Studies in One Dimensional Branching Random Walks

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ming Fang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Ofer Zeitouni

December, 2011

Acknowledgements

I am indebted to my advisor, Dr. Ofer Zeitouni, for his guidance, encouragement and support. Especially, he introduced to me the model of branching random walks and offered a lot of enlightening discussion on different topics in this thesis; under his encouragement and support, I went to many probability conferences which boosted my research and broadened my horizon; he generously supported my research with his funding NSF grant DMS-0804133.

I am grateful to my committee chair, Dr. Maury Bramson. He provided related references to several projects in this thesis; he proofread this thesis very carefully and explained in detail how to improve the mathematical writing.

I would also like to thank the other two committee members: Dr. John Baxter and Dr. Bert Fristedt. Thank you for taking time out of your busy schedule to help me with the thesis.

My gratitude also extends, but does not limit, to Dr. Nicolai V. Krylov, Dr. Arnd Scheel, Dr. Yueyun Hu, Dr. Jonathon Peterson, Xiaoqin Guo, Teng Wang and Xu Li. They all helped me in various aspects during my graduate study.

Dedication

To my father, Meishan Fang, and my mother, Ying Peng, who unconditionally help and support my growth in all aspects to the best of their ability.

To my wife, Dr. Lingyi Dong, who has been and will always be with me together through thick and thin.

To all my relatives and friends, with whom I have shared happiness and sorrow.

Abstract

This thesis deals with three problems arising from branching random walks.

The first problem studies the leftmost path (compared with the leftmost particle) of branching random walks. Let $\mathbb{T}$ denote a rooted b-ary tree and let $\{S_v\}_{v\in\mathbb{T}}$ denote a branching random walk indexed by the vertices of the tree, where the increments are i.i.d. and possess a logarithmic moment generating function $\Lambda(\cdot)$. Let m_n denote the minimum of the variables S_v over all vertices at the nth generation, denoted by $\mathbb{D}_n$. Under mild conditions, m_n/n converges almost surely to a constant, which for convenience may be taken to be 0. With $\bar{S}_v = \max\{S_w : w \text{ is on the geodesic connecting the root to } v\}$, define $L_n = \min_{v\in\mathbb{D}_n} \bar{S}_v$. We prove that $L_n/n^{1/3}$ converges almost surely to an explicit constant l_0.

The second problem studies the tightness of maxima (the displacement of the rightmost particle) of generalized branching random walks on the real line $\mathbb{R}$ that allow time dependence and local dependence between siblings. At time n, $F_n(\cdot)$ is used to denote the distribution function of the maximum. Under appropriate tail assumptions on the branching laws and offspring displacement distributions, we prove that $F_n(\cdot - Med(F_n))$ is tight in n. The main part of the argument is to demonstrate the exponential decay of the right tail $1 - F_n(\cdot - Med(F_n))$.

The third problem studies the maximum of branching random walks in a class of time inhomogeneous environments. Specifically, binary branching random walks with Gaussian increments will be considered, where the variances of the increments change over time macroscopically. We find the asymptotics of the maximum up to an $O_P(1)$ (stochastically bounded) error, and focus on the following phenomena: the profile of the variance matters, both to the leading (velocity) term and to the logarithmic correction term, and the latter exhibits a phase transition.

Contents

Chapter 1

Introduction

This thesis studies some properties of branching random walks (BRWs) in dimension 1. In Chapter 1, we introduce the model of branching random walks and other related models, among which are branching Brownian motion (BBM), independent random walks and random walks in random environments on trees. These different models are introduced for the following reasons. The branching Brownian motion is a continuous time version of branching random walk, so many methods to study these two models can be borrowed from each other. Independent random walks can serve as a starting point in understanding branching random walks and as a comparison benchmark. Specifically, understanding the difference between BRWs and the model of independent random walks is the key in studying the consistent minimum, see Chapter 2, and the maximum of BRW in time inhomogeneous environment, see Chapter 4. The connection between branching random walks and random walks in random environments on trees is less obvious, and it is part of the background of the topic in Chapter 2.

Also in Chapter 1, we review some known results on the asymptotics and tightness of the maximal displacement. A short and intuitive discussion is presented. We refer the rigorous proofs to some existing papers and to Chapters 3 and 4. The methods discussed in Chapter 1 will be modified and used in the following chapters.

In Chapter 2, we look at the consistent minima of branching random walks. Again we start from independent random walks and a different result for independent random walks. Understanding the difference leads us to the desired result on BRWs. Moment methods and precise large deviation estimates are the main tools in this chapter.

In Chapter 3, we switch to an analytic argument in order to study the tightness of the maxima of a generalized class of branching random walks, generalizing some results from [16]. By deriving a recursion from the tree structure and using a Lyapunov function, we are able to prove the tightness of the maxima of branching random walks, even when the environments are inhomogeneous with respect to time and particles have interactions with their siblings.

In Chapter 4, we consider some simple models of branching random walks with two different laws for the increments. With the help of the tightness result in Chapter 3, we give an aymptotic expansion of the maxima, which is similar to but different from the known result for the homogenous case reviewed in Chapter 1. More interesting is that the results change when we switch the order of the laws of the increments. These simple models offer insights into the more general time inhomogeneous branching random walks.

1.1 Branching Random Walks and Related Models

As the name suggests, a BRW can be viewed as a system of particles performing random walks while branching. We assume, in Chapters 1 and 2, that $b \geq 2$ is a deterministic integer and that $G(\cdot)$ is a distribution function on $\mathbb{R}$.[1]

Definition 1. *A one dimensional* ***Branching random walk*** *can be described as follows. At time 0, a particle is located at 0. Suppose that, at time* n*, a particle* v *is at location* S_v*. At time* $n+1$*,* v *dies and gives birth to* b *offspring. The* b *offspring then move to their new locations* $\{S_v+Y_{v,1},\ldots,S_v+Y_{v,b}\}$*, respectively, with increments independently distributed according to* $G(\cdot)$*. Let* $\mathbb{D}_n$ *be the collection of all the particles at time* n *and* $\mathbb{D}=\cup_n\mathbb{D}_n$ *the collection of all the particles at all time. Then,* $\{S_v\}_{v\in\mathbb{D}}$ *forms a branching random walk indexed by the particles.*

Note that the collection of random variables $\{Y_{v,i}\}_{v\in\mathbb{D},1\leq i\leq b}$ are i.i.d. with a common distribution $G(\cdot)$. This model and the results described below easily generalize to the Galton-Watson setup, that is, we allow a random branching in the model (b is a random variable with some appropriate distribution on $\mathbb{Z}_+$.) For simplicity, we assume b to be

[1] The model under this assumption only involves time-homogeneous randomness $G(\cdot)$ from the random walks. Branching random walks under different settings will be studied in Chapters 3 and 4.

deterministic in Chapter 1, 2 and 4. The results reviewed in Chapter 1 were known to hold for random b, and the results in Chapter 2 and 4 are believed to be true for random b but haven't been checked by the author.

There is another equivalent description of branching random walks via trees.

Definition 1'. *A one dimensional **Branching random walk** can be constructed in two steps. First, we construct a weighted tree. In particular, starting from a rooted b-ary tree $\mathbb{T} = (V, E)$, where V is the set of vertices and E is the set of edges, we assign i.i.d. G-distributed random variables X_e to each edge $e \in E$ in the tree. Second, for each vertex $v \in V$, we assign a random variable S_v by summing the random variables on the edges along the geodesic from the root to v. Then $\{S_v\}_{v\in V}$ forms a branching random walk.*

Remark $\{X_e\}_{e\in E}$ in Definition 1' is the same set as $\{Y_{v,i}\}_{v\in\mathbb{D},1\leq i\leq b}$ in Definition 1 with different labeling. X_e is understood as the weight of e, which accrues additively, and S_v is the accumulated weight of the path from the root to v. If we interpret X_e as the displacement between the two vertices connected by e, then this fits Definition 1 of branching random walks with S_v denoting the displacement of v with respect to the root. If we interpret X_e as the time needed to pass the edge e, then S_v is the percolation time to v from the root.

We introduce some more notation for later use. Let $v \in \mathbb{D}_n$ denote a particle at time n. v^k denotes the ancestor of v at time k for $0 \leq k \leq n$, and $v < u$ means that u is a descendent of v. For any $u, v \in \mathbb{D}$, $u \wedge v$ denotes the largest common ancestor of two particles u and v, and uv denotes the edge connecting u and v if u and v are adjacent (i.e., there is an edge connecting u and v).

For any fixed $v \in \mathbb{D}_n$, $S_v = \sum_{k=0}^{n-1} X_{v^k v^{k+1}}$ is a random walk with increments distributed according to $G(\cdot)$. There are b^n particles at time n and therefore, for each n, $\{S_v\}_{v\in\mathbb{D}_n}$ is a collection of b^n random walks with the same distribution. However, the b^n random walks are not independent, which makes the questions on branching random walks less trivial and more interesting. In fact, for any particles u and v, S_u and S_v both depend on $S_{u\wedge v}$, the position of their largest common ancestor. Note that we can get three independent copies of random walks from this description: $S_{u\wedge v}$, $S_u - S_{u\wedge v}$ and $S_v - S_{u\wedge v}$. This kind of independence makes it possible for us to carry out a probabilistic analysis later in the thesis.

It is natural to introduce a related model called independent random walks, where the number of random walks grows exponentially with respect to the number of steps. Specifically, recall that $b \geq 2$ is a deterministic integer and that $G(\cdot)$ is a distribution function on $\mathbb{R}$.

Definition 2. *The model of **independent random walks** of length n is a collection of b^n independent random walks of n steps $\{S_n^k\}_{k=1}^{b^n}$:*

$$S_n^k = X_1^k + X_2^k + \cdots + X_n^k, \qquad k = 1, 2, \ldots, b^n, \tag{1.1}$$

where all the random variables $\{X_i^k\}_{i=1,\ldots,n;\, k=1,\ldots,b^n}$ are i.i.d. with distribution $G(\cdot)$.

