rceil$. Now similarly to (3.4.2), we use the definition of M and that $\mathcal{X}_N(s)$ is non-decreasing in s, and then use a telescoping sum to write

$$\begin{aligned}
\frac{\mathbb{E}[\mathcal{X}_N(n)]}{n} \le \frac{\mathbb{E}[\mathcal{X}_N(Mt_N)]}{n} &= \frac{1}{n}\sum_{k=0}^{M-1} \mathbb{E}[\mathcal{X}_N((k+1)t_N) - \mathcal{X}_N(kt_N)] \\
&\le \frac{1}{n} M\mathbb{E}[\mathcal{X}_N(t_N)] \\
&\le \frac{1}{n}\left(\frac{n}{t_N} + 1\right)\mathbb{E}[\mathcal{X}_N(t_N)] \qquad (3.5.2)
\end{aligned}$$

where in the second inequality we use Lemma 3.3.5, and the third follows since $M \le n/t_N + 1$. Now letting n go to infinity and applying Proposition 3.2.1, we conclude that for any fixed N and A,

$$v_N \le \frac{\mathbb{E}\left[\mathcal{X}_N(t_N(A))\right]}{t_N(A)}. \tag{3.5.3}$$

We will now use Lemma 3.5.1 to establish an upper bound on the right-hand side. Take $c_2 > \log 2$ and let $\varepsilon > 0$ be such that $c_2 > (1+\varepsilon)\log 2$. For this ε, choose $\delta > 0$ and $A > 0$ as in Lemma 3.5.1, and take $\gamma \in (0, \delta)$. Let us also define

$$K_N := (\log 2)(1+\varepsilon) A L_N.$$

Then partitioning the expectation on the right-hand side of (3.5.3), we get

$$\begin{aligned}\frac{\mathbb{E}[\mathcal{X}_N(t_N)]}{t_N} \le \frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)\le K_N\}}\right] &+ \frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)\in(K_N,N^\gamma]\}}\right] \\ &+ \frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)>N^\gamma\}}\right].\end{aligned} \tag{3.5.4}$$

Recall from (3.5.1) that $t_N = \lceil A\log N\rceil$. For the first term on the right-hand side of (3.5.4) we have

$$\frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)\le K_N\}}\right] \le \frac{K_N}{t_N} \le (\log 2)(1+\varepsilon)L_N(\log N)^{-1}. \tag{3.5.5}$$

For the second term,

$$\frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)\in(K_N,N^\gamma]\}}\right] \le \frac{1}{t_N}\mathbb{P}\left(\mathcal{X}_N(t_N) > K_N\right)N^\gamma \le \frac{N^{\gamma-\delta}}{t_N} < \varepsilon, \tag{3.5.6}$$

for N sufficiently large, where the second inequality follows by Lemma 3.5.1, and the third since $\gamma < \delta$.

For the third term on the right-hand side of (3.5.4), we will use the identity that for any non-negative random variable Y with $\mathbb{E}[Y] < \infty$, and for any fixed $x > 0$,

$$\mathbb{E}\left[Y\mathbb{1}_{\{Y>x\}}\right] = \int_x^\infty \mathbb{P}(Y>y)\,dy + x\mathbb{P}(Y>x).$$

To prove the identity, we write the first probability on the right-hand side as the expectation of an indicator, and apply Fubini's theorem to obtain

$$\begin{aligned}\int_x^\infty \mathbb{P}(Y>y)\,dy = \mathbb{E}\left[\int_x^\infty \mathbb{1}_{\{Y>y\}}dy\right] = \mathbb{E}\left[\int_x^Y dy\mathbb{1}_{\{Y>x\}}\right] &= \mathbb{E}\left[(Y-x)\mathbb{1}_{\{Y>x\}}\right] \\ &= \mathbb{E}\left[Y\mathbb{1}_{\{Y>x\}}\right] - x\mathbb{P}(Y>x),\end{aligned}$$

which shows the identity. Using the identity for the third term on the right-hand side of (3.5.4) we get

$$\frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)>N^\gamma\}}\right] \le \frac{1}{t_N}\int_{N^\gamma}^\infty \mathbb{P}\left(\mathcal{X}_N(t_N)>y\right)dy + \frac{1}{t_N}N^\gamma\mathbb{P}\left(\mathcal{X}_N(t_N)>N^\gamma\right). \tag{3.5.7}$$

Now notice that the event $\{\mathcal{X}_N(t_N)>y\}$ implies that at least one jump of size larger than y/t_N occurred in the N-BRW process before time t_N. Thus, since the total number of jumps before time t_N is $2Nt_N$, we have

$$\mathbb{P}\left(\mathcal{X}_N(t_N)>y\right) \le 2Nt_N e^{-g(y/t_N)} \le 2Nt_N e^{-(y/t_N)^{\beta/2}},$$

for y sufficiently large, where the second inequality comes from Lemma 3.1.3 with $\varepsilon<1/2$. Substituting into (3.5.7) we get

$$\frac{1}{t_N}\mathbb{E}\left[\mathcal{X}_N(t_N)\mathbb{1}_{\{\mathcal{X}_N(t_N)>N^\gamma\}}\right] \le \frac{1}{t_N}2Nt_N\left(\int_{N^\gamma}^\infty e^{-(y/t_N)^{\beta/2}}dy + N^\gamma e^{-(N^\gamma/t_N)^{\beta/2}}\right) < \varepsilon, \tag{3.5.8}$$

for N sufficiently large. Now (3.5.3)-(3.5.8) show that

$$v_N \le \frac{\mathbb{E}[\mathcal{X}_N(t_N)]}{t_N} \le (\log 2)(1+\varepsilon)L_N(\log N)^{-1} + 2\varepsilon,$$

for N sufficiently large, and by Corollary 3.3.9 and since we chose ε such that $c_2>(1+\varepsilon)\log 2$, we are done with the proof of Proposition 3.3.7. □

3.5.2 Lemma 3.5.2 implies Lemma 3.5.1 – bounding the rightmost position

We now prove Lemma 3.5.1 using Lemma 3.5.2, and then prove Lemma 3.5.2 in the next section.

Proof of Lemma 3.5.1. Take $\varepsilon>0$ and $\delta_1>0$. Let A and $\tilde{\varepsilon}$ be as in Lemma 3.5.2. Let $\mathbf{B}_N$ denote the set of big jumps before time $t_N=t_N(A)$ which are larger than $(1+\tilde{\varepsilon})L_N$:

$$\mathbf{B}_N = \mathbf{B}_N(A,\tilde{\varepsilon}) := \{(i,b,k)\in[N]\times\{1,2\}\times[\![0,t_N(A)-1]\!] :\ X_{i,b,k}>(1+\tilde{\varepsilon})L_N\}.$$

Now we split the event in Lemma 3.5.1 based on whether a big jump occurred before time t_N:

$$\begin{aligned}\mathbb{P}(\mathcal{X}_N(t_N)>(\log 2)(1+\varepsilon)AL_N) &= \mathbb{P}(\{\mathcal{X}_N(t_N)>(\log 2)(1+\varepsilon)AL_N\}\cap\{\mathbf{B}_N\neq\emptyset\})\\ &\quad+\mathbb{P}(\{\mathcal{X}_N(t_N)>(\log 2)(1+\varepsilon)AL_N\}\cap\{\mathbf{B}_N=\emptyset\}).\end{aligned} \tag{3.5.9}$$

Recall that we assume $\mathcal{X}_i(0)=0$ for all $i\in[N]$. The event in the second term on the right-hand side implies that there is a path between time 0 and t_N such that no big jumps happen on the path, and the sum of the jumps on the path is more than $(\log 2)(1+\varepsilon)AL_N$. (Recall the notation for paths from (2.2.10)) Hence, (3.5.9) implies

$$\begin{aligned}&\mathbb{P}(\mathcal{X}_N(t_N) > (\log 2)(1+\varepsilon)AL_N)\\ &\leq \mathbb{P}(\mathbf{B}_N \neq \emptyset) + \mathbb{P}\left(\exists i_1, i_2 \in [N] : \sum_{(i,b,k)\in P_{i_1,0}^{i_2,t_N}} X_{i,b,k}\mathbb{1}_{\{X_{i,b,k}\leq(1+\tilde{\varepsilon})L_N\}} > (\log 2)(1+\varepsilon)AL_N\right).\end{aligned} \tag{3.5.10}$$

Indeed, a particular case of the event in the second term in the right-hand side of (3.5.10) is when all the indicators in the sum are 1; that is, there is a path which goes further than $(\log 2)(1+\varepsilon)AL_N$, and every jump on the path is smaller than $(1+\tilde{\varepsilon})L_N$. Thus, this event indeed contains the second event on the right-hand side of (3.5.9).

Recall the construction of the N-BRW from N independent BRWs from Section 2.4.1. Also recall Lemma 2.4.1, which allows us to bound the second probability on the right-hand side of (3.5.10) by the probability that there exists a path in at least one of the N independent BRWs which moves more than $(\log 2)(1+\varepsilon)AL_N$ without big jumps. We will use the notation of Section 2.4.1 to formalise this. Furthermore, for the first probability on the right-hand side of (3.5.10) we use a union bound and that in the N-BRW there are altogether $2Nt_N$ jumps before time t_N. Therefore, we continue (3.5.10) as follows:

$$\begin{aligned}&\mathbb{P}(\mathcal{X}_N(t_N) > (\log 2)(1+\varepsilon)AL_N)\\ &\leq 2Nt_N e^{-g((1+\tilde{\varepsilon})L_N)}\\ &\quad + \mathbb{P}\left(\exists j \in [N], v \in \{1,2\}^{t_N} : \sum_{w\preceq v} Y_{j,w}\mathbb{1}_{\{Y_{j,w}\leq(1+\tilde{\varepsilon})L_N\}} > (\log 2)(1+\varepsilon)AL_N\right),\end{aligned}$$

where on the right-hand side, j is the index of one of the N independent BRWs, (j,v) is the label of a time-t_N particle, and $(Y_{j,w},\ w \preceq v)$ are the i.i.d. jumps of all the ancestors of particle (j,v).

The total number of paths in N independent BRWs from time 0 up to t_N is $N2^{t_N}$. Indeed, there are N possible values of j and 2^{t_N} possible values of v in the event above. Furthermore, the sum of i.i.d. jumps in that event is distributed as $\tilde{S}_{t_N}$ from Lemma 3.5.2. To bound the first term on the right-hand side above, we pick $\eta_1 > 0$ such that $(1-\eta_1)(1+\tilde{\varepsilon})^\beta > 1$, and we apply Lemma 3.3.8(b). Then we obtain that for N sufficiently large,

$$\begin{aligned}\mathbb{P}(\mathcal{X}_N(t_N) > (\log 2)(1+\varepsilon)AL_N) &\leq 2Nt_N e^{-(1-\eta_1)(1+\tilde{\varepsilon})^\beta \log N}\\ &\quad + N2^{t_N}\mathbb{P}\left(\tilde{S}_{t_N} > (\log 2)(1+\varepsilon)AL_N\right)\\ &\leq 2t_N N^{1-(1-\eta_1)(1+\tilde{\varepsilon})^\beta} + 2N^{1+A\log 2 - A\log 2 - 1 - \delta_1}\\ &= 2t_N N^{1-(1-\eta_1)(1+\tilde{\varepsilon})^\beta} + 2N^{-\delta_1},\end{aligned}$$

where in the second inequality we used Lemma 3.5.2 and that $t_N \le A\log N + 1$. Now any $0 < \delta < \min((1-\eta_1)(1+\tilde{\varepsilon})^\beta - 1, \delta_1)$ satisfies the statement of Lemma 3.5.1 for N sufficiently large, thus we are done with the proof of this result. □

3.5.3 Proof of Lemma 3.5.2 – large deviation for paths without big jumps

Proof of Lemma 3.5.2. Without loss of generality we can assume that we have $\varepsilon \in (0,1)$ and it is small enough that for all $\varepsilon' \in (0,\varepsilon)$,

$$(1+\varepsilon')^{-1} > 1 - \tfrac{3}{2}\varepsilon', \tag{3.5.11}$$

and that

$$1 - 1/\beta + \varepsilon < 0. \tag{3.5.12}$$

Let $\delta_1 > 0$. Let $\tilde{\varepsilon} = \varepsilon/100$, and define $\tilde{Y}_j := Y_j \mathbb{1}_{\{Y_j \le (1+\tilde{\varepsilon})L_N\}}$ for $j \ge 1$, and $\tilde{Y} := Y\mathbb{1}_{\{Y \le (1+\tilde{\varepsilon})L_N\}}$, where Y has the same distribution as Y_1. We will need some constants in the course of the proof, which we choose as follows:

(a) Let $\delta_2 := 2\tilde{\varepsilon}$.

(b) Set $A > 0$ such that $A\log 2(\frac{\varepsilon}{2} - \delta_2 - \frac{\varepsilon}{2}\delta_2) > 1 + \delta_1$.

(c) Set $c > 0$ such that $cA^{1-1/\beta} = 1 - \delta_2$.

(d) Set $\varepsilon_1 > 0$ such that $(1+\varepsilon_1)(1-\delta_2) < (1+\tilde{\varepsilon})^{-1}$.

(e) Set $\varepsilon_2 \in (0,1)$ such that $(1+\varepsilon_2)(1+\varepsilon_1)(1-\delta_2) < (1-\varepsilon_2)(1+\tilde{\varepsilon})^{-1}$.

The choice in (b) is possible because $\varepsilon/2 = 50\tilde{\varepsilon} = 25\delta_2 > (1+\varepsilon/2)\delta_2$, where in the first inequality we use (a) and in the second the assumption that $\varepsilon < 1$. The choice in (d) is possible because, by the assumption in (3.5.11), and since $\tilde{\varepsilon} < \varepsilon$, we have

$$(1+\tilde{\varepsilon})^{-1} > 1 - \tfrac{3}{2}\tilde{\varepsilon} > 1 - \delta_2,$$

where the second inequality follows by (a). The choice in (e) is possible because of (d). Finally, we set

$$x = (1+\varepsilon/2)A^{1-1/\beta}\log 2. \tag{3.5.13}$$

Now we start to bound the probability in Lemma 3.5.2. First we claim that for N sufficiently large,

$$xg^{-1}(t_N) \le (\log 2)(1+\varepsilon)AL_N. \tag{3.5.14}$$

To see this, let $\eta_2 > 0$ be sufficiently small that

$$(1+\tfrac{\varepsilon}{2})(A+2\eta_2)^{1/\beta} \le (1+\varepsilon)A^{1/\beta}.$$

Then, by the definition of x, and since $t_N \le (A+\eta_2)\log N$ for N large enough and g^{-1} is non-decreasing (see the paragraph after (3.2.5)), we have

$$\begin{aligned} xg^{-1}(t_N) &\le (\log 2)A^{1-1/\beta}(1+\tfrac{\varepsilon}{2})g^{-1}((A+\eta_2)\log N) \\ &\le (\log 2)A^{1-1/\beta}(1+\tfrac{\varepsilon}{2})(A+2\eta_2)^{1/\beta}L_N \\ &\le (\log 2)A(1+\varepsilon)L_N, \end{aligned}$$

for N sufficiently large, where the second inequality follows by Lemma 3.3.8(c) and the third by the choice of η_2. This finishes the proof of our claim in (3.5.14), which implies that for N sufficiently large,

$$\mathbb{P}\left(\tilde{S}_{t_N} > (\log 2)(1+\varepsilon)AL_N\right) \le \mathbb{P}\left(\tilde{S}_{t_N} > xg^{-1}(t_N)\right). \tag{3.5.15}$$

We will now use an argument similar to Gantert's proof of Theorem 3 in [27] to establish an upper bound on the right-hand side of (3.5.15). Recall that we chose $c>0$ in (c). We start by applying Markov's inequality and using independence to obtain

$$\begin{aligned} \mathbb{P}\left(\tilde{S}_{t_N} > xg^{-1}(t_N)\right) &= \mathbb{P}\left(\exp\left(\frac{ct_N}{g^{-1}(t_N)}\sum_{j=1}^{t_N}\tilde{Y}_j\right) > \exp\left(\frac{ct_N}{g^{-1}(t_N)}xg^{-1}(t_N)\right)\right) \\ &\le \mathbb{E}\left[\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right]^{t_N} e^{-ct_N x}. \end{aligned} \tag{3.5.16}$$

From now on we will be focussing on the exponential moment of $\tilde{Y}$ on the right-hand side. The following two facts will be helpful in our calculations. First, for any $y>0$ we have

$$\log y \le y-1; \tag{3.5.17}$$

and second, for any $z \ge 0$

$$e^z - 1 \le ze^z. \tag{3.5.18}$$

To check (3.5.17), we first note that the two sides are equal for $y=1$. For $y\in(0,1)$ we have

$$\log 1 - \log y = \int_y^1 \frac{1}{x}dx \ge 1-y,$$

which shows (3.5.17); and for $y>1$, we have

$$\log y - \log 1 = \int_1^y \frac{1}{x}dx \le y-1,$$

which again shows (3.5.17). To see (3.5.18) we can say that we have equality at $z=0$, and the derivative of the right-hand side is greater than the derivative of the left-hand side for all $z>0$.

Applying (3.5.17) to the expectation in (3.5.16), and then using (3.5.18) for the exponent in the expectation, we obtain

$$\log \mathbb{E}\left[\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] \leq \mathbb{E}\left[\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] - 1 \leq \frac{ct_N}{g^{-1}(t_N)}\mathbb{E}\left[\tilde{Y}\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right]. \tag{3.5.19}$$

Next we apply Hölder's inequality, which says that for real numbers $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ and for random variables Z_1, Z_2, we have

$$\mathbb{E}[|Z_1 Z_2|] \leq \mathbb{E}[|Z_1|^p]^{1/p}\mathbb{E}[|Z_2|^q]^{1/q}.$$

We apply Hölder's inequality with $Z_1 = \tilde{Y}$, $Z_2 = \exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)$, $p = \frac{1+\varepsilon_1}{\varepsilon_1}$, and $q = 1+\varepsilon_1$, where ε_1 was chosen in (d). Then we obtain

$$\mathbb{E}\left[\tilde{Y}\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] \leq \mathbb{E}\left[(\tilde{Y})^{\frac{1+\varepsilon_1}{\varepsilon_1}}\right]^{\frac{\varepsilon_1}{1+\varepsilon_1}} \mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right]^{\frac{1}{1+\varepsilon_1}}. \tag{3.5.20}$$

Note that the first factor on the right-hand side can be bounded above by a constant independent of N, since $Y \geq \tilde{Y}$ has finite moments (by applying Lemma 3.1.3 to g). Therefore, (3.5.19) and (3.5.20) show that there exists a constant $C_1 > 0$ such that for all N,

$$\log \mathbb{E}\left[\exp\left(\frac{ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] \leq \frac{C_1 t_N}{g^{-1}(t_N)}\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right], \tag{3.5.21}$$

where we also used that the expectation above is at least 1.

We now claim that there exists a constant C_2 such that for all large enough N,

$$\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] < C_2. \tag{3.5.22}$$

After proving this claim, we will be able to conclude the proof of Lemma 3.5.2 very quickly.

It will be useful in the proof of (3.5.22) to handle the left-hand side with an indicator on Y, so we will use the upper bound below:

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\right] \\
&= \mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{\tilde{Y}\geq 1\}}\right] + \mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{\tilde{Y}<1\}}\right] \\
&\leq \mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{Y\geq 1\}}\right] + e^{(1+\varepsilon_1)ct_N/g^{-1}(t_N)},
\end{aligned} \tag{3.5.23}$$

where in the inequality we used that $\left\{\tilde{Y} \geq 1\right\} \subseteq \{Y \geq 1\}$.

Now we give an upper bound on $t_N/g^{-1}(t_N)$ which we will use in our further calculations. Let $\eta' \in (0,1)$ be sufficiently small that $(1+\eta')/(1-\eta') \leq 1+\varepsilon_2$, where ε_2 was chosen in (e). Then $A\log N \leq t_N \leq (1+\eta')A\log N$ for N sufficiently large. Using that

g^{-1} is non-decreasing, and then applying Lemma 3.3.8(c), we get for N sufficiently large,

$$\frac{t_N}{g^{-1}(t_N)} \leq \frac{(1+\eta')A\log N}{g^{-1}(A\log N)} \leq \frac{(1+\eta')A\log N}{(1-\eta')A^{1/\beta}L_N} \leq (1+\varepsilon_2)A^{1-1/\beta}\frac{\log N}{L_N}. \tag{3.5.24}$$

Then it follows that the second term on the right-hand side of (3.5.23) is bounded by

$$e^{(1+\varepsilon_1)ct_N/g^{-1}(t_N)} \leq \exp\left((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\frac{\log N}{L_N}\right) \leq 2, \tag{3.5.25}$$

for N sufficiently large, where the second inequality follows by Corollary 3.3.9.

For the first term on the right-hand side of (3.5.23), we apply Lemma 3.1.2 with the random variable Y, and with $K_1 = 1$, $K_2 = (1+\tilde{\varepsilon})L_N$, $v = (1+\varepsilon_1)ct_N/g^{-1}(t_N)$. We obtain

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{Y\geq 1\}}\right] \\
&\leq \int_1^{(1+\tilde{\varepsilon})L_N} \frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)} e^{\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}s}\mathbb{P}(Y>s)ds + e^{\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}}\mathbb{P}(Y\geq 1).
\end{aligned}$$

Now using (3.5.24) and (3.5.25), it follows that for N sufficiently large,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{Y\geq 1\}}\right] \\
&\leq \int_1^{(1+\tilde{\varepsilon})L_N} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\frac{\log N}{L_N} \\
&\qquad \times \exp\left((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\frac{\log N}{L_N}s - g(s)\right) ds + 2.
\end{aligned}$$

Finally, we make the change of variables $u = s/L_N$ to arrive at a form which will be convenient to estimate:

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{Y\geq 1\}}\right] \\
&\leq \int_{1/L_N}^{1+\tilde{\varepsilon}} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \\
&\qquad \times \exp\left((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}(\log N)u - g(uL_N)\right) du + 2
\end{aligned} \tag{3.5.26}$$

for N sufficiently large.

To handle the integral on the right-hand side of (3.5.26), we split the domain into three subintervals. Let K be as in (3.2.3) (we assume without loss of generality that $K > 1$), and consider the integral on the interval $[1/L_N, K/L_N]$. Then we have

$$\begin{aligned}
&\int_{1/L_N}^{K/L_N} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \exp\Big((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}(\log N)u - g(uL_N)\Big)\,du \\
&\le \frac{K}{L_N}(1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \exp\left((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}(\log N)\frac{K}{L_N}\right) \\
&\le \varepsilon, \qquad (3.5.27)
\end{aligned}$$

for N sufficiently large, by Corollary 3.3.9.

Next, we assume N is sufficiently large that $L_N > K$, and consider the interval $(K/L_N, 1)$. By the assumption in (3.2.3) and then by Lemma 3.3.8(b), for N sufficiently large, for all $u \in (K/L_N, 1)$ we have

$$\frac{g(uL_N)}{uL_N} \ge \frac{g(L_N)}{L_N} \ge (1-\varepsilon_2)\frac{\log N}{L_N}.$$

Then for the integral in (3.5.26) on this subinterval for N sufficiently large we get

$$\begin{aligned}
&\int_{K/L_N}^{1} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \exp\Big((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}(\log N)u - g(uL_N)\Big)\,du \\
&\le \int_{K/L_N}^{1} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \\
&\quad \times \exp\Big(\Big((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta} - (1-\varepsilon_2)\Big)\,u\log N\Big)\,du \\
&\le C_3, \qquad (3.5.28)
\end{aligned}$$

for some constant $C_3 > 0$, where the last inequality can be seen as follows. By the choices in (c) and (e) we have

$$1-\varepsilon_2-(1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta} =: c_1 > 0.$$

Then calculating the integral in the second line, there exists a constant $C_3 > 0$ such that the integral is at most

$$C_3\left[-e^{-c_1 u\log N}\right]_{K/L_N}^{1} \le C_3.$$

This establishes (3.5.28).

Finally, we bound the integral in (3.5.26) on the interval $[1, 1+\tilde{\varepsilon}]$. On this subinterval, for N sufficiently large, for all $u \in [1, 1+\tilde{\varepsilon}]$, since g is non-decreasing we have

$$g(uL_N) \ge g(L_N) \ge (1-\varepsilon_2)\log N \ge (1-\varepsilon_2)(1+\tilde{\varepsilon})^{-1}u\log N,$$

where we applied Lemma 3.3.8(b) in the second inequality. Then for N sufficiently large, we can write

$$\begin{aligned}
&\int_1^{1+\tilde{\varepsilon}} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \exp\Big((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}(\log N)u - g(uL_N)\Big)\,du \\
&\le \int_1^{1+\tilde{\varepsilon}} (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta}\log N \\
&\qquad \times \exp\Big(\Big((1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta} - (1-\varepsilon_2)(1+\tilde{\varepsilon})^{-1}\Big)\,u\log N\Big)\,du \\
&\le \varepsilon, \qquad (3.5.29)
\end{aligned}$$

where the last inequality follows in a similar way to (3.5.28). Indeed, by the choices in (c) and (e) we have

$$(1-\varepsilon_2)(1+\tilde{\varepsilon})^{-1} - (1+\varepsilon_1)c(1+\varepsilon_2)A^{1-1/\beta} =: c_2 > 0.$$

Then there exists a constant $C_4 > 0$ such that the integral in the second line is at most

$$C_4\left[-e^{-c_2 u\log N}\right]_1^{1+\tilde{\varepsilon}} \le \varepsilon$$

for N sufficiently large. Hence, by (3.5.26)-(3.5.29) we conclude that for N sufficiently large,

$$\mathbb{E}\left[\exp\left(\frac{(1+\varepsilon_1)ct_N}{g^{-1}(t_N)}\tilde{Y}\right)\mathbb{1}_{\{Y\ge 1\}}\right] \le C_3 + 2\varepsilon + 2, \qquad (3.5.30)$$

which together with (3.5.23) and (3.5.25) shows (3.5.22).

Now we put (3.5.16), (3.5.21) and (3.5.22) together to obtain that for N sufficiently large

$$\begin{aligned}
\mathbb{P}\left(\tilde{S}_{t_N} > xg^{-1}(t_N)\right) &\le \exp\left(t_N\frac{C_1 t_N}{g^{-1}(t_N)}C_2 - cxt_N\right) \\
&\le \exp\left(t_N C_1 C_2(1+\varepsilon_2)A^{1-1/\beta}\frac{\log N}{(\log N)^{1/\beta-\varepsilon}} - cxt_N\right),
\end{aligned}$$

where in the second inequality we used (3.5.24) and Corollary 3.3.9. Since $t_N \ge A\log N$, by the definition of x in (3.5.13) and by (c) we have

$$cxt_N \ge c(\log 2)A^{1-1/\beta}(1+\varepsilon/2)A\log N = (1-\delta_2)(1+\varepsilon/2)(\log 2)A\log N.$$

We therefore conclude

$$\mathbb{P}\left(\tilde{S}_{t_N} > xg^{-1}(t_N)\right) \le N^{C_1C_2(1+\varepsilon_2)A^{1-1/\beta}\frac{t_N}{(\log N)^{1/\beta-\varepsilon}} - (1-\delta_2)(1+\varepsilon/2)(\log 2)A} \le N^{-A\log 2 - 1 - \delta_1}$$

for N sufficiently large, where the second inequality follows by (3.5.12) and (b). By (3.5.15), this finishes the proof of Lemma 3.5.2. □

Chapter 4

Genealogy of the N-particle branching random walk with stretched exponential tails

In this chapter we state a result on the genealogies of the N-BRW when the jump distribution X has stretched exponential tails given by $\mathbb{P}(X > x) = e^{-x^\beta}$ with $\beta \in (0, 1/2)$. The result says that for any large time t, with at least probability of order $(\log N)^{-1/2}$, there exists a time $k \in [t - 2\lceil \log_2 N \rceil, t - \lceil \log_2 N \rceil]$, such that a positive proportion of the time-t population descends from a single time-k particle. We then give a summary of the proof of this result. For some of the intermediate statements we will include the proof, but in many cases we will omit the details. Altogether we aim to give an idea of why the result is true and why we could not prove a stronger result with our method. This chapter is based on joint work with Sarah Penington.