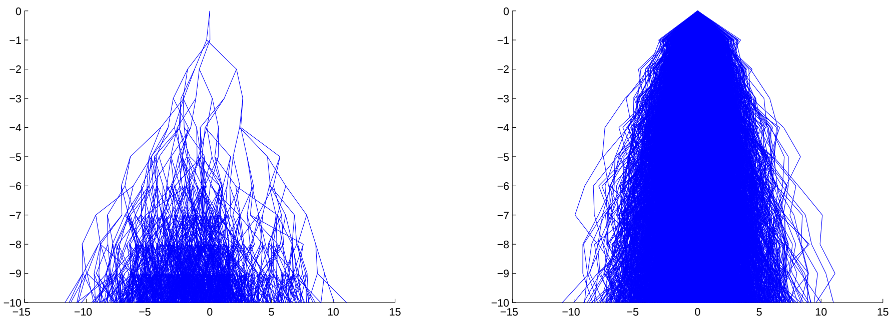
The number of independent random walks matches the number of particles of a branching random walk at time n. But the model of independent random walks assumes complete spatial independence. Not surprisingly, the independence makes many results for independent random walks trivial, while the analysis of branching random walks will be more involved, as the dependence between particles of branching random walks puts additional constraints on the behavior of the walks. In order to prove results on branching random walks, one can sometimes first guess (often hand-waving) analogs for independent random walks, and then incorporate the difference between branching random walks and independent random walks. See Figure 1.1 for a simulation of branching random walks and independent random walks, and note their obviously different behaviors.

We conclude this section with two more related models. The first is branching Brownian motion, a continuous time analog of the branching random walk. [1]

Definition 3. ***Branching Brownian motion*** *is a system of particles performing Brownian motion while branching. The system starts from one particle at location 0 at time 0. Life lengths for particles are independent exponential(1) random variables. When a particle dies, it gives birth to b offspring. The offspring then follow b independent Brownian paths starting from the position of their parent.*

Sometimes there is great similarity between properties of branching random walks and branching brownian motion. Methods can be borrowed from each other. See, for

[1] Other continuous models are also studied, for example, for the branching OU process, see Engländer [27], etc.



(a) Binary Branching Gaussian Random Walk

(b) 2^{10} Independent Gaussian Random Walks

Figure 1.1: Simulations of Branching Random Walks and Independent random Walks.

example, Bramson [14] and Addario-Berry and Reed [1]. But it is also worthwhile to point out that some branching random walks do not always behave in a similar way as the branching Brownian motion, even if random walks converge to Brownian motion according to Donsker's invariance principle. This is beyond the scope of this thesis, but one can find more discussions in Bramson [15].

The other model we introduce is random walks in random environments on trees, which can be defined as follows.

Definition 4. *Given a b-ary tree $\mathbb{T} = (V, E)$ and i.i.d. random weights $\omega = \{p_e\}_{e\in E}$ associated to each edge $e \in E$, a **random walk in random environment** on $\mathbb{T}$ is a time-homogeneous Markovian chain Y_n taking values in V with transition probabilities*

$$P^\omega(S_{n+1} = v | S_n = u) = \begin{cases} \frac{p_{uv}}{\sum_{\overleftarrow{w}=u} p_{uw}+1}, & \text{if } \overleftarrow{v} = u, \\ \frac{1}{\sum_{\overleftarrow{w}=u} p_{uw}+1}, & \text{if } \overleftarrow{u} = v, \\ 0, & \text{otherwise}, \end{cases}$$

for each fixed choice of ω. Here, for any $v \in V$, $\overleftarrow{v}$ denotes the parent of the node v.

Definition 1' of branching random walks and Definition 4 of random walks in random environments on trees both involve a tree with random weights associated to each edge. However, the weights in Definition 1' accrue additively. The weights in Definition

4 roughly represent the probabilities of a walk following a particular edge, and thus accrue multiplicatively if we calculate the probability of a walk following a path. Of course, applying a logarithmic transformation, multiplication can easily be turned into addition. This heuristic builds a non-rigorous connection between branching random walks and random walks in random environments on a tree. This connection is part of the background of Chapter 2. For more rigorous statements and proofs, we refer to Hu and Shi [47].

1.2 Review of Known Results on Branching Random Walks

In the model of branching random walks, see Definition 1, one interesting problem is about the maximum (M_n) and minimum (m_n) of the displacement of particles at time n, i.e., the displacement of the rightmost and leftmost particles at time n:[3]

$$\mathcal{M}_n = \max_{v \in \mathbb{D}_n} S_v, \tag{1.2}$$

and

$$m_n = \min_{v \in \mathbb{D}_n} S_v. \tag{1.3}$$

The asymptotic behavior of the maximum M_n depends on the distribution of the increments, $G(\cdot)$. In this section, instead of quoting all different known results on M_n, we focus on the following special one, through which we illustrate intuition and methods relevant to this thesis. In what follows, for a random variable X, we write $Med(X) = \sup\{x : P(X \leq x) \leq \frac{1}{2}\}$ for the median of X. For simplicity, we deal with the Gaussian case.

Theorem 1. *In Definition 1, let $b = 2$ and $G(\cdot)$ be a standard normal distribution. Then the maximum M_n satisfies the following.*

(i) The sequence $\{M_n - Med(M_n)\}_n$ is tight;

(ii) $Med(M_n) = \sqrt{2\log 2}n - \frac{3}{2\sqrt{2\log 2}} \log n + O(1)$.

Hereinafter $O(1)$ is a generic bounded number, i.e., there exists an $R \in \mathbb{R}$ such that $O(1) \in [-R, R]$.

[3] The properties of maxima and minima are easily translated to each other by multiplying all the increments by -1.

This model is exactly the discrete time analog of branching Brownian motion. The first result of this kind was shown for branching Brownian motion in Bramson [14]. Later, Addario-Berry and Reed [1] gave a more general condition on branching random walks for such results to hold. In the rest of this section, we explain the heuristics for the above theorem by comparing branching random walks with independent random walks, see Definition 2. We also present an approach using a recursion to prove tightness, which will be generalized in Chapter 3.

1.2.1 Moment Methods

We postpone the argument for part (i) in Theorem 1 to the next subsection. In this subsection, we will mainly explain the reason why one can expect part (ii) in Theorem 1 to be true (assuming part (i)). Let us begin with the following theorem for independent random walks.

Theorem 1'. *For the model of independent random walks in Definition 2, let $b = 2$ and $G(\cdot)$ be a standard normal distribution. Then the maximum $M'_n = \max_{k=1}^{2^n} S_n^k$ satisfies*

(i) The sequence $\{M'_n - Med(M'_n)\}_n$ is tight;

(ii) $Med(M'_n) = \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}} \log n + O(1)$.

The proof of Theorem 1' is easy because of the spatial independence of the model. By Definition 2, $\{S_n^k\}_{k=1}^{2^n}$ is a collection of independent normal random variables with mean zero and variance n. Thus,

$$P(M'_n \le x) = \prod_{k=1}^{2^n} P(S_n^k \le x) = \left(1 - \frac{1}{\sqrt{2\pi n}} \int_x^\infty e^{-\frac{y^2}{2n}} dy\right)^{2^n}.$$

Using the fact that $(\frac{1}{x} - \frac{1}{x^3})e^{-x^2/2} \le \int_x^\infty e^{-\frac{y^2}{2}} dy \le \frac{1}{x} e^{-x^2/2}$ for $x > 0$, one can bound $P(M'_n \le x)$ for large x, say $x > 100\sqrt{n}$, as

$$\exp(-c_1 e^{-\frac{x^2}{2n} - \log\frac{x}{\sqrt{n}} + (\log 2)n}) \le P(M'_n \le x) \le \exp(-c_2 e^{-\frac{x^2}{2n} - \log\frac{x}{\sqrt{n}} + (\log 2)n}),$$

where c_1 and c_2 are positive constants. If $x = \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}} \log n + y$, then

$$-\frac{x^2}{2n} - \log\frac{x}{\sqrt{n}} + (\log 2)n = -\left(\sqrt{2\log 2}\right) y + f(n, y),$$

where, for each fixed y, $f(n,y) \to -\log\sqrt{2\log 2}$ as $n \to \infty$. This indicates that the sequence $\{M'_n - Med(M'_n)\}_n$ is tight and

$$Med(M'_n) = \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + O(1).$$

This proof does not generalize to a proof of Theorem 1 due to the dependence between the walks in the BRW model. Therefore, we next describe, in the independent random walks setup, an alternative proof based on moment methods, which does extend to branching random walks. Namely, let

$$N^{ind}_{n,y} = \sum_{k=1}^{2^n} 1_{\{S^k_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y\}} \tag{1.4}$$

be the number of random walks whose displacement are larger than $\sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y$ at time n. Then, the first moment of $N^{ind}_{n,y}$ is

$$\begin{aligned} EN^{ind}_{n,y} &= \sum_{k=1}^{2^n} P(S^k_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y) \\ &= 2^n P(S^1_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y), \end{aligned}$$

from which we can deduce, using standard asymptotics of the normal distribution, that there exist constants c_3, c_4 and c_5 such that

$$c_3 e^{-c_4 y} \leq EN^{ind}_{n,y} \leq c_5 e^{-c_4 y}. \tag{1.5}$$

Using the independence of S^k_n and S^j_n for $k \neq j$, one can compute the second moment of $N^{ind}_{n,y}$ as follows

$$\begin{aligned} E(N^{ind}_{n,y})^2 &= E\sum_{k,j=1}^{2^n} 1_{\{S^k_n, S^j_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y\}} \\ &= \sum_{k\neq j=1}^{2^n} P(S^k_n, S^j_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y) \\ &\quad + \sum_{k=1}^{2^n} P(S^k_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y) \\ &= 2^n(2^n - 1)\left(P(S^1_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y)\right)^2 + EN^{ind}_{n,y} \\ &= (1 - 2^{-n})(EN^{ind}_{n,y})^2 + EN^{ind}_{n,y}. \end{aligned} \tag{1.6}$$

From (1.5), a first moment method (Chebyshev's inequality) implies the upper bound, i.e.,

$$P(M'_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y) = P(N^{ind}_{n,y} \geq 1) \leq EN^{ind}_{n,y} \leq c_5 e^{-c_4 y},$$

which can be made as small as we wish by choosing y large enough. From (1.5) and (1.6), a second moment method implies the lower bound, i.e.,

$$\begin{aligned} & P(M'_n > \sqrt{2\log 2}n - \frac{1}{2\sqrt{2\log 2}}\log n + y) = P(N^{ind}_{n,y} > 0) \\ \geq \quad & \frac{(EN^{ind}_{n,y})^2}{E(N^{ind}_{n,y})^2} = \frac{(EN^{ind}_{n,y})^2}{(1-2^{-n})(EN^{ind}_{n,y})^2 + EN^{ind}_{n,y}}, \end{aligned}$$

which can be made as close to 1 as possible by choosing n large and y very negative (thus $EN^{ind}_{n,y}$ is very large). This completes the proof of Theorem 1' using moment methods.