4.1 Reminder of notation

Consider the N-BRW (as defined in Chapter 2 in Section 2.2.1, but note that we will work with a different jump distribution) and recall the following notation from Chapter 2. We refer to the glossary of notation in Section 2.7 for references to the section or equation where the notation is defined or first appears.

- $[n] = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$, and $[\![a, b]\!] = [a, b] \cap \mathbb{N}_0$ for $0 \le a \le b$
- $\mathcal{X}(n) = \{\mathcal{X}_1(n) \le \dots \le \mathcal{X}_N(n)\}$: ordered positions of the N particles at time n
- $X_{i,b,n}$: jump size of the ith particle's bth offspring at time n
- $\ell_N = \lceil \log_2 N \rceil$
- (i, n): ith particle from the left at time n

- $(i,n) \lesssim (j,n+k)$: particle (i,n) is the time-n ancestor of particle $(j,n+k)$
- $P_{i,n}^{j,n+k}$: path (sequence of jumps) between particles (i,n) and $(j,n+k)$ if $(i,n) \lesssim (j,n+k)$
- $G_x(n)$: set of particles to the right of or at position x at time n (see (2.2.7))
- $\mathcal{N}_{i,n}(n+k)$: set of time-$(n+k)$ descendants of the ith particle at time n (for $k \geq 0$)
- $\mathcal{N}_{i,n}^b(n+k)$: set of time-$(n+k)$ descendants of the ith particle's bth offspring at time n (for $k \geq 1$)
- $\mathcal{F}_n$: σ-algebra generated by the jumps $(X_{i,b,m},\ i \in [N], b \in \{1,2\}, m < n)$ (Recall that jumps made at time n are independent of $\mathcal{F}_n$; also see Section 2.2.4.)

We will also use the following two lemmas which were already stated and proved in Chapters 2 and 3.

Lemma 4.1.1. *Let $x \in \mathbb{R}$ and $n, k \in \mathbb{N}_0$. Then*

$$|G_x(n+k)| \geq \min\Big(N,\ 2^k |G_x(n)|\Big).$$

This statement is Lemma 2.2.3 in Chapter 2, where we prove this lemma.

Lemma 4.1.2. *[2, Lemma 1] Consider a pair of N-BRWs $(\mathcal{X}(n), \mathcal{X}^*(n))_{n\in\mathbb{N}_0}$, where $(\mathcal{X}(n))_{n\in\mathbb{N}_0}$ is determined by the jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and the initial configuration $\mathcal{X}(0)$, and $(\mathcal{X}^*(n))_{n\in\mathbb{N}_0}$ is determined by the same jumps $(X_{i,b,n})_{i\in[N],b\in\{1,2\},n\in\mathbb{N}_0}$ and the initial configuration $\mathcal{X}^*(0)$. Suppose that there exists a time $k \in \mathbb{N}_0$ such that $\mathcal{X}_i(k) \leq \mathcal{X}_i^*(k)$ for all $i \in [N]$. Then*

$$\mathcal{X}_i(n) \leq \mathcal{X}_i^*(n),$$

for all $n \geq k$ and $i \in [N]$.

This statement is Lemma 3.3.2 in Chapter 3, where we include the proof.

4.2 The main result

Let X be a random variable with stretched exponential tails given by

$$\mathbb{P}(X > x) = e^{-x^\beta}, \tag{4.2.1}$$

for $x \geq 0$ and for some $\beta \in (0, 1/2)$. That is, in this section we also assume that the jump distribution has a density given by

$$f(x) = \beta x^{\beta-1} e^{-x^\beta} \tag{4.2.2}$$

for all $x > 0$. We also define

$$L_N := (\log N)^{1/\beta}. \tag{4.2.3}$$

This is a special case of the definition of L_N in Chapter 3 (see (3.2.4)).

Consider the N-BRW with stretched exponential jump distribution, given by the random variable X. The decay of the tail of this jump distribution is faster than polynomial and slower than exponential. As the genealogies show completely different behaviours in the exponentially decaying case from the polynomially decaying one, we aim to investigate whether there is a change of behaviour in the stretched exponential case as we change the value of β. In the result we state below we study the genealogies of the N-BRW when $\beta \in (0, 1/2)$. In the future we are planning to study the case when $\beta \in [1/2, 1)$ (see Section 4.7).

To study the genealogies of the N-BRW with a stretched exponential jump distribution requires a more precise analysis than in the polynomial case. This can be seen for example by considering that the largest jump in a single time step of the N-BRW (i.e. the maximum of $2N$ jumps) is close to the size of the largest jump in of order $\log N$ steps (and both are close to L_N). Indeed, one can check that the largest jump in a single time step is greater than of size $(\log N - C/\log N)^{1/\beta}$ with high probability, and the largest jump in of order $\log N$ time steps is less than $(\log N + C \log\log N/\log N)^{1/\beta}$ with high probability for some large enough constant $C > 1$. Thus, there will be a large number of jumps close to the size of the largest in $2\ell_N$ steps, which is a problem we did not have in the polynomial case.

In this chapter we will assume that N is a power of 2, so

$$N = 2^{\ell_N}.$$

Let us also introduce the notation

$$x_j := \mathcal{X}_{N-2^j+1}(0), \tag{4.2.4}$$

that is, x_j denotes the position of the 2^j-th particle from the right at time zero. We can now state the main result of this chapter.

Theorem 4.2.1. *There exist $C > 0$ and $c > 0$ such that for N sufficiently large, for any $t \geq 2\ell_N$,*

$$\mathbb{P}\left(\exists k \in [\![t - 2\ell_N, t - \ell_N]\!],\ i \in [N] : |\mathcal{N}_{i,k}(t)| \geq cN\right) \geq \frac{C}{(\log N)^{1/2}}.$$

In the proof it will be practical to show a stronger statement, where the event also includes that all particles at time t are to the right of a well-chosen position y^*, which we will specify later on. Furthermore, to prove Theorem 4.2.1 we will condition on the configuration at time $t - 2\ell_N$, but without loss of generality, we will assume $t = 2\ell_N$ and prove results for any initial configuration, so the notation is slightly simpler. We define

$$P_N(y,c) := \mathbb{P}\left(\exists k \in [\![t-2\ell_N, t-\ell_N]\!],\ i\in[N] : |\mathcal{N}_{i,k}(t)| \geq cN \text{ and } \mathcal{X}_1(2\ell_N) > y\right) \quad (4.2.5)$$

and we will prove

$$P_N := P_N(y^*, c) \geq \frac{C}{(\log N)^{1/2}}, \quad (4.2.6)$$

for a well-chosen y^*.

Including the event $\{\mathcal{X}_1(2\ell_N) > y\}$ is helpful because of the following idea. We aim to say that for a well-chosen $y = y^*$, in order for N particles to reach position y^*, an unusually big jump must happen during the time interval $[0, \ell_N - 1]$, and we aim to show that the probability that such a big jump occurs and at least cN time-t particles descend from that big jump is at least of order $(\log N)^{-1/2}$.

We remark that the largest jump at time ℓ_N is not necessarily the one with the most descendants. If the largest jump is of size M and the particle performing this jump starts from position x at time ℓ_N, then other particles starting from to the right of position x might end up to the right of position $x+M$ even with smaller jumps than M.

We made an attempt to show that starting from an arbitrary configuration at time zero, by time ℓ_N the configuration should be such that the largest jump takes the lead with high probability. Our idea was the following. Let $\mathcal{X}_1(\ell_N) := x$. The largest jump at any time step is likely to be of size close to L_N. We were aiming to show $G_{x+\delta L_N}(\ell_N) = o(N^{(1-\delta)^\beta})$, no matter what the initial configuration is. If this were true, then the diameter of the particle cloud at time ℓ_N would be about L_N, and the largest jump from position $x+\delta$ would likely to be of size less than $(1-\delta)L_N$, and so it would end up to the left of position $x+L_N$. Since the number of particles very close to x is close to N in this case, there would be jumps of size roughly L_N starting from close to x, ending up roughly at position $L_N + x$. Then the largest of these jumps could take the lead, and studying the gap between this leader particle and the rest of the population, we could try to prove that the leader particle will have at least cN descendants at time $2\ell_N$. However, for certain choices of initial particle configuration we could not show that $G_{x+\delta}(\ell_N)$ will decay faster than $N^{(1-\delta)^\beta}$, in which case the above argument fails. In the following we will discuss a different approach which leads to the proof of Theorem 4.2.1

4.2.1 Outline of the proof of Theorem 4.2.1

Now we turn to the ideas of the proof of Theorem 4.2.1. In Section 4.4.1, we will define $r(k)$ for $k \in [\![1, \ell_N]\!]$, such that, for any $y \in \mathbb{R}$, on the event $\{\mathcal{X}_N(k) > y - r(k)\}$,

$$\mathbb{P}\left(\mathcal{X}_1(2\ell_N) > y \mid \mathcal{F}_k\right) > 1/2.$$

Note that on the event in the definition of $P_N(y,c)$ in (4.2.5), we have $\mathcal{X}_1(2\ell_N) > y$. A particular way for N particles to end up to the right of position y by time $2\ell_N$, is to have a particle which makes a jump at some time $k \in [\![0, \ell_N - 1]\!]$ which arrives to the right of position $y - r(k+1)$. Provided this happens, there is a positive constant probability that $\{\mathcal{X}_1(2\ell_N) > y\}$ occurs. In our proof we show that the probability that such jump happens, and the particle performing the jump has at least cN time-$2\ell_N$ descendants is at least of order $(\log N)^{-1/2}$. Let $A(y)$ denote the following event:

$$A(y) := \bigcup_{(i,b,k)\in[N]\times\{1,2\}\times[0,\ell_N-1]} \{\mathcal{X}_i(k) + X_{i,b,k} > y - r(k+1)\}.$$

Let $c > 0$ a small constant. In our proof, for some position y^*, we will find a lower bound on $P_N(y^*, c)$ and an upper bound on the expected number of particles to the right of y^* (divided by N), both given roughly by $\mathbb{P}(A(y^*))$. Comparing the two bounds leads to the proof of Theorem 4.2.1.

We will define a function $h(y)$ roughly as

$$h(y) \approx C_1 N (\log N)^{1/2} \mathbb{P}(A(y)),$$

for some $C_1 > 0$, and set y^* such that

$$h(y^*) = cN.$$

(The precise definition of $h(y)$ can be found in Section 4.4.2.) One of the main steps to prove Theorem 4.2.1 is to show

$$\mathbb{E}[|G_{y^*}(2\ell_N)|] \leq h(y^*) = cN.$$

We discuss this step in Section 4.6. The other main step of the proof of Theorem 4.2.1 is to show (using that $\mathbb{E}[|G_{y^*}|] \leq cN$) that for some $C_2 > 0$,

$$P_N(y^*, c) \geq C_2 \mathbb{P}(A(y^*)). \tag{4.2.7}$$

We give a sketch proof of this in Section 4.5. Now the definitions of $h(y)$ and y^* give

$$\mathbb{P}(A(y^*)) \approx \frac{h(y^*)}{C_1 N (\log N)^{1/2}} = \frac{c}{C_1 (\log N)^{1/2}},$$

which together with (4.2.7) shows (4.2.6), from which we can conclude Theorem 4.2.1.

The motivation for this argument comes from the fact that

$$\mathbb{E}[|G_y(t)|] = \mathbb{E}\left[|G_y(t)|\mathbb{1}_{A(y)}\right] + \mathbb{E}\left[|G_y(t)|\mathbb{1}_{(A(y))^c}\right] \leq N\mathbb{P}(A(y)) + \mathbb{E}\left[|G_y(t)|\mathbb{1}_{(A(y))^c}\right].$$

If we had shown $\mathbb{E}\left[|G_{y^*}(t)|\mathbb{1}_{(A(y^*))^c}\right] \leq \mathbb{E}\left[|G_{y^*}(t)|\mathbb{1}_{A(y^*)}\right]$, it would have meant that there is a jump at some time k which arrives to the right of $y^* - r(k+1)$ on most paths which end up to the right of y^* at time t. This would have implied that P_N is of constant order rather than of order $(\log N)^{-1/2}$ (because we could have defined $h(y)$ without the $(\log N)^{1/2}$ term). Our proof however does not show this; in the following we explain the main steps of the above argument and give an idea of where the $(\log N)^{1/2}$ term comes from.

4.3 Fairly big jumps and large deviation result

One of the main tasks in proving Theorem 4.2.1 is to upper bound $\mathbb{E}[|G_y(2\ell_N)|]$, i.e. the expected number of particles to the right of some position y at time $2\ell_N$. The first useful property which we will need for this upper bound is stated in Lemma 4.3.1 below. Recall that we assume $\beta \in (0, 1/2)$ in (4.2.1). Let

$$\delta_N := (\log N)^{-\gamma}, \text{ for some } \gamma \in (1/(1-\beta), 1/\beta), \tag{4.3.1}$$

and let

$$\rho_N := 100\delta_N^{1-\beta} L_N. \tag{4.3.2}$$

From now on, jumps of size larger than $\delta_N L_N$ will be called *fairly big jumps*.

Lemma 4.3.1 says that the probability that a path without fairly big jumps moves more than ρ_N in $2\ell_N$ time is very small.

Lemma 4.3.1. *Let X_i, $i = 1, 2, \ldots$ be i.i.d. random variables with the same distribution as X in* (4.2.1)*. Then for N sufficiently large,*

$$\mathbb{P}\left(\sum_{i=1}^{2\ell_N} X_i \mathbb{1}_{\{X_i \le \delta_N L_N\}} > \rho_N\right) < N^{-49}.$$

The proof of this lemma is very similar to the proof of Lemma 3.5.2 in the proof of the asymptotic speed result. A consequence of Lemma 4.3.1 is, that we will be able to assume that the sum of jumps which are smaller than $\delta_N L_N$ is less than ρ_N on every path between time 0 and $2\ell_N$. Let H_N denote the event that we have just described. Recall the notation for paths from (2.2.10). We let

$$H_N := \left\{ \sum_{(i,b,k) \in P_{i_1,0}^{i_2,2\ell_N}} X_{i,b,k} \mathbb{1}_{\{X_{i,b,k} \le \delta_N L_N\}} < \rho_N, \forall i_1, i_2 \in [N] \right\}. \tag{4.3.3}$$

Next, we state a consequence of Lemma 4.3.1 in terms of $\mathbb{E}[|G_y(2\ell_N)|]$, because this is how we will use it.

Corollary 4.3.2. *For all $\varepsilon > 0$, for N sufficiently large,*

$$\mathbb{E}[|G_y(2\ell_N)|] \le \mathbb{E}[|G_y(2\ell_N)| \mathbb{1}_{H_N}] + \varepsilon.$$

Proof. Consider the paths in N independent BRWs coupled with the N-BRW (see Section 2.4.1). The total number of such paths is $N2^{2\ell_N}$ by time $2\ell_N$. By Lemma 2.4.1, the event H_N^c implies that there must be at least one among the $N2^{2\ell_N}$ paths which moves more than ρ_N without fairly big jumps. Therefore, by Lemma 4.3.1,

$$\mathbb{E}\left[|G_y(2\ell_N)|\mathbb{1}_{H_N^c}\right] \leq N2^{2\ell_N}N^{-49} < \varepsilon, \tag{4.3.4}$$

for N sufficiently large. □

The next lemma gives an upper bound on the probability that a path moves a distance more than x with the restriction that every jump on the path is between sizes d and M (the lemma will be useful for us with setting $d = \delta_N L_N$ and $M \approx L_N$, in which case d is much smaller than M).

Assume that there are m jumps on such a path. Note that if M is close x then a single jump can still cover roughly the whole distance x, so in this case the restriction that jumps are smaller than M is less significant.

The first part of the lemma below says that if M is not a significant restriction, then the upper bound is given by roughly the probability that a jump is of size greater than $x-(m-1)d$ and the remaining $m-1$ jumps are greater than d. Then it can be checked that this upper bound is largest when $m = 1$, that is, when the distance is covered by a single jump.

The second part says that if M is a more significant restriction, then the upper bound is given by roughly the probability that a jump is of size greater than M, another jump is of size greater than $x - M - (m-2)d$, and the remaining $m-2$ jumps are of size greater than d. If M is much larger than d then it can be seen that this probability is largest when $m = 2$.

Lemma 4.3.3. *Let X_i, $i = 1, 2, \ldots$ be i.i.d. random variables with the same distribution as X in (4.2.1). Let $m \geq 2$ be an integer, let $x, d, M \in \mathbb{R}$ with $0 < d \leq M$ and $x > md$. Let $S_m := \sum_{i=1}^m X_i$. Then*

- *If $x - M - (m-1)d < 0$ then*

$$\mathbb{P}\left(S_m > x,\ d < X_i < M,\ i = 1, \ldots, m\right) \leq (1 + x^\beta)^{m-1} e^{-(m-1)d^\beta - (x-(m-1)d)^\beta}.$$

- *If $x - M - (m-1)d \geq 0$ then*

$$\mathbb{P}\left(S_m > x,\ d < X_i < M,\ i = 1, \ldots, m\right) \leq (1+x^\beta)^{m-2} M^\beta e^{-(m-2)d^\beta - M^\beta - (x-M-(m-2)d)^\beta}.$$

Lemma 4.3.3 is needed to show that the most likely way for paths to arrive to the right of position y^* at time $2\ell_N$ is to have only two fairly big jumps of size larger than $\delta_N L_N$ on the path. We do not include the proof of this lemma.

4.4 Proof of Theorem 4.2.1

4.4.1 Choice of $r(k)$

For $k \in [\![1, \ell_N - 1]\!]$, we define $r(k)$ by

$$(k+1)2^{2\ell_N - k}\mathbb{P}\left(X > r(k)\right) = 32N, \tag{4.4.1}$$

and for $k = \ell_N$ we let

$$r(\ell_N) = 0. \tag{4.4.2}$$

With the next lemma we show that $r(k)$ is as we described in Section 4.2.1; that is, given that there is at leasts one particle to the right of position $y - r(k)$ at time k, we will have $\mathcal{X}_1(2\ell_N) > y$ with a positive constant probability.

Lemma 4.4.1. *For all $k \in [\![0, \ell_N]\!]$ and for all $y \in \mathbb{R}$, on the event $\{\mathcal{X}_N(k) > y - r(k)\}$,*

$$\mathbb{P}\left(\mathcal{X}_1(2\ell_N) > y \mid \mathcal{F}_k\right) > 1 - e^{-6}.$$

Sketch proof. For $k = \ell_N$, on the event $\{\mathcal{X}_N(\ell_N) > y\}$, we deterministically have $\mathcal{X}_1(2\ell_N) > y$ by Lemma 4.1.1.

Now we deal with the case when $k \in [\![1, \ell_N - 1]\!]$. Let $X_{s,j}$, $s, j = 1, 2 \ldots$ be i.i.d. random variables with the same distribution as X. For $k \in [\![1, \ell_N - 1]\!]$, on the event $\{\mathcal{X}_N(k) > y - r(k)\}$ we can lower bound $|G_y(2\ell_N)|$ by roughly

$$\left(\frac{1}{2}\sum_{s=k+1}^{2\ell_N - 1}\sum_{j=1}^{2(2^{s-k}\wedge N)} \mathbb{1}_{\{X_{s,j} > r(k)\}}(2^{2\ell_N - s - 1} \wedge N)\right) \wedge N. \tag{4.4.3}$$

The reason for this is the following. On the event $\{\mathcal{X}_N(k) > y - r(k)\}$, by Lemma 4.1.1, there must be at least $2^{s-k} \wedge N$ particles to the right of position $y - r(k)$ at any time $s \in [\![k+1, 2\ell_N - 1]\!]$. Then the children of these particles make at least $2(2^{s-k} \wedge N)$ jumps, which arrive to the right of position y at time $s+1$ if the jump is larger than $r(k)$. A time-$(s+1)$ particle to the right of y will have $2^{2\ell_N - s - 1}$ surviving descendants unless some of the descendants are killed in the selection steps by time $2\ell_N$, but that is only possible if there are N particles to the right of position y by time $2\ell_N$.

There is an issue about double counting particles that are already to the right of y and make a jump of size $r(k)$. It is possible (but not straightforward) to show that the inequality holds with a factor of $1/2$ in front of the sum.

Now,

$$\sum_{s=\ell_N}^{k+\ell_N} 2^{s-k} \cdot 2^{2\ell_N - s - 1} = (k+1)2^{2\ell_N - k - 1};$$

thus, since the random variables $X_{s,j}$ are i.i.d. and by the definition of $r(k)$ in (4.4.1), we have

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{2}\sum_{s=k+1}^{2\ell_N-1}\sum_{j=1}^{2(2^{s-k}\wedge N)}\mathbb{1}_{\{X_{s,j}>r(k)\}}(2^{2\ell_N-s-1}\wedge N)\right]\\&\geq\frac{1}{2}\sum_{s=\ell_N}^{k+\ell_N}2(2^{s-k}\wedge N)\mathbb{P}\left(X>r(k)\right)\left(2^{2\ell_N-s-1}\wedge N\right)\\&\geq 16N.\end{aligned}$$

From here one can use Theorem 2.3(c) in [35] from McDiarmid to conclude the lemma. We omit the further details of this argument. □

Observe that by (4.4.1), we have

$$r(k)=(\log N-(\log 2)k+\log(k+1)-\log 32)^{1/\beta} \tag{4.4.4}$$

for $k\in[\![0,\ell_N-1]\!]$. Thus, we think of $r(k)$ as a monotone decreasing function in k which is roughly given by $(\log N-(\log 2)k)^{1/\beta}$.

Consider paths starting from the right of position x_j (see (4.2.4)). Then (by Lemma 4.1.1) at time ℓ_N-j we will have $|G_{x_j}(\ell_N-j)|=N$, and it is likely that at least one of the particles in $G_{x_j}(\ell_N-j)$ will perform a jump of size roughly L_N. Then again by Lemma 4.1.1 there will be at least 2^j particles to the right of position roughly x_j+L_N at time ℓ_N, and it is likely that one of the 2^j particles will make a jump of size roughly $(j\log 2)^{1/\beta}$ in the next step. As a result (by Lemma 4.1.1) there will be N particles to the right of roughly $x_j+L_N+(j\log 2)^{1/\beta}$ at time $2\ell_N$. Note that by (4.4.4), we have

$$(j\log 2)^{1/\beta}\approx r(\ell_N-j). \tag{4.4.5}$$

Also recall that we will want to choose y^* such that $\mathbb{E}\left[G_{y^*}(2\ell_N)\right]\leq cN$ for some $c\in(0,1)$. Therefore, we will think of y^*-x_j as a slightly larger distance than

$$L_N+(j\log 2)^{1/\beta}\approx L_N+r(\ell_N-j)$$

for all $j\in[\![0,\ell_N]\!]$.

In the proof of Theorem 4.2.1 we need another function, $R(k,y)$ as well. For $k\in[\![0,\ell_N-1]\!]$, we choose $R(k,y)$ such that all particles are to the right of $y-R(k,y)$ at time k with probability greater than 1/2; that is, we choose $R(k,y)$ such that

$$\mathbb{P}\left(\mathcal{X}_1(k)>y-R(k,y)\right)>1/2. \tag{4.4.6}$$

We note here that we will not present every step of the proof of Theorem 4.2.1, in particular, we will not deal with most steps involving the function $R(k,y)$. We therefore do not give a precise definition of this function; instead we will refer to the property in (4.4.6) when we talk about $R(k,y)$.

4.4.2 Proof of Theorem 4.2.1

Let $K_1, K_2 > 1$ be some large constants, and define

$$h_1(y) := K_1(\log N)^{1/2} N \sum_{k=0}^{\ell_N-1} \sum_{j=0}^{\ell_N-k} 2^{j+k} \mathbb{P}\left(X > y - x_j - r(k+1) - \rho_N\right), \tag{4.4.7}$$

$$h_2(y) := K_2(\log N)^{1/2} N^2 \sum_{k=0}^{\ell_N-1} \mathbb{P}\left(X > R(k,y) - r(k+1) - \rho_N\right) \tag{4.4.8}$$

and

$$h(y) := h_1(y) \vee h_2(y),$$

where x_j, ρ_N, $r(k)$ are defined in (4.2.4), (4.3.2), (4.4.1)-(4.4.2) respectively and $R(k,y)$ is as in (4.4.6).

We define y^* by

$$h(y^*) = cN, \tag{4.4.9}$$

for some $c \in (0, 1/100)$. This is possible because both $h_1(y)$ and $h_2(y)$ are continuous functions of y (assuming (4.2.1)), and for a very small y both functions are larger than N, and for a very large y they are $o(N)$. Now we state the two main intermediate results that we need to prove Theorem 4.2.1.

Proposition 4.4.2. *For N sufficiently large,*

$$\mathbb{E}\left[|[G_{y^*}(2\ell_N)]\right] \leq h(y^*) = cN.$$

We describe the ideas of the proof of Proposition 4.4.2 in Section 4.6.

Proposition 4.4.3. *There exists $C_1 > 0$ such that for N sufficiently large,*

$$P_N \geq C_1 \sum_{k=0}^{\ell_N-1} \sum_{j=0}^{\ell_N-k} 2^{j+k} \mathbb{P}\left(X > y^* - x_j - r(k) - \rho_N\right),$$

and

$$P_N \geq C_1 \sum_{k=0}^{\ell_N-1} N\mathbb{P}\left(X > R(k,y^*) - r(k) - \rho_N\right),$$

where P_N is given by (4.2.6)*, y^* by* (4.4.9)*, x_j, ρ_N, $r(k)$ are defined in* (4.2.4), (4.3.2), (4.4.1)-(4.4.2) *respectively, and $R(k,y)$ is as in* (4.4.6).

We give a sketch proof of the first inequality in Proposition 4.4.3 in Section 4.5. The proof of the second inequality is very similar to the first, but we will not give any detail on this part.

Proof of Theorem 4.2.1. By Proposition 4.4.3 and by the definitions of the functions h_1, h_2 and h, and by the choice of y^*,

$$P_N \geq \frac{C_1}{K_1(\log N)^{1/2}N}h_1(y^*) \vee \frac{C_2}{K_2(\log N)^{1/2}N}h_2(y^*) = \frac{C_3}{(\log N)^{1/2}N}h(y^*) = \frac{C_3 c}{(\log N)^{1/2}},$$

which shows the inequality in (4.2.6), from which we can conclude Theorem 4.2.1. □

Remark. *We use Proposition 4.4.2 in the proof of Proposition 4.4.3.*

Remark. *The function $h_1(y)$ is more convenient for showing our calculations than the function $h_2(y)$. In the rest of this argument we will prove the first lower bound in Proposition 4.4.3, and we will outline the ideas for proving that the size of a subset of $G_{y^*}(2\ell_N)$ can be bounded above by $\varepsilon h_1(y^*)$ for any constant $\varepsilon > 0$. For other subsets we need a slightly different argument which gives the upper bound $h_2(y^*)$. We will not go into detail on that part of the proof.*

4.5 Sketch proof of Proposition 4.4.3 – lower bound on P_N

In this section we discuss the main steps of the proof of the first inequality of Proposition 4.4.3. Including the term ρ_N (defined in (4.3.2)) in that inequality makes the proof slightly longer but does not change the key ideas, partly because $\rho_N = o(L_N/(\log N))$ as $N \to \infty$. Thus, we will go over the proof that

$$P_N \geq C_1 \sum_{k=0}^{\ell_N-1} \sum_{j=0}^{\ell_N-k} 2^{j+k}\mathbb{P}\left(X > y^* - x_j - r(k+1)\right), \tag{4.5.1}$$

noting that the precise argument would need some adjustment.