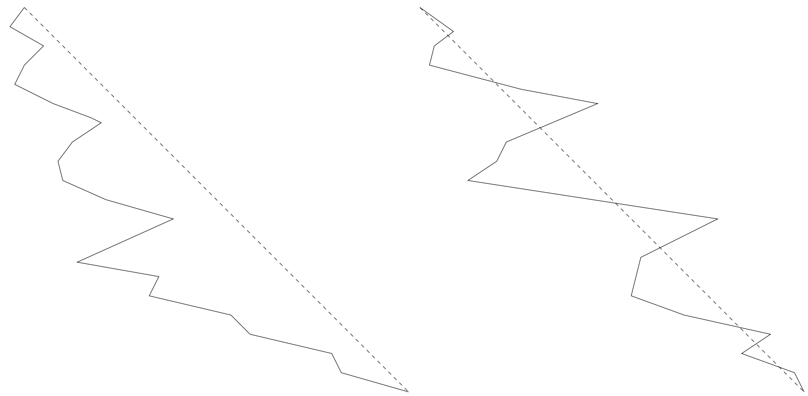
To adapt the above argument to branching random walks, one needs to understand the similarity and difference between the two models in Theorem 1 and Theorem 1'. If we define

$$N^{brw}_{n,y} = \sum_{v\in\mathbb{D}_n} 1_{\{S_v > \sqrt{2\log 2}n - \frac{3}{2\sqrt{2\log 2}}\log n + y\}}, \tag{1.7}$$

the first moment $EN^{brw}_{n,y}$, by a similar estimate as (1.5), will go to ∞ as $n \to \infty$, failing to obtain the right bound.

If we compare the results in Theorem 1 and Theorem 1', both M_n and M'_n are linear in n (with the same speed) in the leading order term, and the difference lies in the second order correction. The similarity in the order n terms comes from the point of view of large deviations. The difference comes from the different constraints along the path. As roughly illustrated in Figure 1.2, branching random walks typically do not fluctuate to the right of the straight line leading to the maximum (see figure 1.2(a)) due to the intricate dependence of the particles, but independent random walks get enough room on both sides to fluctuate at the intermediate times (see figure 1.2(b)). This fluctuation difference is because, at any intermediate time, there are less random walks in branching random walks than in independent random walks. For example, at time $n/2$, there are $2^{n/2}$ walks (particles) in branching random walks, but there are 2^n walks in independent random walks.

We can modify the set in (1.7) by taking account of the intermediate fluctuation constraints, and prove Theorem 1 using a ballot theorem. We content ourselves with the



(a) Branching Random Walk (b) Independent Random Walks

Figure 1.2: Typical path leading to the maximal displacement. Vertical direction: time; horizontal direction: Space.

intuition in the above paragraph; more details can be found in the papers by Bramson [14] and Addario-Berry and Reed [1]. Part of the rigorous proof will be repeated within the argument in Chapter 4.

More importantly, we will apply similar ideas and methods, as discussed above, in Chapter 2 and 4 to prove other facts concerning branching random walks.

1.2.2 Recursion

In this subsection, we discuss one method for proving part (i) of Theorem 1. There are several different methods available to prove tightness, for example, there is a simple argument by Dekking and Host [20] which can be used to prove tightness combined with some probability estimates. The method relevant to this thesis is the one used in Bramson and Zeitouni [16]. This method involves a recursion on $F_n(\cdot)$, the distribution of M_n.

The continuous analog of the recursion is the well-known KPP (Kolmogorov-Petrovsky-Piscounov) equation. Consider binary branching Brownian motion, see Definition 3 with $b = 2$. Let $\mathfrak{n}(t)$ be the number of particles at time t and let $\{x_1(t), \ldots, x_{\mathfrak{n}(t)(t)}\}$ be the displacement of those particles, then the maximal displacement at time t is defined as

$M_t^{bbm} = \max_{k=1}^{\mathtt{n}(t)} x_k(t)$. The distribution of the maximum, $u(t,x) = P(M_t^{bbm} \le x)$, satisfies the initial value problem

$$\begin{cases} u_t(t,x) = \frac{1}{2}u_{xx}(t,x) + u(u-1), \\ u(0,x) = 1_{[0,\infty)}(x). \end{cases} \tag{1.8}$$

This is a special case of the KPP equation. The derivation can be found in Mckean [64] and Bramson [14], and an analog for branching random walks will be presented in the next paragraph. KPP equations have been and are still studied extensively in the field of partial differential equations. A well-known characteristic of the solution is the traveling wave phenomenon, that is, $u(t, m(t) + x) \to w(x)$ as $t \to \infty$ for some distribution function $w(x)$ and some centering term $m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \log t + O(1)$, as was proved by Bramsom [14].

Let us come back to branching random walks. Recall that M_n is the maximum at time n and $F_n(\cdot)$ is its distribution. Let u and v be the two particles at time 1, then $M_n^u = \max_{\{w\in\mathbb{D}_{n+1}, u<w\}}(S_w - S_u)$ and $M_n^v = \max_{\{w\in\mathbb{D}_{n+1}, v<w\}}(S_w - S_v)$ have the same distribution as M_n. Thus

$$\begin{aligned} F_{n+1}(x) &= P(M_{n+1} \le x) = P(S_u + M_n^u < x, S_v + M_n^v \le x) \\ &= \left(P(S_u + M_n^u < x)\right)^2 = \left(G * F_n(x)\right)^2, \end{aligned}$$

where $*$ means the usual convolution. We can then deduce a recursion on the tail $\bar{F}_n(x) = 1 - F(x)$ as

$$\bar{F}_{n+1}(x) = 1 - \left(G * (1-\bar{F}_n)(x)\right)^2 = 1 - \left(1 - G * \bar{F}_n(x)\right)^2 = Q\left(G * \bar{F}_n(x)\right), \tag{1.9}$$

where $Q(x) = 1 - (1-x)^2 = 2x - x^2$. Note that $\bar{F}_0(x) = 1_{(-\infty,0)}(x)$ since $M_0 = 0$. The idea to prove the tightness is that $\bar{F}_n$ is never too flat for any n large, that is, for fixed $\eta \in (0,1)$, there exist an $\epsilon = \epsilon(\eta) > 0$, an N and an M such that, if $n > N$ and $\bar{F}_n(x - M) \le \eta$, then $\bar{F}_n(x - M) \ge (1+\epsilon)\bar{F}_n(x)$.

Intuitively, there are two operators in (1.9): the convolution and Q. The convolution has an effect of flattening $\bar{F}_n$ out, but the Q operator makes the curve steep again. To capture the change of flatness, Bramson and Zeitouni employed a Lyapunov function

$$L(u) = \sup_{x:u(x)\in(0,\delta_0]} l(u;x), \tag{1.10}$$

where

$$l(u;x) = \log\left(\frac{1}{u(x)}\right) + \log_b\left(1 + \epsilon_1 - \frac{u(x-M)}{u(x)}\right)_+ .$$

Here, δ_0, b and ϵ_1 are constants chosen to make the argument work, and $u : \mathbb{R} \to [0,1]$ is a decreasing function with $\lim_{x\to-\infty} u(x) = 1$ and $\lim_{x\to\infty} u(x) = 0$. With this choice of the Lyapunov function, they showed that $\sup_n L(\bar{F}_n) < \infty$, which implied that $\bar{F}_n$ cannot be too flat in the right tail. This, together with some other auxiliary lemmas, easily showed that $\bar{F}_n$ cannot be too flat in the left tail either. Thus, the family $\{M_n - Med(M_n)\}_n$ is tight.

We will extend the argument above in Chapter 3 to handle the tightness of the maximal displacement of branching random walks in random medium with local dependences.

1.3 Summary of Results in This Thesis

With the intuition gained on the simplified BRW model from Section 1.2, we now explain the results in the following three chapters and some of the heuristics in proving the results. The three chapters are based on three papers (two of which, those from chapters 2 and 4, coauthored with Ofer Zeitouni), and can be read independently. A short summary of the models and results, without detailed assumptions, is presented below. We refer the details and precise statements to the introductions of the corresponding chapters.

In Chapter 2, we study the consistent minimum of the BRW model defined in Definition 1 and Definition 1'. The minimum m_n, defined by (1.3), is the displacement of the leftmost particle at time n. From Theorem 1 in Section 1.2 and the symmetry between m_n and M_n as commented in footnote 3, we know that

$$\lim_{n\to\infty} \frac{m_n}{n} = c, \quad a.s.,$$

for some constant c. The consistent minimum describes the 'leftmost' path up to time n. To characterize the 'leftmost' path, we define

$$L_n = \min_{v\in\mathbb{D}_n} \max_{k=0}^{n} (S_{v^k} - ck),$$

which describes the maximal deviation along the 'leftmost' path from m_k for $k \le n$.

From the intuition on the path leading to m_n, see Figure 1.2 for the path to M_n, for some intermediate level k, the particle on that path deviates (to the right) from the leftmost particle (roughly at ck) by a relatively large distance (at least order $n^{1/2}$). We can find other paths, along which the maximal deviation from the leftmost particles among all intermediate levels is smaller than that along the path leading to m_n. The notation L_n is the minimum of such maximal deviations, and the path achieving L_n is the 'leftmost' path.

Regarding L_n, we prove the following, by understanding the profile of the path achieving L_n and then applying similar moment methods as in Section 1.2.1 to some appropriate set of walks.

Theorem in Chapter 2. *Under appropriate assumptions on $G(\cdot)$, we have*

$$\lim_{n\to\infty} \frac{L_n}{n^{1/3}} = l_0, \quad a.s.$$

for some explicit constant l_0.