4.5.1 Good events

Let us define the function $f(i,k)$ for $i \in [N]$ and $k \in [\![0, \ell_N - 1]\!]$ as

$$f(i,k) := \lceil \log_2(N - i + 1) \rceil - k, \tag{4.5.2}$$

that is, we have $f(i,k) = j$ if

$$N - 2^{j+k} + 1 \leq i \leq N - 2^{j+k-1}. \tag{4.5.3}$$

Then by the definition of x_j and by Lemma 4.1.1, we have

$$\mathcal{X}_i(k) \geq x_{f(i,k)}, \tag{4.5.4}$$

for all $i \in [N]$ and $k \in [0, \ell_N - 1]$, and it is possible that we have $\mathcal{X}_i(k) < x_{f(i,k)-1}$. Then (4.5.4) implies that

$$\begin{aligned}\{X_{i,b,k} > y^* - r(k+1) - x_{f(i,k)}\} &\subseteq \{\mathcal{X}_i(k) + X_{i,b,k} > y^* - r(k+1)\} \\ &= \{\mathcal{X}_l(k+1) > y^* - r(k+1)\},\end{aligned} \tag{4.5.5}$$

if $l \in \mathcal{N}^b_{i,k}(k+1)$.

Next, for $i \in [N]$, $b \in \{1,2\}$ and $k \in [\![0, \ell_N - 1]\!]$ we first define then explain the following events:

- $A_{i,b,k} := \{X_{i,b,k} > y^* - r(k+1) - x_{f(i,k)}\}$,
- $B_{i,b,k} := \bigcap_{k_1 \in [\![0,\ell_N]\!], i_1 \in [N] \setminus \mathcal{N}^b_{i,k}(k_1)} \{\mathcal{X}_{i_1}(k_1) \leq y^* - r(k_1)\}$,
- $C_{i,b,k} := \{|G_{y^*}(2\ell_N) \setminus \mathcal{N}^b_{i,k}(2\ell_N)| \leq c'N\}$, for some constant $c' \in (100c, 1)$, with c as in (4.4.9),
- $D = \{|G_{y^*}(2\ell_N)| = N\}$.

The event $A_{i,b,k}$ says that the jump $X_{i,b,k}$ is larger than $y^* - r(k+1) - x_{f(i,k)}$, and thus by (4.5.5) it implies that there is at least one particle to the right of position $y^* - r(k+1)$ at time $k+1$; that is, on the event $A_{i,b,k}$ we have

$$\mathcal{X}_N(k+1) \geq y^* - r(k+1). \tag{4.5.6}$$

(Note that if the particle performing the jump $X_{i,b,k}$ does not survive the selection step, then there is not one, but N particles to the right of $y^* - r(k+1)$ at time $k+1$.) This property will be important because of Lemma 4.4.1.

The event $B_{i,b,k}$ means that every single particle which is not descended from the jump $X_{i,b,k}$ is to the left of position $y^* - r(k_1)$ at all times in the interval $[\![0, \ell_N]\!]$ (thus, on the event $A_{i,b,k} \cap B_{i,b,k}$ the particle that performed the jump $X_{i,b,k}$ survives the selection step). The event $C_{i,b,k}$ says that the number of time-$2\ell_N$ particles, which are to the right of position y^* and are not descended from the jump $X_{i,b,k}$, is less than $c'N$. The event D says that all particles are to the right of y^* at time $2\ell_N$.

We also define the event

$$E := \bigcup_{(i,b,k) \in [N] \times \{1,2\} \times [\![0,\ell_N - 1]\!]} C_{i,b,k} \cap D,$$

and notice that the event E is equivalent to the event in the definition of P_N (see (4.2.5)-(4.2.6)); that is,

$$P_N = \mathbb{P}(E). \tag{4.5.7}$$

Hence, to prove (4.5.1), it will be enough to lower bound the probability of the event

$$\bigcup_{(i,b,k)\in[N]\times\{1,2\}\times[\![0,\ell_N-1]\!]} A_{i,b,k}\cap B_{i,b,k}\cap C_{i,b,k}\cap D\subseteq E. \tag{4.5.8}$$

First we claim that for $k'\leq k$ we have

$$B_{i,b,k}\cap A_{i',b',k'}=\emptyset, \tag{4.5.9}$$

if $(i,b,k)\neq(i',b',k')$. Now we prove this claim.

On the event $B_{i,b,k}$ we have

$$\mathcal{X}_j(k'+1)\leq y^*-r(k'+1) \tag{4.5.10}$$

for all $j\notin\mathcal{N}^b_{i,k}(k'+1)$. Let $j^*\in\mathcal{N}^{b'}_{i',k'}(k'+1)$. If $k'<k$ then $j^*\notin\mathcal{N}^b_{i,k}(k'+1)$, because in this case $\mathcal{N}^b_{i,k}(k'+1)=\emptyset$. If $k'=k$ then $j^*\notin\mathcal{N}^b_{i,k}(k'+1)$ still holds, since in this case $j^*\in\mathcal{N}^{b'}_{i',k'}(k'+1)=\mathcal{N}^{b'}_{i',k}(k+1)$, and $\mathcal{N}^{b'}_{i',k}(k+1)\cap\mathcal{N}^b_{i,k}(k+1)=\emptyset$. We conclude that (4.5.10) applies for $j=j^*$, thus

$$\mathcal{X}_{j^*}(k'+1)\leq y^*-r(k'+1).$$

On the other hand, on the event $A_{i',b',k'}$ we have $X_{i',b',k'}>y^*-r(k'+1)-x_{f(i',k')}$, which implies $\mathcal{X}_{j^*}(k'+1)>y^*-r(k')$ because of (4.5.5). This concludes the proof of (4.5.9) for $k'\leq k$, and as a consequence we have

$$A_{i,b,k}\cap B_{i,b,k}\cap A_{i',b',k'}\cap B_{i',b',k'}=\emptyset$$

for all $(i,b,k)\neq(i',b',k')$. We then conclude that the set of events

$$\{A_{i,b,k}\cap B_{i,b,k}\cap C_{i,b,k}\cap D,\ (i,b,k)\times\{1,2\}\times[\![0,\ell_N-1]\!]\}$$

consists of pairwise disjoint events. Therefore, by (4.5.8),

$$\mathbb{P}(E)\geq\sum_{(i,b,k)\in[N]\times\{1,2\}\times[\![0,\ell_N-1]\!]}\mathbb{P}\left(A_{i,b,k}\cap B_{i,b,k}\cap C_{i,b,k}\cap D\right). \tag{4.5.11}$$

4.5.2 Coupling with modified N-BRWs

At the moment the event $A_{i,b,k}$ is not independent of $B_{i,b,k}$ or $C_{i,b,k}$, because it affects for example the number of particles in $\mathcal{N}^b_{i,k}(k_1)$ in the definition of $B_{i,b,k}$, and also the number of surviving particles not in $\mathcal{N}^b_{i,k}(2\ell_N)$ in the definition of $C_{i,b,k}$. To overcome this problem, we define the modified processes $\mathcal{X}^{i,b,k}$ for $(i,b,k)\in[N]\times\{1,2\}\times[\![0,\ell_N-1]\!]$ as follows. Fix $(i^*,b^*,k^*)\in[N]\times\{1,2\}\times[\![0,\ell_N-1]\!]$. Let $(\mathcal{X}^{i^*,b^*,k^*}(n))_{n\in[\![0,k^*]\!]}$ be the N-BRW determined by the original initial configuration $\mathcal{X}(0)$ and by the jumps $(X_{i,b,k})_{(i,b,k)\in[N]\times\{1,2\}\times[\![0,k^*-1]\!]}$, so that we have

$$(\mathcal{X}^{i^*,b^*,k^*}(n))_{n\in[\![0,k^*]\!]} = (\mathcal{X}(n))_{n\in[\![0,k^*]\!]}.$$

Then, in the selection step at time k^*+1 we first kill the particle that made the jump X_{i^*,b^*,k^*}, and then keep the N rightmost particles. That is, $\mathcal{X}^{i^*,b^*,k^*}(k^*+1) = \{\mathcal{X}_1^{i^*,b^*,k^*}(k^*+1) \leq \cdots \leq \mathcal{X}_N^{i^*,b^*,k^*}(k^*+1)\}$ is given by the N largest numbers from the collection

$$(\mathcal{X}_i^{i^*,b^*,k^*}(k^*) + X_{i,b,k^*})_{(i,b)\in([N]\times\{1,2\})\setminus(i^*,b^*)}.$$

After time k^*+1 the process $(\mathcal{X}^{i^*,b^*,k^*}(n))_{n\in[\![k^*+2,2\ell_N]\!]}$ is determined by the jumps $(X_{i,b,k})_{(i,b,k)\in[N]\times\{1,2\}\times[\![k^*+1,2\ell_N-1]\!]}$ as follows. For all $k\in[\![0,2\ell_N]\!]$ let $\mathcal{N}(k)$ denote the set of time-k descendants of the particle that made the jump X_{i^*,b^*,k^*}, and let D_k denote the number of these descendants:

$$\mathcal{N}(k) := \mathcal{N}_{i^*,k^*}^{b^*}(k), \tag{4.5.12}$$

and

$$D_k := |\mathcal{N}(k)|.$$

(Note that for $k\leq k^*$ we have $\mathcal{N}(k)=\emptyset$.) Furthermore, the ordered indices in $\mathcal{N}(k)$ will be denoted by

$$n_1(k) < n_2(k) < \cdots < n_{D_k}(k),$$

and the ordered indices in $[N]\setminus\mathcal{N}(k)$ by

$$n_1^c(k) < n_2^c(k) < \cdots < n_{N-D_k}^c(k).$$

We now re-index the jumps $X_{i,b,k}$ in such a way that the children of the $N-D_k$ rightmost particles in $\mathcal{X}^{i^*,b^*,k^*}$ perform the same jumps as the children of the $N-D_k$ particles in $\mathcal{X}$ which are not descended from the jump X_{i^*,b^*,k^*}. The remaining D_k (leftmost) particles' children in $\mathcal{X}^{i^*,b^*,k^*}$ will perform the jumps of the D_k particles' children which descend from the jump X_{i^*,b^*,k^*}.

Formally, for all $k\in[0,2\ell_N]$ we define the bijection $\sigma_k:[N]\to[N]$ (from the indices of $\mathcal{X}^{i^*,b^*,k^*}(k)$ to the indices of $\mathcal{X}(k)$) by

$$\sigma_k(i) = \sigma_k^{i^*,b^*,k^*}(i) := \begin{cases} n_i(k), & \text{if } i\leq D_k \\ n_{i-D_k}^c(k) & \text{if } i > D_k, \end{cases} \tag{4.5.13}$$

and then we have

$$\sigma_k^{-1}(j) = \begin{cases} |\mathcal{N}(k)\cap[1,j]|, & \text{if } j\in\mathcal{N}(k) \\ |([N]\setminus\mathcal{N}(k))\cap[1,j]| + D_k, & \text{if } j\notin\mathcal{N}(k). \end{cases} \tag{4.5.14}$$

Now let $(\mathcal{X}^{i^*,b^*,k^*}(n))_{n\in[\![k^*+2,2\ell_N]\!]}$ be the N-BRW determined by the jumps $(X_{\sigma_k(i),b,k})_{(i,b,k)\in[N]\times\{1,2\}\times[\![k^*+1,2\ell_N-1]\!]}$ the usual way: for all $k\in[\![k^*+1,2\ell_N]\!]$, the ordered positions $\{\mathcal{X}_1^{i^*,b^*,k^*}(k+1) \le \cdots \le \mathcal{X}_N^{i^*,b^*,k^*}(k+1)\} = \mathcal{X}^{i^*,b^*,k^*}(k+1)$ are given by the N largest numbers from the collection

$$(\mathcal{X}_i^{i^*,b^*,k^*}(k) + X_{\sigma_k(i),b,k})_{(i,b)\in[N]\times\{1,2\}}.$$

Note that with this definition, $\mathcal{X}^{i^*,b^*,k^*}$ is independent of the jump X_{i^*,b^*,k^*}. Also note that for all $k \le k^*$, we have $D_k = 0$ and hence $\sigma_k(i) = i$ for all $i \in [N]$ in this case.

We also define another process

$$(\tilde{\mathcal{X}}^{i^*,b^*,k^*}(k))_{k\in[\![0,2\ell_N]\!]},$$

as the N-BRW determined by the jumps $(X_{\sigma_k(i),b,k})_{(i,b,k)\in[N]\times\{1,2\}\times[\![0,2\ell_N-1]\!]}$ and by the initial configuration $\mathcal{X}(0)$. That is, $(\tilde{\mathcal{X}}^{i^*,b^*,k^*}(k))_{k\in[\![0,2\ell_N]\!]}$ is distributed as an N-BRW, and, at all times k, the ith particle's children make exactly the same jumps as the ith particle's children at time k in the process $(\mathcal{X}^{i^*,b^*,k^*}(k))_{k\in[\![0,2\ell_N]\!]}$. The difference between the two processes is that in $(\tilde{\mathcal{X}}^{i^*,b^*,k^*}(k))_{k\in[\![0,2\ell_N]\!]}$ the selection step is not modified at time k^*+1.

With the lemmas below we compare the processes $\mathcal{X}$, $\mathcal{X}^{i^*,b^*,k^*}$ and $\tilde{\mathcal{X}}^{i^*,b^*,k^*}$. First, Lemma 4.5.1 says that $\tilde{\mathcal{X}}^{i^*,b^*,k^*}(k)$ dominates $\mathcal{X}^{i^*,b^*,k^*}(k)$ for all $k \in [\![0, 2\ell_N]\!]$, and that the positions of particles not descended from the jump X_{i^*,b^*,k^*} in $\mathcal{X}$ are dominated by the corresponding positions in $\mathcal{X}^{i^*,b^*,k^*}$.

Lemma 4.5.1. *For all $(i^*, b^*, k^*) \in [N] \times \{1, 2\} \times [\![0, \ell_N - 1]\!]$ and $(i, k) \in [N] \times [\![0, 2\ell_N]\!]$, we have*

$$\mathcal{X}_i^{i^*,b^*,k^*}(k) \le \tilde{\mathcal{X}}_i^{i^*,b^*,k^*}(k),$$

and if $i > D_k$ then

$$\mathcal{X}_{\sigma_k(i)}(k) \le \mathcal{X}_i^{i^*,b^*,k^*}(k).$$

The main point of the proof of this lemma is to check that the statements hold at time k^*+1 (where we have the modified selection step in $\mathcal{X}^{i^*,b^*,k^*}$), and then use Lemma 4.1.2 and a similar argument to the proof of Lemma 4.1.2 to conclude the result. We will not include the details of this proof.