In Chapter 3, we extend the tightness result (see part (i) in Theorem 1) to a wider class of branching random walks, where the increments of siblings may depend on each other and the laws of the branching and increments may vary with respect to time. Instead of the deterministic constant b and the unique distribution function $G(\cdot)$ in Definition 1, we have a sequence of distributions $\{p_{n,k}\}_{n\geq 0,k\geq 1}$ on $\mathbb{Z}_+$ and a sequence of distributions $\{G_{n,k}(x_1,\ldots,x_k)\}_{n\geq 0,k\geq 1}$ on $\mathbb{R}^k$. A particle at time n independently gives birth to k offspring with probability $p_{n,k}$, and the increments of the k offspring are distributed according to a joint distribution $G_{n,k}(x_1,\ldots,x_k)$ and independent of everything else. For $m \leq n$, let M_n^m denote the maximum at time n of a BRW starting from location 0 at time m, $F_n^m(\cdot)$ the distribution of M_n^m and $\bar{F}_n^m(\cdot) = 1 - F_n^m(\cdot)$. Similar to (1.9), we have a recursion on $\bar{F}_n^m(\cdot)$: for $m < n$,

$$\bar{F}_n^m(x) = 1 - \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x - y_i)\right) d^k G_{m,k}(y_1,\ldots,y_k). \tag{1.11}$$

The tightness result follows from a modified analysis of the recursion (1.11) based on Bramson and Zeitouni [16], where they derived the tightness for the recursion (1.9). The recursion (1.11) can be reduced to recursive inequalities of simpler forms. The argument in Bramson and Zeitouni [16], which analyzed the recursion (1.9) and its pointwise nonlinearity, is adapted to study the recursive inequalities and to handle some global nonlinearity. We prove

Theorem in Chapter 3. *Assume that $G_{n,k}$ are tight, their marginal distributions have uniform super-exponential decay right tails, $p_{n,k}$ have uniformly bounded support. Then, under some additional technical assumptions, the sequence $\{M_n - Med(M_n)\}_n$ is tight.*

In Chapter 4, we consider the maximal displacement of BRWs in certain time inhomogeneous environments. Specifically, in Definition 1, $b = 2$ (binary branching), and the increments are still independent but distributed as $N(0, \sigma_1^2)$ before time $n/2$ and $N(0, \sigma_2^2)$ after time $n/2$, if we consider the model up to time n. The maxima (also denoted by M_n), re-centered by their medians, are still tight by the result from Chapter 3. Our goal is to characterize the mean of M_n up to an $O(1)$ error. The argument depends on moment methods and analysis of the best profile leading the maximum. A large deviation calculation provides a good hint on the difference between the optimal profiles in the homogeneous and inhomogeneous environments. We obtain the following

Theorem in Chapter 4. *The maximum M_n of the BRW in the time inhomogeneous environment described above satisfies that the sequence $\{M_n - Med(M_n)\}_n$ is tight and*

$$Med(M_n) = (\sqrt{2 \log 2}\sigma_{\text{eff}})n - \beta \frac{\sigma_{\text{eff}}}{\sqrt{2 \log 2}} \log n + O(1)$$

for some σ_{eff} and β, which change with the different ordering of σ_1^2 and σ_2^2 as follows:

(i) when $\sigma_1^2 < \sigma_2^2$ (increasing variances), $\sigma_{\text{eff}} = \sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}$ and $\beta = \frac{1}{2}$;

(ii) when $\sigma_1^2 > \sigma_2^2$ (decreasing variances), $\sigma_{\text{eff}} = \frac{\sigma_1+\sigma_2}{2}$ and $\beta = 3$.

The homogeneous case ($\sigma_1^2 = \sigma_2^2$) corresponds to part (ii) in Theorem 1. An interesting phenomenon here is: the profile of the environments matters, both in the leading (velocity) term and in the logarithmic correction term, and the latter exhibits a phase transition in the sense that the logarithmic term is discontinuous as $\sigma_1^2 \to \sigma_2^2$.

Chapter 2

Consistently Minimal Displacement of Branching Random Walks

2.1 Introduction

Recall as in Chapter 1 that a branching random walk is a process describing a particle performing random walk while branching. In this chapter, we consider the 1-dimensional case as follows. At time 0, there is one particle at location 0. At time 1, the particle splits into b particles ($b \in \mathbb{Z}_+$ deterministic and $b \geq 2$ to avoid trivial cases), each of which moves independently to a new position according to some distribution function $F(x)$. Then at time 2, each of the b particles splits again into b particles, which again move independently according to the distribution function $F(x)$. The splitting and moving continue at each integer time and are independent of each other. This procedure produces a 1-dimensional branching random walk.

To describe the relation between particles, we associate to each particle a vertex in a b-ary rooted tree $\mathbb{T} = \{V, E\}$ with root o, where each vertex has b children; V is the set of vertices in $\mathbb{T}$ and E is the set of edges in $\mathbb{T}$. The root o is associated with the original particle. The b children of a vertex $v \in V$ correspond to the b particles from the splitting of the particle corresponding to v. In particular, the vertices whose distance from o is n, denoted by $\mathbb{D}_n$, correspond to particles at time n. To describe the displacement between particles, we assign i.i.d. random variables X_e with common distribution $F(x)$ to each edge $e \in E$. (Throughout, we let $e = uv$ denote the edge e connecting two vertices $u, v \in V$.) For each vertex $v \in V$, we use $|v|$ to denote its distance from o and use v^k to denote the ancestor of v in $\mathbb{D}_k$ for any $0 \le k \le |v|$. Then the positions of particles at time n can be described by $\{S_v | v \in \mathbb{D}_n\}$, where for $v \in \mathbb{D}_n$, $S_v = \sum_{i=0}^{n-1} X_{v^i v^{i+1}}$.

The limiting behavior of the maximal displacement $M_n = \max_{v \in \mathbb{D}_n} S_v$ or the minimal displacement $m_n = \min_{v \in \mathbb{D}_n} S_v$ as $n \to \infty$ has been extensively studied in the literature (See in particular Bramson [14],[15], Addario-Berry and Reed [1], and references therein.) Throughout this chapter, we assume that

$$Ee^{\lambda X_e} < \infty \quad \text{for some } \lambda < 0 \text{ and some } \lambda > 0. \tag{2.1}$$

Then the Fenchel-Legendre transform of the log-moment generating function $\Lambda(\lambda) = \log Ee^{\lambda X_e}$,

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}} (\lambda x - \Lambda(\lambda)), \tag{2.2}$$

is the large deviation rate function (see [21, Ch. 1,2]) of a random walk with step distribution $F(x)$. In addition to (2.1), we also assume that, for some $\lambda_- < 0$ and $\lambda_+ > 0$ in the interior of $\{\lambda : \Lambda(\lambda) < \infty\}$,

$$\lambda_\pm \Lambda'(\lambda_\pm) - \Lambda(\lambda_\pm) = \log b, \tag{2.3}$$

which implies that $\Lambda^*(\Lambda'(\lambda_\pm)) = \log b$. These assumptions imply that

$$M := \lim_{n\to\infty} \frac{M_n}{n} = \Lambda'(\lambda_+) \quad \text{and} \quad m := \lim_{n\to\infty} \frac{m_n}{n} = \Lambda'(\lambda_-) \quad a.s.. \tag{2.4}$$

See [1] for more details on (2.4).

The *offset* of the branching random walk is defined as the minimal deviation of the path up to time n from the line leading to mn (roughly, the minimal position at time n). Explicitly, set

$$L_n = \min_{v \in \mathbb{D}_n} \max_{k=0}^{n} (S_{v^k} - mk). \tag{2.5}$$

(See Figure 2.1 for a pictorial description of L_3.) Without loss of generality, subtracting the deterministic constant $\Lambda'(\lambda_-)$ from each increment $\{X_e\}$, we can and will assume that

$$m = \Lambda'(\lambda_-) = 0. \tag{2.6}$$

Under this assumption, (2.3) and (2.5) simplify to

$$-\Lambda(\lambda_-) = \log b, \tag{2.3$'$}$$

$$L_n = \min_{v \in \mathbb{D}_n} \max_{k=0}^{n} S_{v^k}. \tag{2.5$'$}$$

In the process of studying random walks in random environments on trees, Hu and

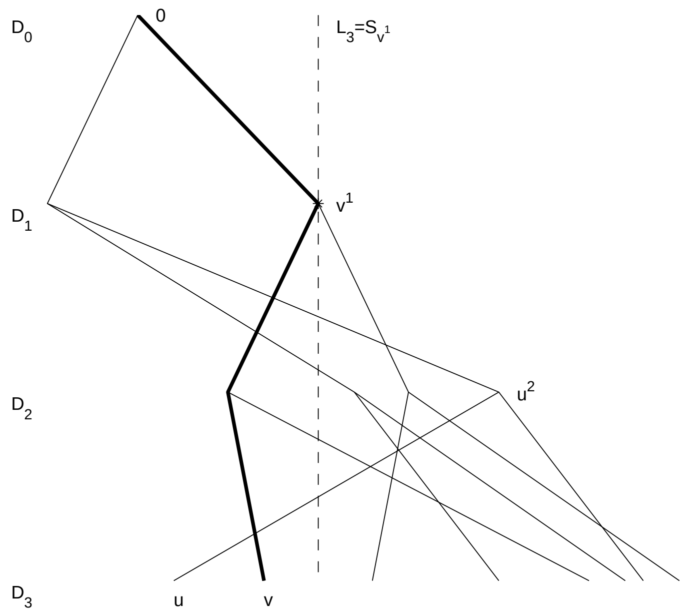


Figure 2.1: Figure for L_3 when m=0 and b=2

Shi [47] (2007) discovered that the offset has order $n^{1/3}$ in the following sense: there exist constants $c_1, c_2 > 0$ such that

$$c_1 \le \liminf_{n\to\infty} \frac{L_n}{n^{1/3}} \le \limsup_{n\to\infty} \frac{L_n}{n^{1/3}} \le c_2. \tag{2.7}$$

They raised the question as to whether the limit of $L_n/n^{1/3}$ exists. In this note, we answer this affirmatively and prove the following.

Theorem 2. *Under assumption (2.1) and (2.3) and with* $l_0 = \sqrt[3]{\frac{3\pi^2\sigma_Q^2}{2|\lambda_-|}}$, *it holds that*

$$\lim_{n\to\infty} \frac{L_n}{n^{1/3}} = l_0 \quad a.s. \tag{2.8}$$

In the expression for l_0, $\lambda_- < 0$ by the definition (2.3) and σ_Q^2 is the variance that will be defined in (2.10).