For all $z \in \mathbb{R}$, $n \in [\![0, 2\ell_N]\!]$ and $(i, b, k) \in [N] \times \{1, 2\} \times [\![0, \ell_N - 1]\!]$, let us write $G_z^{i,b,k}(n)$ and $\tilde{G}_z^{i,b,k}(n)$ for the sets of particles to the right of or at position z at time n in the processes $\mathcal{X}^{i,b,k}$ and $\tilde{\mathcal{X}}^{i,b,k}$ respectively. The next lemma is about $G_z^{i,b,k}(n)$ and it is a very similar statement to Lemma 4.1.1.

Lemma 4.5.2. *Let $k_1 \in [\![0, 2\ell_N]\!]$, $k_2 \in [\![k_1, 2\ell_N]\!]$ and $z \in \mathbb{R}$. Then for all $(i, b, k) \in [N] \times \{1, 2\} \times [\![0, \ell_N - 1]\!]$,*

$$|G_z^{i,b,k}(k_2)| \ge \left(2^{k_2-k_1-1}|G_z^{i,b,k}(k_1)|\right) \wedge N.$$

We do not have exactly the same statement as in Lemma 4.1.1 for $G_z^{i,b,k}(k_2)$, because the modified selection step might make a small difference if it happens during the time interval $[k_1, k_2]$. Then it is possible that the particle that makes the jump $X_{i,b,k}$ descends from the set $G_z^{i,b,k}(k_1)$, and since this particle is killed, there will be fewer particles in the set $G_z^{i,b,k}(k_2)$. Taking this into account, the lower bound from Lemma 4.1.1 still works if we subtract 1 from the exponent of 2 on the right-hand side. We will not write out the precise proof of this statement.

Now we introduce the events $B'_{i,b,k}$ and $C'_{i,b,k}$ which are similar to the events $B_{i,b,k}$ and $C_{i,b,k}$, but they depend on the process $\mathcal{X}^{i,b,k}$. For all $(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]$, let

$$B'_{i,b,k} := \bigcap_{k_1 \in [\![0,\ell_N]\!]} \{\mathcal{X}_N^{i,b,k}(k_1) \le y^* - r(k_1)\} \tag{4.5.15}$$

and

$$C'_{i,b,k} := \{|G_{y^*}^{i,b,k}(2\ell_N)| \le c'N\}, \tag{4.5.16}$$

where c' is as in the definition of $C_{i,b,k}$. The event $B'_{i,b,k}$ means that for all times $k_1 \in [\![0, \ell_N]\!]$, in $\mathcal{X}^{i,b,k}(k_1)$, every particle is to the left of position $y^* - r(k_1)$ at time k_1. The event $C'_{i,b,k}$ says that the number of time-$2\ell_N$ particles in $\mathcal{X}^{i,b,k}$ to the right of position y^* is at most $c'N$. The messages of both events are the same as those of the events $B_{i,b,k}$ and $C_{i,b,k}$ considering that in $\mathcal{X}^{i,b,k}$ we remove the particle that performs the jump $X_{i,b,k}$ and hence that particle has no descendants. In Lemma 4.5.3 we claim that in fact the events $B'_{i,b,k}$ and $C'_{i,b,k}$ imply the events $B_{i,b,k}$ and $C_{i,b,k}$ respectively.

Lemma 4.5.3. *For all $(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]$, $n \in [\![0, 2\ell_N]\!]$ and $z \in \mathbb{R}$, we have*

$$B'_{i,b,k} \subseteq B_{i,b,k}, \quad C'_{i,b,k} \subseteq C_{i,b,k}, \quad \text{and} \quad |G_z^{i,b,k}(n)| \le |\tilde{G}_z^{i,b,k}(n)|.$$

Proof. Fix $(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]$ and assume that $B'_{i,b,k}$ occurs. We will use the notation introduced in (4.5.12). Take $k_1 \in [\![0, \ell_N]\!]$, $i_1 \in [N] \setminus \mathcal{N}(k_1)$, and let $i_2 := \sigma_{k_1}^{-1}(i_1)$. Then $i_2 > D_{k_1}$ by the definition of $\sigma_{k_1}^{-1}$ in (4.5.14). Hence, by Lemma 4.5.1, on the event $B'_{i,b,k}$, we have

$$\mathcal{X}_{i_1}(k_1) \le \mathcal{X}_{i_2}^{i,b,k}(k_1) \le \mathcal{X}_N^{i,b,k}(k_1) \le y^* - r(k_1),$$

which shows that the event $B_{i,b,k}$ must occur.

In order to prove $C'_{i,b,k} \subseteq C_{i,b,k}$, we check that

$$|G_{y^*}(2\ell_N) \setminus \mathcal{N}(2\ell_N)| \le |G_{y^*}^{i,b,k}(2\ell_N)|. \tag{4.5.17}$$

Take $i_1 \in G_{y^*}(2\ell_N) \setminus \mathcal{N}(2\ell_N)$, and let $i_2 = \sigma_{2\ell_N}^{-1}(i_1)$. Then similarly as before, $i_2 > D_{2\ell_N}$. Thus, by Lemma 4.5.1,

$$y^* \le \mathcal{X}_{i_1}(2\ell_N) \le \mathcal{X}_{i_2}^{i,b,k}(2\ell_N),$$

that is, $i_2 \in G_{y^*}^{i,b,k}(2\ell_N)$. Therefore, since $\sigma_{2\ell_N} : [N] \to [N]$ is a bijection, $\sigma_{2\ell_N}^{-1}$ defines an injection from $G_{y^*}(2\ell_N) \setminus \mathcal{N}(2\ell_N)$ to $G_{y^*}^{i,b,k}(2\ell_N)$, which shows (4.5.17), and so $C'_{i,b,k} \subseteq C_{i,b,k}$

as well.

Finally, the first statement of Lemma 4.5.1 immediately implies $|G_z^{i,b,k}(n)| \le |\tilde{G}_z^{i,b,k}(n)|$, so we are done with the proof of Lemma 4.5.3. □

4.5.3 Conclusion of proof of Proposition 4.4.3

Since $\mathcal{X}^{i,b,k}$ is independent of the jump $X_{i,b,k}$, we also have that the events $B'_{i,b,k}$ and $C'_{i,b,k}$ are independent of the event $A_{i,b,k}$. Therefore, by Lemma 4.5.3 and then by independence, we have

$$\begin{aligned}\mathbb{P}\left(A_{i,b,k} \cap B_{i,b,k} \cap C_{i,b,k} \cap D\right) &\ge \mathbb{P}\left(A_{i,b,k} \cap B'_{i,b,k} \cap C'_{i,b,k}\right) - \mathbb{P}\left(A_{i,b,k} \cap D^c\right) \\ &= \mathbb{P}\left(A_{i,b,k}\right) \mathbb{P}\left(B'_{i,b,k} \cap C'_{i,b,k}\right) - \mathbb{P}\left(A_{i,b,k} \cap D^c\right),\end{aligned}$$

for all $(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]$. Then continuing (4.5.11), we obtain

$$\mathbb{P}(E) \ge \sum_{(i,b,k)\in[N]\times\{1,2\}\times[\![0,\ell_N-1]\!]} \left(\mathbb{P}\left(A_{i,b,k}\right) \mathbb{P}\left(B'_{i,b,k} \cap C'_{i,b,k}\right) - \mathbb{P}\left(A_{i,b,k} \cap D^c\right)\right). \quad (4.5.18)$$

Fix $(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]$. We have

$$\mathbb{P}\left(A_{i,b,k} \cap D^c\right) = \mathbb{E}\left[\mathbb{P}\left(A_{i,b,k} \cap D^c \mid \mathcal{F}_{k+1}\right)\right] = \mathbb{E}\left[\mathbb{P}\left(D^c \mid \mathcal{F}_{k+1}\right) \mathbb{1}_{A_{i,b,k}}\right].$$

Recall that on the event $A_{i,b,k}$ we have $\mathcal{X}_N(k+1) > y^* - r(k+1)$ as we explained in (4.5.6). Hence, by Lemma 4.4.1 we have on the event $A_{i,b,k}$,

$$\mathbb{P}\left(D^c \mid \mathcal{F}_{k+1}\right) < e^{-6},$$

and therefore we conclude

$$\mathbb{P}\left(A_{i,b,k} \cap D^c\right) \le e^{-6}\mathbb{P}\left(A_{i,b,k}\right). \quad (4.5.19)$$

We deal with the first term in the sum in (4.5.18) as follows. We have

$$\mathbb{P}\left(B'_{i,b,k} \cap C'_{i,b,k}\right) = \mathbb{P}\left(C'_{i,b,k}\right) - \mathbb{P}\left((B'_{i,b,k})^c \cap C'_{i,b,k}\right). \quad (4.5.20)$$

By the definition of $C'_{i,b,k}$ and by the third statement of Lemma 4.5.3,

$$\mathbb{P}\left((C'_{i,b,k})^c\right) \le \mathbb{P}\left(|\tilde{G}_{y^*}^{i,b,k}(2\ell_N)| > c'N\right) = \mathbb{P}\left(|G_{y^*}(2\ell_N)| > c'N\right) \le \frac{\mathbb{E}\left[|G_{y^*}(2\ell_N)|\right]}{c'N} \le \frac{c}{c'},$$

where the equality holds because $\tilde{\mathcal{X}}^{i,b,s}$ is distributed as $\mathcal{X}$, then we apply Markov's inequality, and the third inequality follows by Proposition 4.4.2. Therefore, we have

$$\mathbb{P}\left(C'_{i,b,k}\right) \ge 1 - \frac{c}{c'}. \quad (4.5.21)$$

Let τ denote the first time k_1 before time ℓ_N when the rightmost particle in $\mathcal{X}^{i,b,k}$ is to the right of $y - r(k_1)$, if there is any such time:

$$\tau := \inf\{k_1 \in [\![0, \ell_N]\!] : \mathcal{X}_N^{i,b,k}(k_1) > y - r(k_1)\},$$

and let $\tau = \ell_N + 1$ if there is no such time up to time ℓ_N. Then we have $B'_{i,b,k} \in \mathcal{F}_\tau$, since for any $s \in [\![0, \ell_N]\!]$, we have $\{\tau \leq s\} \cap B'_{i,b,k} = \emptyset$, and for any $s \geq \ell_N + 1$ we have $\{\tau \leq s\} \cap B'_{i,b,k} = B'_{i,b,k} \in \mathcal{F}_s$. Thus

$$\mathbb{P}\left((B'_{i,b,k})^c \cap C'_{i,b,k}\right) = \mathbb{E}\left[\mathbb{P}\left(C'_{i,b,k} \mid \mathcal{F}_\tau\right) \mathbb{1}_{(B'_{i,b,k})^c}\right]. \tag{4.5.22}$$

On the event $(B'_{i,b,k})^c$ we have $\tau \in [\![0, \ell_N]\!]$ and $\mathcal{X}_N^{i,b,k}(\tau) > y^* - r(\tau)$. Also note that

$$\mathbb{P}\left(C'_{i,b,k} \mid \mathcal{F}_\tau\right) \leq \mathbb{P}\left(|G_{y^*}^{i,b,k}(2\ell_N)| < N \mid \mathcal{F}_\tau\right), \tag{4.5.23}$$

by the definition of the event $C'_{i,b,k}$.

Now Lemma 4.4.1 does not exactly apply, because we are considering the process $\mathcal{X}^{i,b,k}$ rather than $\mathcal{X}$. However, the idea for giving an upper bound on this conditional probability on the event $\{\mathcal{X}_N^{i,b,k}(\tau) > y^* - r(\tau)\}$ remains the same: we count the number of time-$2\ell_N$ descendants of particles which made a jump of size $r(\tau)$ and which are descended from to the right of position $y^* - r(\tau)$. To lower bound the number of particles to the right of $y^* - r(\tau)$ at some time s, one can use Lemma 4.5.2 instead of Lemma 4.1.1. Apart from that, one can repeat the proof of Lemma 4.4.1 to arrive at the following upper bound: on the event $(B'_{i,b,k})^c$,

$$\mathbb{P}\left(|G_{y^*}^{i,b,k}(2\ell_N)| < N \mid \mathcal{F}_\tau\right) < e^{-9/8}. \tag{4.5.24}$$

(We will not give more details on this calculation.) Therefore, by (4.5.22) and (4.5.23),

$$\mathbb{P}\left((B'_{i,b,k})^c \cap C'_{i,b,k}\right) \leq e^{-9/8},$$

and then by (4.5.20) and (4.5.21),

$$\mathbb{P}\left(B'_{i,b,k} \cap C'_{i,b,k}\right) \geq 1 - \tfrac{c}{c'} - e^{-9/8}.$$

Now from (4.5.18) and (4.5.19) we have

$$\begin{aligned}\mathbb{P}(E) &\geq \sum_{(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]} \mathbb{P}(A_{i,b,k})\,(1 - \tfrac{c}{c'} - e^{-9/8} - e^{-6}) \\ &\geq \sum_{(i,b,k) \in [N] \times \{1,2\} \times [\![0, \ell_N - 1]\!]} \mathbb{P}\left(X > y^* - r(k+1) - x_{f(i,k)}\right)/2,\end{aligned}$$

by the definition of $A_{i,b,k}$ and because $1 - \frac{c}{c'} - e^{-9/8} - e^{-6} > 1/2$ by the choices of c and c' in the definition of $C_{i,b,k}$. Then, considering that if $f(i,k) = j$ then $j \leq \ell_N - k$ by (4.5.2), and then using (4.5.3), we have

$$\begin{aligned}\mathbb{P}(E) &\geq \sum_{k=0}^{\ell_N-1}\sum_{j=0}^{\ell_N-k}\sum_{i:\, f(i,k)=j} 2\mathbb{P}\left(X > y^* - r(k+1) - x_j\right)/2\\ &\geq \frac{1}{2}\sum_{k=0}^{\ell_N-1}\sum_{j=0}^{\ell_N-k} 2^{j+k}\mathbb{P}\left(X > y^* - r(k+1) - x_j\right),\end{aligned}$$

which together with (4.5.7) shows (4.5.1). This is what we wanted to prove; we do not include the adjustments we have to make for a precise proof of the first inequality of Proposition 4.4.3. To prove the second inequality, one can give a very similar argument to the one in the present subsection. In order to write $R(k)-r(k)$ instead of $y^*-x_j-r(k+1)$ one can use the property in (4.4.6)

4.6 Ideas for the proof of Proposition 4.4.2 – upper bound on $\mathbb{E}[|G_{y^*}(t)|]$

Recall that for $j \in [\![0, \ell_N]\!]$, we defined the positions x_j as

$$x_j := \mathcal{X}_{N-2^j+1}(0).$$

Now let us define the sets $\mathbf{S}_j$ as follows. For $j \in [\ell_N - 1]$ let

$$\mathbf{S}_j := \{(N-2^{j+1}+1, 0), \ldots, (N-2^j, 0)\},$$

and let

$$\mathbf{S}_0 := \{(N-1, 0), (N, 0)\}.$$

Then there are 2^j particles in each set $\mathbf{S}_j$ for $j \in [\ell_N - 1]$, and 2^{j+1} for $j = 0$. Furthermore, a path starting from the set $\mathbf{S}_j$ needs to move at least distance $y^* - x_j$ to arrive in the set $G_{y^*}(2\ell_N)$.