The proof of the theorem is divided into two parts - the lower bound (2.21) and the upper bound (2.32). In Section 2, we review a result from Mogul'skii [66], which will be the key estimate in our proof. In Section 3, we apply a first moment argument (with some modification) in order to study the minimal positions for intermediate levels with the restriction that the walks do not exceed $ln^{1/3}$ for some $l > 0$ at all time. This yields the lower bound for L_n. In section 4, we apply a second moment argument to the lower bound $P(L_n \leq ln^{1/3})$ for certain values of l. Compared with standard applications of the second moment method in related problems, the analysis here requires the control of second order terms in the large deviation estimates. Truncation of the tree is then used to get independence and complete the proof of the upper bound.

The offset is determined by a competition between two terms: a displacement term (whose cost is exponential in the displacement) and an entropy term (reflecting the difficulty in keeping the walk confined in a narrow tube, and with cost proportional to the exponent of the time divided by width squared; this is made precise in Theorem 3). Roughly speaking, in a time interval of length Δt and displacement width Δl, the cost is of the form $e^{c_1\Delta l - c_2 \Delta t/(\Delta l)^2}$. One then sees that the optimum is achieved at Δl proportional to $(\Delta t)^{1/3}$. This gives the scaling on $n^{1/3}$ to the displacement. In the actual proof, when optimizing the cost, a certain curve $s(t)$, see (2.17), emerges. The curve $s(t)$ reflects the location of the minimal position of intermediate levels, and plays an important role also in the second moment computation, see a discussion in Section 2.5.1.

2.2 An Auxiliary Estimate: the Absorption Problem for Random Walk

We derive in this section some estimates for random walk with i.i.d. increments $\{X_i\}_{i\geq 1}$ distributed according to a law P with $P((-\infty, x]) = F(x)$ satisfying (2.1), (2.3) and (2.6). Define

$$S_n(t) = \frac{X_0 + X_1 + \cdots + X_k}{n^{1/3}} \quad \text{for } \frac{k}{n} \leq t < \frac{k+1}{n},\ k = 0, 1, \ldots, n-1,$$

where $X_0 = 0$. Note that due to (2.6), $EX_i > 0$. Introduce the auxiliary law

$$\frac{dQ}{dP} = e^{\lambda_- X_1 - \Lambda(\lambda_-)}. \tag{2.9}$$

Under Q, $E_Q X_1 = 0$. The variance of X_1 under Q is denoted by

$$\sigma_Q^2 = E_Q X_1^2. \tag{2.10}$$

In the following estimates, $f_1(t)$ and $f_2(t)$, which may take the value $\pm\infty$, are right-continuous and piecewise constant functions on $[0, 1]$. Let $G := \{(t, g(t)) \in \mathbb{R}^2 : 0 \leq t \leq 1, f_1(t) < g(t) < f_2(t)\}$ be a region in $\mathbb{R}^2$ bounded by the graphs of $f_1(t)$ and $f_2(t)$. Assume also that G contains the graph of a continuous function.

Theorem 3. *(Mogul'skii [66, Theorem 3]) Under the above assumptions,*

$$Q(S_n(\cdot) \in G) = e^{-\frac{\pi^2\sigma_Q^2}{2} H_2(G) n^{1/3} + o(n^{1/3})}, \tag{2.11}$$

where

$$H_2(G) = \int_0^1 \frac{1}{(f_1(t) - f_2(t))^2} dt. \tag{2.12}$$

In the following, we will need to control the dependence of the estimate (2.11) on the starting point.

Corollary 1. *With notation and assumptions as in Theorem 3, for any $\epsilon > 0$, there is a $\delta > 0$ such that, for any interval $I \subset (f_1(0), f_2(0))$ with length $|I| \leq \delta$, we have*

$$\sup_{x \in I} Q(x + S_n(\cdot) \in G) \leq e^{-(\frac{\pi^2\sigma_Q^2}{2} H_2(G) - \epsilon) n^{1/3} + o(n^{1/3})}. \tag{2.13}$$

Proof Let $I = (a, b)$ and $G_x := \{(t, y) : 0 \leq t \leq 1, f_1(t) - x < y < f_2(t) - x\}$ be the shift of G by x. Set $G' = G_a \cup G_b$. We have

$$\sup_{x \in I} Q(x + S_n(\cdot) \in G) = \sup_{x \in I} Q(S_n(\cdot) \in G_x) \leq Q(S_n(\cdot) \in G') = e^{-\frac{\pi^2\sigma_Q^2}{2} H_2(G') n^{1/3} + o(n^{1/3})}.$$

Since $H_2(G') = \int_0^1 \frac{1}{(f_2(t) - f_1(t) + (b-a))^2} dt \uparrow H_2(G)$ as $|I| = (b - a) \to 0$ uniformly in the position of I, the lemma is proved. $\square$

2.3 Lower Bound

Consider the branching random walk up to level n. In this and the next section, we estimate the number of particles that stay constantly below $ln^{1/3}$, i.e.,

$$N_n^l = \sum_{v \in \mathbb{D}_n} 1_{\{S_{v^k} \le ln^{1/3} \text{ for } k=0,1,\ldots,n\}}. \tag{2.14}$$

In order to get a lower bound on the offset, we apply a first moment method with a small modification: while it is natural to just calculate the first moment of N_n^l, such a computation ignores the constraint on the number of particles at level k imposed by the tree structure. In particular, EN_n^l for branching random walks is the same as the one for b^n independent random walks. An easy first and second moment argument shows that the limit in (2.8) is 0 for b^n independent random walks, and thus no useful upper bound can be derived in this way.

To address this issue, we use a more delicate first moment argument. Namely, we look at the vertices not only at level n but also at some intermediate levels. Divide the interval $[0, n]$ into $1/\epsilon$ equidistant levels, with $1/\epsilon$ an integer. Define recursively, for any $\delta > 0$,

$$\begin{cases} s_0 = 0, w_0 = l + \delta; \\ s_k = s_{k-1} - \frac{\pi^2 \sigma_Q^2}{2\lambda_- w_{k-1}^2}\epsilon, \; w_k = l + \delta - s_k \text{ for } k = 1, \ldots, \frac{1}{\epsilon}. \end{cases} \tag{2.15}$$

For particles staying below $ln^{1/3}$, s_k will be interpreted as values such that the walks between times $k\epsilon n$ and $(k+1)\epsilon n$ never go below $(s_k - \delta)n^{1/3}$, and $w_k n^{1/3}$ will correspond to the width of the window $W_k = ((s_k - \delta)n^{1/3}, ln^{1/3})$ that we allow between level $k\epsilon n$ and $(k+1)\epsilon n$ when considering those walks that do not go below $(s_k - \delta)n^{1/3}$ or go above $ln^{1/3}$.

Before calculating the first moment, consider the recursion (2.15) for s_k. Rewrite it as

$$s_k = s_{k-1} - \frac{\pi^2 \sigma_Q^2}{2\lambda_-(l + \delta - s_{k-1})^2}\epsilon. \tag{2.16}$$

This is an Euler's approximation sequence for the solution of the following differential equation

$$s'(t) = -\frac{\pi^2 \sigma_Q^2}{2\lambda_-(\alpha - s(t))^2}, \; s(0) = 0, \tag{2.17}$$

where $\alpha = l + \delta$. The above initial value problem has the solution $s^{\alpha}(t) = \alpha + \sqrt[3]{-\frac{3\pi^2\sigma_Q^2}{2\lambda_-}t - \alpha^3}$. Here we find

$$l_0 = \sqrt[3]{\frac{3\pi^2\sigma_Q^2}{-2\lambda_-}} \tag{2.18}$$

such that $s^{l_0}(1) = l_0$.

For any $l_1 < l_0$, we can choose $\delta > 0$ and $l_1 + \delta < l_0$. In this case, $s^{l_1+\delta}(1) > l_1 + \delta > l_1$. If we choose such l_1 and δ in (2.15), it is easy to check that the sequence $\{s_k\}_{k=0}^{\frac{1}{\epsilon}}$ will be greater than l_1 somewhere in the sequence. Define

$$K = \min\{k : s_k \geq l_1\}. \tag{2.19}$$

For fixed $\gamma > 0$ small enough, we can choose ϵ small such that

$$K\epsilon < 1 - \gamma. \tag{2.20}$$

For $k < K - 1$, let Z_k denote the number of vertices v between level $k\epsilon n$ and $(k+1)\epsilon n$ with $S_v < (s_k - \delta)n^{1/3}$. Denote by Z_{K-1} the number of vertices w between level $(K-1)\epsilon n$ and n with $S_w < (s_{K-1} - \delta)n^{1/3}$. Denote by Z the number vertices $v \in \mathbb{D}_n$ whose associated walks stay in W_k between level $k\epsilon n$ and $(k+1)\epsilon n$ for $k < K$ and then stay in W_{K-1} up to level n. Explicitly,

$$Z_0 = \sum_{i=1}^{\lfloor \epsilon n \rfloor} \sum_{v \in \mathbb{D}_i} 1_{\{S_v < -\delta n^{1/3}\}},$$

$$Z_k = \sum_{i=\lfloor k\epsilon n \rfloor + 1}^{\lfloor (k+1)\epsilon n \rfloor} \sum_{v \in \mathbb{D}_i} 1_{\{S_v < (s_k - \delta)n^{1/3},\ S_{v_d} \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<k\}},\quad 0 < k < K-1,$$

$$Z_{K-1} = \sum_{i=\lfloor K\epsilon n \rfloor + 1}^{n} \sum_{v \in \mathbb{D}_i} 1_{\{S_v < (s_{K-1} - \delta)n^{1/3},\ S_{v_d} \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<K-1\}},$$

$$Z = \sum_{v \in \mathbb{D}_n} 1_{\{S_{v_d} \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<K,\ S_{v_d} \in W_{K-1} \text{ for } K\epsilon n \leq d \leq n\}}.$$

Observe that

$$N_n^{l_1} \leq \sum_{k=0}^{K-1} Z_k + Z.$$

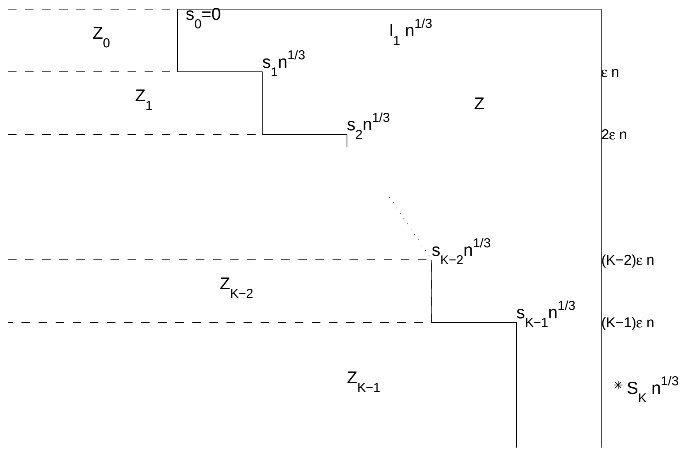


Figure 2.2: The relation between Z_k's and s_k's.