Recall the definition of the event H_N from (4.3.3). When bounding $\mathbb{E}[|G_{y^*}(t)|]$, by Corollary 4.3.2, we only need to consider the scenario when on every path between time 0 and $2\ell_N$, the sum of jumps which are smaller than $\delta_N L_N$ is at most ρ_N. Then, if a path starts from $\mathbf{S}_j$, then the fairly big jumps (jumps larger than $\delta_N L_N$) on the path must add up to at least $y^* - x_j - \rho_N$ in order for the path to arrive in the set $G_{y^*}(2\ell_N)$. Let B_{δ_N} denote the set of fairly big jumps during the time interval $[0, 2\ell_N]$:

$$B_{\delta_N} := \{(i,b,k) \in [N] \times \{1,2\} \times [\![0, 2\ell_N - 1]\!] :\ X_{i,b,k} > \delta_N L_N\}.$$

We need to handle different paths in different ways depending on the number of fairly big jumps on each path. Let G_k denote the subset of $G_{y^*}(2\ell_N)$ for which there are exactly

k fairly big jumps on the path leading to each particle in G_k from time zero:

$$G_k := G_{y^*}(2\ell_N) \cap \left\{ j \in [N] : \left| P_{i,0}^{j,2\ell_N} \cap B_{\delta_N} \right| = k, \text{ if } (i,0) \lesssim (j, 2\ell_N) \right\}. \tag{4.6.1}$$

It is easy to check that by the definitions of y^* and the function h_1 in (4.4.9) and (4.4.7), that we must have

$$y^* - x_j > L_N \tag{4.6.2}$$

for all $j \in [\![0, \ell_N]\!]$. Therefore, the number of paths leading to $G_{y^*}(2\ell_N)$ without any fairly big jump on the path is zero on the event H_N, and thus

$$\mathbb{E}[|G_0| \mathbb{1}_{H_N}] = 0. \tag{4.6.3}$$

4.6.1 Paths with exactly one fairly big jump

Next, we count the expected number of particles in G_1. We claim that

$$\mathbb{E}[|G_1| \mathbb{1}_{H_N}] \leq 5 \sum_{k=0}^{\ell_N - 1} \sum_{j=0}^{\ell_N - k} 2^{j+k} \mathbb{P}(X > y^* - r(k+1) - x_j - \rho_N) N. \tag{4.6.4}$$

We prove this claim as follows. Let k_1 denote the time of the single fairly big jump which happens on a path between time zero and the set G_1.

At time $k_1 \in [\![0, \ell_N - 1]\!]$ and for all $j \in [\![0, (\ell_N - k_1) \wedge (\ell_N - 1)]\!]$, there are at most 2^{j+k_1+1} (surviving) particles descended from $\mathbf{S}_j$ which have not made a fairly big jump yet (we have the '+1' in the exponent because there are two particles in $\mathbf{S}_0$). The at most 2^{j+k_1+2} children of these particles need to make a fairly big jump of size at least $y^* - x_j - \rho_N$ to get to the right of y^* by time $2\ell_N$, and a particle making such a jump at time k_1 will have at most N descendants at time $2\ell_N$.

At time k_1, for $j = \ell_N - k_1$ there are at most $N = 2^{j+k_1}$ particles descended from the set $\bigcup_{j'=j}^{\ell_N - 1} \mathbf{S}_{j'}$ which have not made a fairly big jump yet. The at most $2N$ children of these particles have to make a fairly big jump of size at least $y^* - x_j - \rho_N$ to get to the right of y^* by time $2\ell_N$, and a particle making such a jump at time k_1 will have at most N descendants at time $2\ell_N$.

At time $k_1 \in [\![\ell_N, 2\ell_N]\!]$, particles that have not made a fairly big jump yet have to make a fairly big jump of size at least $y^* - \mathcal{X}_N(0) - \rho_N$ to get to the right of y^* by time $2\ell_N$. A particle performing such a jump at time k_1 can have at most $2^{2\ell_N - k_1 - 1}$ descendants at time $2\ell_N$.

Therefore,

$$\begin{aligned}
&\mathbb{E}[|G_1|\mathbb{1}_{H_N}] \\
&\leq \sum_{k_1=0}^{\ell_N-1}\sum_{j=0}^{(\ell_N-k_1)\wedge(\ell_N-1)} 2^{j+k_1+2}\mathbb{P}(X > y - x_j - \rho_N)\,N \\
&\quad + \sum_{k_1=\ell_N}^{2\ell_N-1} 2N\mathbb{P}(X > y - \mathcal{X}_N(0) - \rho_N)\,2^{2\ell_N-k_1-1} \\
&\leq \sum_{k_1=0}^{\ell_N-1}\sum_{j=0}^{\ell_N-k_1} 2^{j+k_1+2}\mathbb{P}(X > y - x_j - \rho_N)\,N + \sum_{k_1'=0}^{\ell_N-1} 2N\mathbb{P}(X > y - x_0 - \rho_N)\,2^{k_1'}, \quad (4.6.5)
\end{aligned}$$

where we took $k_1' = 2\ell_N - k_1 - 1$, and the second sum on the right-hand side can be upper bounded by the first sum (even by the $j = 0$ term of the first sum). Since $r(k_1) \geq 0$ for all $k_1 \in [\ell_N]$, we conclude (4.6.4).

4.6.2 Paths with exactly two fairly big jumps

In this section we discuss the ideas to bound $\mathbb{E}[|G_2|\mathbb{1}_{H_N}]$. We distinguish several cases within the set G_2 based on when the two fairly big jumps occur and whether the first or the second jump is larger. We will explain one particular case, and we note that the other cases work similarly.

We say that a particle is in the set $G_{2,1}$, if all of the following hold:

- the particle is in G_2
- the first fairly big jump is larger than the second on the path from time zero to the particle
- the first fairly big jump occurs before time ℓ_N.

Other subsets of G_2 include the cases when the first fairly big jump occurs after time ℓ_N and when the second fairly big jump is larger than the first. These cases can be handled using a similar method to the proof we give for the set $G_{2,1}$. In Proposition 4.6.1 we bound the size of the set $G_{2,1}$.

Proposition 4.6.1. *For all $\varepsilon > 0$, for N sufficiently large,*

$$\mathbb{E}[|G_{2,1}|\mathbb{1}_{H_N}] \leq \varepsilon h_1(y^*).$$

The lemma below is a key step of the proof of Proposition 4.6.1. The lemma gives an upper bound for the probability that the sum of two jumps is larger than $y^* - x_j - \rho_N$, assuming that the first (larger) jump is at most of size $y^* - x_j - r(k_1+1) - \rho_N$. The upper bound is expressed in terms of the probability that the first jump is larger than $y^* - x_j - r(k_1+1) - \rho_N$ and the second is larger than $r(k+1)$.

Lemma 4.6.2. *Take $k_1 \in [\![0, \ell_N - 1]\!]$ and $j \in [\![0, \ell_N - k_1]\!]$, and suppose that $r(k_1+1) \leq (y^* - x_j - \rho_N)/2$. Then for all $\varepsilon > 0$ and for K_1 as in (4.4.7), for N sufficiently large,*

$$\int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} f(z)\mathbb{P}(X > y^* - x_j - \rho_N - z)\,dz$$
$$\leq \varepsilon K_1 (\log N)^{1/2} e^{-r(k_1+1)^\beta - (y^*-x_j-\rho_N-r(k_1+1))^\beta},$$

where the function $f(z)$ is the density given by (4.2.2).

We now prove Proposition 4.6.1 using Lemma 4.6.2, then discuss the lemma. In particular we will comment on the factor of $(\log N)^{1/2}$.

Proof of Proposition 4.6.1. Let $G^a_{2,1} \subseteq G_{2,1}$ denote the set of particles in $G_{2,1}$ for which the first fairly big jump is larger than $y^* - x_j - r(k_1+1) - \rho_N$, if it occurs at time k_1. Then, by the same argument as the one we gave for (4.6.5), we have

$$\mathbb{E}\left[|G^a_{2,1}|\mathbb{1}_{H_N}\right] \leq \sum_{k_1=0}^{\ell_N-1}\sum_{j=0}^{\ell_N-k_1} 2^{j+k_1+2}\mathbb{P}(X > y^* - x_j - r(k_1+1) - \rho_N)\,N. \tag{4.6.6}$$

Next, let $G^b_{2,1} \subseteq G_{2,1}$ denote the set of particles in $G_{2,1}$ for which the first fairly big jump is at most of size $y^* - x_j - r(k_1+1) - \rho_N$, if it occurs at time k_1. Then we claim that we can give the following upper bound on the expected number of particles in $G^b_{2,1}$:

$$\mathbb{E}\left[|G^b_{2,1}|\mathbb{1}_{H_N}\right] \leq \sum_{k_1=0}^{\ell_N-1}\sum_{j=1}^{(\ell_N-k_1)\wedge(\ell_N-1)}\int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} 2^{j+k_1+2} f(z)$$
$$\sum_{k_2=k_1+1}^{2\ell_N-1} 2(2^{k_2-k_1-1}\wedge N)\mathbb{P}(X > y^* - x_j - \rho_N - z)(2^{2\ell_N-k_2-1}\wedge N)dz. \tag{4.6.7}$$

We explain this formula as follows:

- We integrate over the size z of the first fairly big jump. The integral domain follows by the definition of $G_{2,1}$, by the fact that the sum of the two fairly big jumps has to be at least $y^* - x_j - \rho_N$, and by the definition of $G^b_{2,1}$.
- The first fairly big jump can occur at any time $k_1 \in [\![0, \ell_N - 1]\!]$.
- At time $k_1 \in [\![0, \ell_N - 1]\!]$ and for all $j \in [\![0, (\ell_N - k_1) \wedge (\ell_N - 1)]\!]$, there are at most 2^{k_1+j+2} particles descended from the set $\mathbf{S}_j$ which can attempt to make a jump of size z.
- At time $k_1 \in [\![1, \ell_N - 1]\!]$ for $j = \ell_N - k_1$, there are at most $2N = 2^{k_1+j+1}$ particles descended from the set $\bigcup_{j'=j}^{\ell_N-1}\mathbf{S}_{j'}$ which can attempt to make a jump of size z, with density $f(z)$.

- The second fairly big jump can happen at any time $k_2 \in [\![k_1+1, 2\ell_N-1]\!]$.
- At any time k_2 there are at most $2\cdot(2^{k_2-k_1-1}\wedge N)$ jumps descended from any particle that performed a first fairly big jump at time k_1.
- On the event H_N, the size of the second fairly big jump has to be larger than $y^* - x_j - \rho_N - z$, because the two fairly big jumps need to add up to at least $y^* - x_j - \rho_N$ in order for the path to end up in the set G_2.
- Any particle that performed a second jump at time k_2 can have at most $(2^{2\ell_N-k_2-1}\wedge N)$ descendants at time $2\ell_N$.

Now considering only the terms in (4.6.7) which depend on k_2, we have

$$\begin{aligned}
&\sum_{k_2=k_1+1}^{2\ell_N-1}(2^{k_2-k_1-1}\wedge N)(2^{2\ell_N-k_2-1}\wedge N)\\
&=\sum_{k_2=k_1+1}^{\ell_N-1}2^{k_2-k_1-1}N+\mathbb{1}_{\{k_1>0\}}\sum_{k_2=\ell_N}^{k_1+\ell_N}2^{k_2-k_1-1+2\ell_N-k_2-1}+\sum_{k_2=k_1+\ell_N+1}^{2\ell_N-1}N2^{2\ell_N-k_2-1}\\
&\le 2^{2\ell_N-k_1-1}+(k_1+1)2^{2\ell_N-k_1-2}+2^{2\ell_N-k_1-1}.
\end{aligned}$$

Therefore, by (4.6.7) and then by Lemma 4.6.2, we have

$$\begin{aligned}
&\mathbb{E}\left[|G_{2,1}^b|\mathbb{1}_{H_N}\right]\\
&\le\sum_{k_1=0}^{\ell_N-1}\sum_{j=1}^{\ell_N-k_1}\int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N}2(k_1+2)2^{2\ell_N+j}f(z)\mathbb{P}\left(X>y^*-x_j-\rho_N-z\right)dz.\\
&\le\varepsilon K_1(\log N)^{1/2}\\
&\qquad\sum_{k_1=0}^{\ell_N-1}\sum_{j=1}^{\ell_N-k_1}2(k_1+2)2^{2\ell_N+j}\mathbb{P}\left(X>r(k+1)\right)\mathbb{P}\left(X>y^*-x_j-\rho_N-r(k_1+1)\right).
\end{aligned}\tag{4.6.8}$$

Now by the definition of $r(k)$ we have $(k+1)2^{2\ell_N-k_1}\mathbb{P}(X>r(k+1))=32N$, and using that $k_1+2\le 2(k_1+1)$ for all $k_1\in[\![0,\ell_N-1]\!]$, we obtain

$$\mathbb{E}\left[|G_{2,1}^b|\mathbb{1}_{H_N}\right]\le 128\varepsilon K_1(\log N)^{1/2}N\sum_{k_1=0}^{\ell_N-1}\sum_{j=1}^{\ell_N-k_1}2^{j+k_1}\mathbb{P}\left(X>y^*-x_j-\rho_N-r(k_1+1)\right),$$

which, by the definition of h_1 in (4.4.7), concludes the proof of Proposition 4.6.1. □

Discussion about Lemma 4.6.2: Take $k_1\in[\![0,\ell_N-1]\!]$ and $j\in[\![0,\ell_N-k_1]\!]$ and suppose that $r(k_1+1)\le(y^*-x_j-\rho_N)/2$. Consider the integral in Lemma 4.6.2. Substituting the density from (4.2.2) and the tail probability from (4.2.1), we get

$$\begin{aligned}
&\int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} f(z)\mathbb{P}\left(X > y^* - x_j - \rho_N - z\right) dz \\
&= \int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} \beta z^{\beta-1} e^{-z^\beta-(y^*-x_j-\rho_N-z)^\beta} dz \\
&\leq \int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} K'(\log N)^{1-1/\beta} e^{-z^\beta-(y^*-x_j-\rho_N-z)^\beta} dz,
\end{aligned} \tag{4.6.9}$$

for some constant $K' > 1$, where in the inequality we used that $z^{\beta-1} \leq \left(\frac{y^*-x_j-\rho_N}{2}\right)^{\beta-1} \leq K'(\log N)^{1/\beta-1}$ by (4.6.2), (4.3.2) and (4.2.3). Now we claim that for any $\varepsilon > 0$, for N sufficiently large, we have

$$\begin{aligned}
&\int_{\frac{y^*-x_j-\rho_N}{2}}^{y^*-x_j-r(k_1+1)-\rho_N} e^{-z^\beta-(y^*-x_j-\rho_N-z)^\beta} dz \\
&\leq \varepsilon \frac{K_1}{K'} (\log N)^{1/\beta-1/2} e^{-r(k_1+1)^\beta-(y^*-x_j-\rho_N-r(k_1+1))^\beta}.
\end{aligned} \tag{4.6.10}$$

Then (4.6.9) and the claim imply Lemma 4.6.2.

To prove the claim one needs to give a first order estimate on $z^\beta - (y^* - x_j - r(k_1+1) - \rho_N)^\beta$ and on $(y^* - x_j - \rho_N - z)^\beta - r(k_1+1)^\beta$, and then bound the difference between these estimates. Then it is possible to prove that the integral of the exponential of this bound is at most of order $(\log N)^{1/\beta-1/2}$. We will not write out this calculation.

Our original idea was, that if the largest fairly big jump is restricted to be at most $y^* - x_j - \rho_N - r(k+1)$, then the number of paths with this restriction will contribute to $\mathbb{E}[|G_{2,1}|\mathbb{1}_{H_N}]$ by roughly the same order as the number of paths with jumps larger than this size; i.e. we wanted to see $\mathbb{E}\left[|G_{2,1}^a|\mathbb{1}_{H_N}\right] \approx \mathbb{E}\left[|G_{2,1}^b|\mathbb{1}_{H_N}\right]$. This has not turned out to be the case because of the $(\log N)^{1/2}$ factor in Lemma 4.6.2. To get rid of this factor, we would have needed a $(\log N)^{1/\beta-1}$ factor on the right-hand side of (4.6.10) rather than a $(\log N)^{1/\beta-1/2}$ factor. With the calculation below we would like to show that we cannot give a better upper bound on the integral in question: it is possible that there exist j and k_1 such that the left-hand side of (4.6.10) is larger than a constant times the right-hand side of (4.6.10).

Let $r := r(k_1+1)$ and $A := (y^* - x_j - \rho_N)/2$. We assume that

(a) $A - r = c_1(\log N)^{1/\beta-1/2}$ for some constant $c_1 > 0$, and

(b) $A = c_2 L_N$ for some constant $c_2 > 0$.

We will show that with these assumptions, there exists $a > 0$ such that

$$\int_A^{2A-r} e^{-z^\beta-(2A-z)^\beta} dz \geq a(\log N)^{1/\beta-1/2} e^{-r^\beta-(2A-r)^\beta}. \tag{4.6.11}$$

Note that for any $z \in [r, A]$, we have

$$z^\beta \leq r^\beta + \beta(z-r)r^{\beta-1}$$

and

$$(2A-z)^\beta \leq (2A-r)^\beta + \beta(r-z)(2A-r)^{\beta-1}.$$

Therefore, with a change of variables, we have

$$\begin{aligned}\int_A^{2A-r} e^{-z^\beta-(2A-z)^\beta} dz &= \int_r^A e^{-z^\beta-(2A-z)^\beta} dz \\ &\geq e^{-r^\beta-(2A-r)^\beta} \int_r^A e^{-\beta(z-r)(r^{\beta-1}-(2A-r)^{\beta-1})} dz. \qquad (4.6.12)\end{aligned}$$

Now in the exponent of the integrand we have

$$r^{\beta-1}-(2A-r)^{\beta-1} = -\int_r^{2A-r} (\beta-1)x^{\beta-2}dx \leq (1-\beta)(2A-2r)r^{\beta-2} \leq c_3(\log N)^{1/2-1/\beta},$$

for some $c_3 > 0$, where the second inequality follows by the assumptions in (a) and (b). Thus,

$$\int_r^A e^{-\beta(z-r)(r^{\beta-1}-(2A-r)^{\beta-1})} dz \geq \int_r^A e^{-\beta c_3(\log N)^{1/2-1/\beta}(z-r)} dz \geq c_4(\log N)^{1/\beta-1/2},$$

for some $c_4 > 0$, by the assumption in (a). Therefore, with the assumptions in (a) and (b), (4.6.11) holds by (4.6.12), which means that under these assumptions we cannot improve (4.6.10).

The assumptions in (a) and (b) are possibilities that can indeed occur. For example, if $y^* - x_j \approx L_N + r(\ell_N - j)$ for $j = (1-\varepsilon)^\beta \ell_N$ with some small $\varepsilon > 0$, then we have $y^* - x_j \approx (2-\varepsilon)L_N$ (see (4.4.5)). Now it can be checked that there exists $k < \ell_N - j$ such that $(1-\varepsilon/2)L_N - r(k)$ is of order $(\log N)^{1/\beta-1/2}$.

4.6.3 Paths with at least three fairly big jumps

We would like to show that for any $\varepsilon > 0$,

$$\sum_{i=3}^{2\ell_N-1} \mathbb{E}\left[|G_i| \mathbb{1}_{H_N}\right] \leq \varepsilon h(y^*). \qquad (4.6.13)$$

Intuitively, one can think of the following. Consider Lemma 4.3.3 with $x = y^* - x_j$ for some $j \in [\![0, \ell_N - 1]\!]$, $M = L_N^+ = (\log N + C_1 \log\log N)^{1/\beta}$ for some $C_1 > 1$, $d = \delta_N L_N$ (see (4.3.1)), and $m = 3$; and assume $x - M > CL_N$ for some $C \in (0,1)$. Now it is not hard to see that the probability we get from the second part of Lemma 4.3.3 is much smaller than in the case when we have the same assumptions except that we change m to $m = 2$.

Hence, we expect that 'it is not worth' making three fairly big jumps.

Indeed, with the assumptions above, it is enough to apply Lemma 4.3.3 for all the $2^{j+2\ell_N}$ paths in N independent BRWs without selection, to conclude that the number of paths satisfying these assumptions is less than $\varepsilon h(y^*)$ for any $\varepsilon > 0$.

However, for small values of j, it is not necessarily true that $\frac{y^*-x_j-L_N^+}{L_N} > C > 0$. It turns out that when $j < \varepsilon\ell_N$ for some $\varepsilon > 0$, we need to include selection in our argument similarly as we did in the case with two fairly big jumps in (4.6.7). This idea leads to several different cases based on when and how many big jumps happen on the paths. Thus the proof becomes pretty long, but eventually it shows that (4.6.13) holds.

4.7 Conclusion

In the previous section we showed how we go about proving Proposition 4.4.2. Using that result we showed Proposition 4.4.3 in Section 4.5 from which we conclude Theorem 4.2.1 as we described in Section 4.4.2.

In the future we would like to finish the full write-up of this proof, and we would also like to prove an upper bound for the probability of a similar event as the one in Theorem 4.2.1 but with β close to 1. An upper bound of smaller order than $(\log N)^{-1/2}$ would show that there is a change of behaviour in the genealogy of the N-BRW in the stretched exponential case as we change the value of β.

Bibliography

[1] K. B. Athreya and P. E. Ney. *Branching processes.* Courier Corporation, 2004.

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

[3] J. Bérard and P. Maillard. The limiting process of N-particle branching random walk with polynomial tails. *Electron. J. Probab.*, 19, 2014.

[4] J. Berestycki, N. Berestycki, and J. Schweinsberg. The genealogy of branching Brownian motion with absorption. *Ann. Probab.*, 41(2):527–618, 2013.

[5] J. Berestycki, É. Brunet, and S. Penington. Global existence for a free boundary problem of Fisher–KPP type. *Nonlinearity*, 32(10):3912, 2019.

[6] N. Berestycki. Recent progress in coalescent theory. *Ensaios Matematicos*, 16, 09 2009.

[7] I.-J. Bienaymé. *De la loi de multiplication et de la durée des familles: probabilités.* Cosson, 1845.

[8] J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Applied Probability*, 8(3):446–459, 1976.

[9] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

[10] E. Bolthausen and A. S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Communications in Mathematical Physics*, 197(2):247–276, 1998. The original publication is available at www.springerlink.com.

[11] É. Brunet, B. Derrida, A. H. Mueller, and S. Munier. Noisy traveling waves: Effect of selection on genealogies. *Europhys. Lett.*, 76(1):1–7, 2006.

[12] É. Brunet, B. Derrida, A. H. Mueller, and S. Munier. Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Phys. Rev. E*, 76:041104, 2007.

[13] É. Brunet. *Some aspects of the Fisher-KPP equation and the branching Brownian motion.* habilitation, UPMC, 2016.

[14] É. Brunet and B. Derrida. Shift in the velocity of a front due to a cutoff. *Phys. Rev. E*, 56:2597–2604, 1997.

[15] É. Brunet and B. Derrida. Microscopic models of traveling wave equations. *Computer Physics Communications*, 121:376–381, 2000.

[16] A. Cortines and B. Mallein. A N-branching random walk with random selection. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14:117–137, 2017.

[17] A. De Masi, P. A. Ferrari, E. Presutti, and N. Soprano-Loto. Hydrodynamics of the N-BBM process. In *International workshop on Stochastic Dynamics out of Equilibrium*, pages 523–549. Springer, 2017.

[18] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009.

[19] D. Denisov, A. B. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.*, 36(5):1946–1991, 2008.

[20] F. Dumortier, N. Popović, and T. J. Kaper. The critical wave speed for the Fisher–Kolmogorov–Petrowskii–Piscounov equation with cut-off. *Nonlinearity*, 20(4):855, 2007.

[21] R. Durrett. Maxima of branching random walks. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 62:165–170, 1983.

[22] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4 edition, 2010.

[23] R. Durrett and D. Remenik. Brunet–Derrida particle systems, free boundary problems and Wiener–Hopf equations. *Ann. Probab.*, 39(6):2043–2078, 2011.

[24] P. Dyszewski, N. Gantert, and T. Höfelsauer. Large deviations for the maximum of a branching random walk with stretched exponential tails. *Electronic Communications in Probability*, 25:1–13, 2020.

[25] P. Dyszewski, N. Gantert, and T. Höfelsauer. The maximum of a branching random walk with stretched exponential tails. *arXiv preprint arXiv:2004.03871*, 2020.

[26] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, 1937.

[27] N. Gantert. The maximum of a branching random walk with semiexponential increments. *Ann. Probab.*, 28(3):1219–1229, 2000.

[28] N. Gantert, Y. Hu, and Z. Shi. Asymptotics for the survival probability in a killed branching random walk. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(1):111–129, 02 2011.

[29] J. M. Hammersley. Postulates for subadditive processes. *Ann. Probab.*, 2(4):652–680, 1974.

[30] T. E. Harris et al. *The theory of branching processes*, volume 6. Springer Berlin, 1963.

[31] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probab.*, 3(5):790–801, 1975.

[32] J. F. Kingman. On the genealogy of large populations. *Journal of applied probability*, 19(A):27–43, 1982.

[33] A. Kolmogorov, I. Petrovskii, and N. Piscounov. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problem biologique. *Mosc. Univ. Bull. Math*, 1:1–25, 1937.

[34] P. Maillard. Speed and fluctuations of N-particle branching Brownian motion with spatial selection. *Probability Theory and Related Fields*, 166(3-4):1061–1173, 2016.

[35] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pages 195–248. Springer, 1998.

[36] H. P. McKean. Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. *Communications on pure and applied mathematics*, 28(3):323–331, 1975.

[37] P. A. P. Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.

[38] C. Mueller, L. Mytnik, and J. Quastel. Small noise asymptotics of traveling waves. *Markov Processes and Related Fields*, 3, 01 2008.

[39] M. Pain. Velocity of the L-branching brownian motion. *Electronic Journal of Probability*, 21:1–28, 2016.

[40] S. Penington, M. I. Roberts, and Z. Talyigás. Genealogy and spatial distribution of the N-particle branching random walk with polynomial tails. *Electronic Journal of Probability*, 27:1–65, 2022.

[41] J. Pitman. Coalescents with multiple collisions. *The Annals of Probability*, 27(4):1870–1902, 1999.

[42] J. Schweinsberg. Rigorous results for a population model with selection II: genealogy of the population. *Electronic Journal of Probability*, 22:1–54, 2017.

[43] Z. Shi. *Branching random walks.* Springer, 2015.

[44] A. V. Skorokhod. Branching diffusion processes. *Theory of Probability & Its Applications*, 9(3):445–449, 1964.

[45] W. Whitt. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues.* Springer Series in Operations Research and Financial Engineering. Springer New York, New York, NY, 2002.