Using Theorem 3, we provide upper bounds for the first moment of Z_k and Z. Starting with Z_0, we have

$$\begin{aligned} EZ_0 &= \sum_{i=1}^{\lfloor \epsilon n \rfloor} b^i E 1_{\{S_i < -\delta n^{1/3}\}} = \sum_{i=1}^{\lfloor \epsilon n \rfloor} b^i E_Q e^{-\lambda_- S_i + i\Lambda(\lambda_-)} 1_{\{S_i < -\delta n^{1/3}\}} \\ &\leq \sum_{i=1}^{\lfloor \epsilon n \rfloor} e^{\lambda_- \delta n^{1/3}} E_Q 1_{\{S_i < -\delta n^{1/3}\}} \leq \sum_{i=1}^{\lfloor \epsilon n \rfloor} e^{\lambda_- \delta n^{1/3}} \leq e^{\lambda_- \delta n^{1/3} + o(n^{1/3})}, \end{aligned}$$

where we used the change of measure (2.9) in the second equality, and (3′) and the fact that $\lambda_- < 0$ in the first inequality.

For $0 < k < K-1$, using again the change of measure (2.9), we get

$$\begin{aligned} EZ_k &= \sum_{i=\lfloor k\epsilon n \rfloor + 1}^{\lfloor (k+1)\epsilon n \rfloor} b^i E 1_{\{S_i < (s_k - \delta) n^{1/3},\ S_d \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<k\}} \\ &= \sum_{i=\lfloor k\epsilon n \rfloor + 1}^{\lfloor (k+1)\epsilon n \rfloor} E_Q e^{-\lambda_- S_i} 1_{\{S_i < (s_k - \delta) n^{1/3},\ S_d \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<k\}} \\ &\leq e^{-\lambda_- (s_k - \delta) n^{1/3}} \sum_{i=\lfloor k\epsilon n \rfloor + 1}^{\lfloor (k+1)\epsilon n \rfloor} E_Q 1_{\{S_i < (s_k - \delta) n^{1/3},\ S_d \in W_j \text{ for } j\epsilon n \leq d \leq (j+1)\epsilon n \text{ and } j<k\}}. \end{aligned}$$

Therefore,

$$\begin{aligned} EZ_k &\le e^{-\lambda_-(s_k-\delta)n^{1/3}} \sum_{i=\lfloor k\epsilon n\rfloor+1}^{\lfloor (k+1)\epsilon n i\rfloor} Q(S_d\in W_j \text{ for } j\epsilon n\le d\le (j+1)\epsilon n \text{ and } j<k) \\ &= e^{-\lambda_-(s_k-\delta)n^{1/3}} \sum_{i=\lfloor k\epsilon n\rfloor+1}^{\lfloor (k+1)\epsilon n\rfloor} e^{-\sum_{j=0}^{k-1}\frac{\pi^2\sigma_Q^2}{2w_j^2}\epsilon n^{1/3}+o(n^{1/3})} \\ &\le e^{-\lambda_-(s_k-\delta)n^{1/3}-\sum_{j=0}^{k-1}\frac{\pi^2\sigma_Q^2}{2w_j^2}\epsilon n^{1/3}+o(n^{1/3})} = e^{\lambda_-\delta n^{1/3}+o(n^{1/3})}, \end{aligned}$$

where (2.11) with the choice of $G=\cup_{j=0}^{k-1}\left\{[j\epsilon,(j+1)\epsilon)\times W_j/n^{1/3}\right\}\cup\{[k\epsilon,1]\times(-\infty,\infty)\}$ is applied in the first equality, and (2.15) in the second.

The calculation of EZ_{K-1} is almost the same as EZ_k except that we replace the summation limits above by $(K-1)\epsilon n+1$ and n and that we replace the k in the summand by $K-1$. Thus, we get the same upper bound for EZ_{K-1},

$$EZ_{K-1}\le e^{\lambda_-\delta n^{1/3}+o(n^{1/3})}.$$

We estimate EZ similarly as follows. First, use the change of measure (2.9) to get

$$\begin{aligned} EZ &= b^n E1_{\{S_d\in W_j \text{ for } j\epsilon n\le d\le(j+1)\epsilon n \text{ and } j<K,\ S_d\in W_{K-1}, \text{ for } K\epsilon n\le d\le n\}} \\ &= E_Q e^{-\lambda_- S_n} 1_{\{S_d\in W_j \text{ for } j\epsilon n\le d\le(j+1)\epsilon n \text{ and } j<K,\ S_d\in W_{K-1}, \text{ for } K\epsilon n\le d\le n\}} \\ &\le e^{-\lambda_- l_1 n^{1/3}} E_Q 1_{\{S_d\in W_j \text{ for } j\epsilon n\le d\le(j+1)\epsilon n \text{ and } j<K,\ S_d\in W_{K-1}, \text{ for } K\epsilon n\le d\le n\}}. \end{aligned}$$

Then, applying (2.11) with $G=\cup_{j=0}^{K-1}\left\{[j\epsilon,(j+1)\epsilon)\times W_j/n^{1/3}\right\}\cup\{[K\epsilon,1]\times W_{K-1}/n^{1/3}\}$ in the first equality, we get

$$\begin{aligned} EZ &\le e^{-\lambda_- l_1 n^{1/3}} E_Q 1_{\{S_d\in W_j \text{ for } j\epsilon n\le d\le(j+1)\epsilon n \text{ and } j<K,\ S_d\in W_{K-1}, \text{ for } K\epsilon n\le d\le n\}} \\ &= e^{-\lambda_- l_1 n^{1/3}-\sum_{i=0}^{K-1}\frac{\pi^2\sigma_Q^2}{2w_i^2}\epsilon n^{1/3}-\frac{\pi^2\sigma_Q^2}{2w_{K-1}^2}(1-K\epsilon)n^{1/3}+o(n^{1/3})} \\ &\le e^{-\gamma\frac{\pi^2\sigma_Q^2}{2l_1^2}n^{1/3}+o(n^{1/3})}, \end{aligned}$$

where the last inequality is obtained by noting that $l_1\le S_K=-\sum_{i=0}^{K-1}\frac{\pi^2\sigma_Q^2}{2\lambda_- w_i^2}\epsilon$ by (2.19) and (2.15), and then recalling (2.20) and $w_{K-1}<l_1$.

In conclusion, we proved that $E(\sum_{k=0}^{K-1} Z_k + Z) \leq e^{-c_3 n^{1/3} + o(n^{1/3})}$ for some $0 < c_3 < \min\{-\lambda_- \delta, \gamma \frac{\pi^2 \sigma_Q^2}{2l_1^2}\}$. Since $\sum_{k=0}^{K-1} Z_k + Z$ is an integer valued random variable, we have

$$P(\sum_{k=0}^{K-1} Z_k + Z > 0) = P(\sum_{k=0}^{K-1} Z_k + Z \geq 1) \leq E(\sum_{k=0}^{K-1} Z_k + Z) \leq e^{-c_3 n^{1/3} + o(n^{1/3})}.$$

By the Borel-Cantelli lemma, we have $\sum_{k=0}^{K-1} Z_k + Z = 0$ a.s. for all large n. So is $N_n^{l_1} = 0$, which means that $L_n > l_1 n^{1/3}$ a.s. for all large n. Since $l_1 < l_0$ is arbitrary, we conclude that

$$\liminf_{n\to\infty} \frac{L_n}{n^{1/3}} \geq l_0 \quad a.s.. \tag{2.21}$$

This completes the proof of the lower bound in Theorem 2.

2.4 Upper Bound

2.4.1 A Second Moment Method Estimate

In this section, we consider any fixed $l_2 > l_0$. A second moment argument will provide a lower bound for the probability that we can find at least one walk which stays in the interval W_k between level $k\epsilon n$ and $(k+1)\epsilon n$ for all k. A truncation (of the tree) argument will complete the proof of the upper bound.

As a first step, consider the sequence $\{s_k\}$ in (2.15) with $l_2 > l_0$. Then for any $\delta > 0$, it is easy to see that $s^{l_2+\delta}(t)$ is increasing and convex for $0 \leq t \leq 1$. Thus in the Euler's approximation (2.16) of the initial value problem (2.17),

$$s_{\frac{1}{\epsilon}} < s^{l_2+\delta}(1) < s^{l_2}(1) < l_2. \tag{2.22}$$

It follows from (2.15) that

$$w_k \geq \delta \quad \text{for all } \ 0 \leq k \leq \frac{1}{\epsilon} - 1. \tag{2.23}$$

Define $\tilde{N}_n^{l_2}$ as follows:

$$\tilde{N}_n^{l_2} = \sum_{v\in\mathbb{D}_n} 1_{\{S_{vj}\in W_k,\ \text{for}\ k\epsilon n\leq j\leq(k+1)\epsilon n,\ k=0,\ldots,\frac{1}{\epsilon}-1\}}.$$

We will apply second moment method to $\tilde{N}_n^{l_2}$. $E\tilde{N}_n^{l_2}$ is calculated the same way as EZ in the previous section. But this time we consider $G = \{\cup_{j=0}^{\frac{1}{\epsilon}-1} W_j/n^{1/3} \times [j\epsilon, (j+1)\epsilon)\} \cup$

$\{(l_2-\Delta l_2, l_2)\times\{1\}\}$ in (2.11) with $\Delta l_2 \to 0$, so

$$\begin{aligned} E\tilde{N}_n^{l_2} &= b^n E 1_{\{S_j\in W_k,\ \text{for}\ k\epsilon n\le j\le (k+1)\epsilon n,\ k=0,\ldots,\frac{1}{\epsilon}-1\}} \\ &= E_Q e^{-\lambda_- S_n} 1_{\{S_j\in W_k,\ \text{for}\ k\epsilon n\le j\le (k+1)\epsilon n,\ k=0,\ldots,\frac{1}{\epsilon}-1\}} \\ &= e^{(-\lambda_- l_2 - \sum_{k=0}^{\frac{1}{\epsilon}-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon)n^{1/3}+o(n^{1/3})}. \end{aligned} \tag{2.24}$$

From (2.22) and the definition (2.15) of s_k, $-\lambda_- l_2 - \sum_{k=0}^{\frac{1}{\epsilon}-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon > 0$ and thus $E\tilde{N}_n^{l_2}\to\infty$.

Therefore, we will be ready to apply the second moment method after the following calculations:

$$\begin{aligned} E(\tilde{N}_n^{l_2})^2 &= E\sum_{u,v\in\mathbb{D}_n} 1_{\{S_{u^j},S_{v^j}\in W_k,\ \text{for}\ k\epsilon n\le j\le (k+1)\epsilon n,\ k=0,\ldots,\frac{1}{\epsilon}-1\}} \\ &= \sum_{h=0}^{n-1} E \sum_{\substack{u,v\in\mathbb{D}_n\\ u\wedge v\in\mathbb{D}_h}} 1_{\{S_{u^j},S_{v^j}\in W_k,\ \text{for}\ k\epsilon n\le j\le (k+1)\epsilon n,\ k=0,\ldots,\frac{1}{\epsilon}-1\}} \\ &\quad + E\tilde{N}_n^{l_2}. \end{aligned} \tag{2.25}$$

In the last expression above, $u\wedge v$ is the largest common ancestor of u and v.

Write $h = q\epsilon n + r$ for $0\le q\le \frac{1}{\epsilon}-1$ and $0\le r<\epsilon n$. There are $b^{2n-h-1}(b-1)$ indices in the second sum in the right side of (2.25). We estimate the probability for one such pair to stay in W_k's. In order to simplify the notation, define

$$p_1(0,h,x) = P(S_h\in dx,\ S_j\in W_k,\ \text{for}\ k\epsilon n\le j\le (k+1)\epsilon n\wedge h,\ k=0,\ldots,q),$$

$$p_2(h,x,n,y) = P(S_n\in dy,\ S_j\in W_k,\ \text{for}\ h\vee k\epsilon n\le j\le (k+1)\epsilon n,\ k=q,\ldots,n|S_h=x).$$

Similarly, define $q_1(0,h,x)$ and $q_2(h,x,n,y)$ to be the probability of the same events under Q.

Then, we can write $E(\tilde{N}_n^{l_2})^2$ as

$$E(\tilde{N}_n^{l_2})^2 = E\tilde{N}_n^{l_2} + \sum_{h=0}^{n-1} b^{2n-h-1}(b-1)\int_{W_q}\left(\int_{W_n} p_2(h,x,n,y)dy\right)^2 p_1(0,h,x)dx.$$

By a change of measure (2.9), the above quantity is equal to

$$\begin{aligned} & E\tilde{N}_n^{l_2} + \sum_{h=0}^{n-1} b^{2n-h-1}(b-1) \int_{W_q} \left(\int_{W_n} e^{-\lambda_-(y-x)+(n-h)\Lambda(\lambda_-)} q_2(h,x,n,y)dy \right)^2 \\ & \qquad \cdot e^{-\lambda_- x + h\Lambda(\lambda_-)} q_1(0,h,x)dx \\ \leq \quad & E\tilde{N}_n^{l_2} + \sum_{h=0}^{n-1} \frac{b-1}{b} e^{(-2\lambda_- l_2 + \lambda_-(s_q-\delta))n^{1/3}} \\ & \qquad \cdot \int_{W_q} \left(\int_{W_n} q_2(h,x,n,y)dy \right)^2 q_1(0,h,x)dx. \end{aligned} \tag{2.26}$$

We now provide an upper bound for the integral term in the right side of (2.26). We have

$$\begin{aligned} & \int_{W_q} \left(\int_{W_n} q_2(h,x,n,y)dy \right)^2 q_1(0,h,x)dx \\ \leq \quad & \left(\sup_{x\in W_q} \int_{W_n} q_2(h,x,n,y)dy \right)^2 \int_{W_q} q_1(0,h,x)dx. \end{aligned}$$

By Chapman-Kolmogorov equation, the last quantity in the above display can be rewritten as

$$\left(\sup_{x\in W_q} \int_{W_n} \int_{W_{q+1}} q_2\left(h,x,(q+1)\epsilon n, z\right) q_2\left((q+1)\epsilon n, z, n, y\right) dzdy \right)^2 \int_{W_q} q_1(0,q\epsilon n,x)dx,$$

which is equal to

$$\left(\sup_{x\in W_q} \int_{W_{q+1}} q_2(h,x,(q+1)\epsilon n,z) q_2((q+1)\epsilon n, z, n, W_n)dz \right)^2 e^{-\sum_{k=0}^{q-1} \frac{\pi^2\sigma_Q^2}{2w_k^2} n^{1/3} + o(n^{1/3})}.$$

Letting $\cup_i I_i = W_{q+1}$ be any covering of W_{q+1}, then the above quantity is less than or equal to

$$\begin{aligned} & \left(\sup_{x\in W_q} \sum_i \int_{I_i} q_2(h,x,(q+1)\epsilon n,z) q_2((q+1)\epsilon n, z, n, W_n)dz \right)^2 e^{-\sum_{k=0}^{q-1} \frac{\pi^2\sigma_Q^2}{2w_k^2} n^{1/3} + o(n^{1/3})} \\ \leq & \left(\sum_i \sup_{z\in I_i} q_2((q+1)\epsilon n, z, n, W_n) \right)^2 e^{-\sum_{k=0}^{q-1} \frac{\pi^2\sigma_Q^2}{2w_k^2} \epsilon n^{1/3} + o(n^{1/3})}. \end{aligned}$$

Due to (2.13), for any small $\epsilon_1 > 0$, we can choose a finite covering of $\{I_i\}$ and $|I_i| \le \delta_1 n^{1/3}$ such that for each i,

$$\sup_{z\in I_i} q_2((q+1)\epsilon n, z, n, W_n) \le e^{-(\sum_{k=q+1}^{\frac{1}{\epsilon}-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon-\epsilon_1)n^{1/3}+o(n^{1/3})}.$$

After splitting $\sum_{h=0}^{n-1}$ to $\sum_{q=0}^{\frac{1}{\epsilon}-1}\sum_{r=0}^{\epsilon n-1}$ in (2.26), we obtain the upper bound for $E(\tilde{N}_n^{l_2})^2$ as follows,

$$\begin{aligned} & E(\tilde{N}_n^{l_2})^2 \\ \le\ & E\tilde{N}_n^{l_2} + \sum_{q=0}^{1/\epsilon-1} e^{(-2\lambda_- l_2+\lambda_-(s_q-\delta))n^{1/3}-\sum_{k=0}^{q-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon n^{1/3}-2(\sum_{k=q+1}^{\frac{1}{\epsilon}-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon-\epsilon_1)n^{1/3}+o(n^{1/3})} \\ \le\ & \sum_{q=0}^{1/\epsilon-1} e^{(-2\lambda_- l_2+\lambda_-(s_q-\delta))n^{1/3}-\sum_{k=0}^{q-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon n^{1/3}-2(\sum_{k=q+1}^{\frac{1}{\epsilon}-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon-\epsilon_1)n^{1/3}+o(n^{1/3})}. \end{aligned} \tag{2.27}$$

With the bounds for $E\tilde{N}_n^{l_2}$ (2.24) and $E(\tilde{N}_n^{l_2})^2$ (2.27), we have

$$\begin{aligned} P(\tilde{N}_n^{l_2} > 0) &\ge \frac{(E\tilde{N}_n^{l_2})^2}{E(\tilde{N}_n^{l_2})^2} \ge \frac{1}{\sum_{q=0}^{1/\epsilon-1} e^{(\lambda_-(s_q-\delta)+\sum_{k=0}^{q-1}\frac{\pi^2\sigma_Q^2}{2w_k^2}\epsilon+2\frac{\pi^2\sigma_Q^2}{2w_q^2}\epsilon+2\epsilon_1)n^{1/3}+o(n^{1/3})}} \\ &= \frac{1}{\sum_{q=0}^{1/\epsilon-1} e^{(-\lambda_-\delta+\frac{\pi^2\sigma_Q^2}{w_q^2}\epsilon+2\epsilon_1)n^{1/3}+o(n^{1/3})}} \ge e^{(\lambda_-\delta-\frac{\pi^2\sigma_Q^2}{\delta^2}\epsilon-2\epsilon_1)n^{1/3}+o(n^{1/3})} \\ &= e^{-\epsilon_2 n^{1/3}+o(n^{1/3})}, \end{aligned} \tag{2.28}$$

where $\epsilon_2 := -\lambda_-\delta + \frac{\pi^2\sigma_Q^2}{\delta^2}\epsilon + 2\epsilon_1$, and we use (2.15) in the first equality and $w_q \ge \delta$ (see (2.23)) in the last inequality. We can make ϵ_2 arbitrarily small by first choosing δ small then choosing ϵ and ϵ_1 small. Therefore, we get

$$P(L_n \le l_2 n^{1/3}) \ge P(\tilde{N}_n^{l_2} > 0) \ge e^{-\epsilon_2 n^{1/3}+o(n^{1/3})}. \tag{2.29}$$

2.4.2 A Truncation Argument

In view of the lower bound (2.29), we truncate the tree at level $\lfloor \epsilon_3 n^{1/3}\rfloor = \lfloor 2\epsilon_2 n^{1/3}/\log b\rfloor$ to get $b^{\lfloor \epsilon_3 n^{1/3}\rfloor} \ge e^{2\epsilon_2 n^{1/3}}/b$ independent branching random walks. We take care of the path before and after level $\lfloor \epsilon_3 n^{1/3}\rfloor$ separately.

Define L_n^v similarly as L_n for each branching random walk starting from $v \in \mathbb{D}_{\lfloor \epsilon_3 n^{1/3} \rfloor}$, i.e., letting $z = \lfloor \epsilon_3 n^{1/3} \rfloor$,

$$L_n^v = \min_{u \in \mathbb{D}_{z+n}, u^z = v} \max_{k=z}^{z+n} (S_{u^k} - S_v).$$

Then

$$\begin{aligned} P(L_n^v > l_2 n^{1/3} \text{ for every } v) &= (1 - P(L_n \le l_2 n^{1/3}))^{b^{\lfloor \epsilon_3 n^{1/3} \rfloor}} \\ &\le (1 - e^{-\epsilon_2 n^{1/3} + o(n^{1/3})})^{e^{2\epsilon_2 n^{1/3}}/b} \\ &\le e^{-e^{\epsilon_2 n^{1/3} + o(n^{1/3})}}, \end{aligned} \tag{2.30}$$

when n is large. By the Borel-Cantelli lemma, the above double exponential guarantees that almost surely for all large n, there exists a $v \in \mathbb{D}_{\lfloor \epsilon_3 n^{1/3} \rfloor}$ such that

$$L_n^v \le l_2 n^{1/3}. \tag{2.31}$$

This is an upper bound for the deviation of paths after level $\lfloor \epsilon_3 n^{1/3} \rfloor$. We also need to control the paths before that level, which is a standard large deviation computation. Indeed, for q integer (later, we take $q = \lfloor \epsilon_3 n^{1/3} \rfloor$), set

$$\tilde{Z}_q = \sum_{k=1}^{q} \sum_{v \in \mathbb{D}_k} 1_{\{S_v \ge 2Mq\}}.$$

Recall the definition for M in (2.4). Let Q' be defined by $\frac{dQ'}{dP} = e^{\lambda_+ X_e - \Lambda(\lambda_+)}$. We have

$$\begin{aligned} E\tilde{Z}_q &= \sum_{k=1}^{q} b^k E 1_{\{S_k \ge 2Mq\}} = \sum_{k=1}^{q} b^k E_{Q'} e^{-\lambda_+ S_k + k\Lambda(\lambda_+)} 1_{\{S_k \ge 2Mq\}} \\ &\le \sum_{k=1}^{q} b^k e^{-2\lambda_+ Mq + k\Lambda(\lambda_+)} E_{Q'} 1_{\{S_k \ge 2Mq\}} \\ &\le \sum_{k=1}^{q} b^k e^{-\lambda_+ Mk + k\Lambda(\lambda_+)} e^{-\lambda_+ Mq} = e^{-\lambda_+ Mq + o(q)}, \end{aligned}$$

where, in the last equality, we use the definitions of M and λ_+ (see (2.3) and (2.4)). It follows that

$$P(\tilde{Z}_q \ge 1) \le E\tilde{Z}_q \le e^{-\lambda_+ Mq + o(q)}.$$

Again by the Borel-Cantelli lemma, $\tilde{Z}_q = 0$ for all large q almost surely.

Taking $q = \lfloor \epsilon_3 n^{1/3} \rfloor$ and combining with (2.31), we obtain that

$$L_n \le L_{n+\lfloor \epsilon_3 n^{1/3} \rfloor} \le (l_2 + 2M\epsilon_3) n^{1/3}$$

is true for all large n almost surely. That is,

$$\limsup_{n\to\infty} \frac{L_n}{n^{1/3}} \le l_2 + 2M\epsilon_3 \quad a.s.$$

Since $\epsilon_3 > 0$ and $l_2 > l_0$ are arbitrary, we conclude that

$$\limsup_{n\to\infty} \frac{L_n}{n^{1/3}} \le l_0 \quad a.s. \tag{2.32}$$

Together with (2.21), this completes the proof of Theorem 2. □

2.5 Concluding Remarks

2.5.1 The Curve s(t) of (2.17)

We comment in this subsection on the curve $s(t)$ of (2.17) as a solution to an appropriate variational principle. By the computation in Section 2, $s(t)n^{1/3}$ denotes the minimal possible position for vertices at level tn. However, in Section 3, it is not a priori clear that $s(t)$ will be our best choice. To see why $s(t)$ must indeed be the best choice for the upper bound argument, let us consider a general curve $\phi(t) \le l_2$ as the lower bound for the region. Examining the second moment computation, we need

$$\max_t \left\{ -\phi(t) + \int_0^t \frac{c}{(l_2 - \phi(u))^2} du \right\} \le 0$$

to make the argument work, where c is some constant. Define $w(t) = l_2 - \phi(t) \ge 0$. The above condition is equivalent to

$$l_2 \ge \max_t \{ w(t) + \int_0^t \frac{c}{w(u)^2} du \}.$$

Therefore, the best (smallest) upper bound that we can hope is the result of the following optimization problem

$$\min_{w:(0,1)\to\mathbb{R}_+} \max_t \{ w(t) + \int_0^t \frac{c}{w(u)^2} du \}. \tag{2.33}$$

The solution to this variational problem, denoted by $w^*(\cdot)$, satisfies $s(t) = l_2 - w^*(t)$.

2.5.2 Generalizations

Since the approach in this chapter only uses first and second moment methods, it seems to apply, under natural assumptions, to the situation where the b-ary tree is replaced by a Galton-Watson tree whose offspring distribution possesses high enough exponential moments. This is a question subject to further study. We do not pursue such an extension here.

Acknowledgment

While the result in this chapter was being completed, we learned that as part of their study of RWRE on trees, G. Faraud, Y. Hu and Z. Shi had independently obtained Theorem 2, using a related but slightly different method [32]. In particular, their work handles also the case of Galton–Watson trees. We thank Y. Hu for discussing this problem with one of us (O.Z.) and for providing us with the reference [66], which allowed us to skip tedious details in our original proof.

Chapter 3

Tightness for Maxima of Generalized Branching Random Walks

3.1 Introduction

We study the maxima of a class of generalized branching random walks (GBRW), which are governed by a family of branching rules $\{p_{n,k}\}_{n\geq 0,k\geq 1}$ and displacement laws $\{G_{n,k}\}_{n\geq 0,k\geq 1}$. For this class, we assume that $p_{n,k}$ are nonnegative real numbers such that $\sum_{k=1}^{\infty} p_{n,k} = 1$ and $\sum_{k=1}^{\infty} kp_{n,k} < \infty$ for each $n \geq 0$; $G_{n,k}$ are distribution functions on $\mathbb{R}^k$ for each n and k. The GBRW is defined recursively as follows. In order to show the genealogical relationships between particles, we will use the following hierarchical field to label them. At time 0, a particle $o = \overline{1}$ is located at 0. Suppose that, at time n, $v = \overline{1\alpha_1 \ldots \alpha_n}$ ($\alpha_i \in \mathbb{N}$) is a particle at location S_v. At time $n+1$, v dies and gives birth to $K_v \geq 1$ (random) offspring. We denote the offspring of v at generation $n+1$ by $\{\overline{v1}, \ldots, \overline{vK_v}\}$ and their locations by $\{S_v + X_{v,1}, \ldots, S_v + X_{v,K_v}\}$, respectively. Let $\mathbb{D}$ be the collection of all the particles at any time and $\mathbb{D}_n$ the ones alive at time n. We consider the case where the random vectors $\{(K_v, X_{v,1}, \ldots, X_{v,K_v})\}_{v\in\mathbb{D}}$ indexed by particles are independent and have distributions

$$P\left(K_v = k \middle| v \in \mathbb{D}_n; S_u, u \in \cup_{k=0}^{n} \mathbb{D}_k\right) = p_{n,k} \tag{3.1}$$

and

$$P\left(X_{v,1} \le x_1, \ldots, X_{v,K_v} \le x_{K_v} \middle| v \in \mathbb{D}_n; K_v = k; S_u, u \in \cup_{k=0}^{n} \mathbb{D}_k\right)$$
$$= G_{n,k}(x_1, \ldots, x_k) \qquad \text{for } n = 0, 1, \ldots \text{ and } k = 1, 2, \ldots. \tag{3.2}$$

We are interested in the maximal displacement of particles at time n, i.e., $\mathcal{M}_n = \max_{v \in \mathbb{D}_n} S_v$. Let $F_n(\cdot)$ be the distribution function of $\mathcal{M}_n$ and set $\bar{F}_n(\cdot) = 1 - F_n(\cdot)$. Under some assumptions, we want to prove the tightness of the sequence of re-centered distributions $F_n\left(\cdot - Med(F_n)\right)$, where $Med(F_n)$ is the median of F_n. See Section 3.2 and Section 3.5 for two different sets of assumptions under which tightness can be proved.

From the previous description, our GBRW allows time dependence (through the n parameter) and some local dependence (through the joint distribution $G_{n,k}$). We will review some of the existing literature on tightness and make some comparisons with this paper. Dekking and Host [20] (1991) gave a short proof for tightness of $F_n\left(\cdot - Med(F_n)\right)$ when the offspring displacements are all bounded above by a uniform constant. Using moment arguments, Addario-Berry and Reed [1] (2009) proved that $\mathcal{M}_n - E\mathcal{M}_n$ is exponentially tight when the offspring displacements are i.i.d. and satisfy appropriate large deviation assumptions. By modifying the arguments in [1], [14] and [20], it is possible (but not checked yet) to extend the tightness result to the case when the offspring displacements are unbounded and have local dependence between siblings but not time dependence. See [17] (2010) for using this method to prove the tightness of maxima of modified branching random walks derived from Gaussian free field.

Using a different approach, Bramson and Zeitouni [16] (2009) provided an analytic method to prove tightness of the maximal displacement when, among other situations, the offspring displacement distributions depend on time and satisfy certain tail conditions; they assumed that the offspring displacements are i.i.d. and used a recursion to derive their results. When the joint distribution is locally dependent, this recursion (see (3.3) below) loses some of its nice properties; we will therefore not be able to apply directly this approach. Rather, it needs to be modified to take advantage of some recursion bounds, see (3.6) below.

In order to find a recursion, one needs to look at GBRWs starting from particles at some intermediate time. For any integer m and $v = \overline{1\alpha_1 \ldots \alpha_m} \in \mathbb{D}_m$, the process $\{S_u - S_v | u = \overline{1\alpha_1 \ldots \alpha_m \beta_1 \ldots \beta_k} \in \mathbb{D}_{m+k}, \beta_k \in \mathbb{N}, k = 1, 2, \ldots\}$ is a GBRW governed by branching rules $\{p_{n+m,k}\}_{n\geq 0,k\geq 1}$ and displacement laws $\{G_{n+m,k}\}_{n\geq 0,k\geq 1}$. For $n > m$, the maximal displacement (from S_v) at time $n-m$ is denoted by $\mathcal{M}_n^v$. $\{\mathcal{M}_n^v\}_{v\in\mathbb{D}_m}$ are i.i.d. random variables whose distribution is denoted by $F_n^m(\cdot)$. Again set $\bar{F}_n^m(\cdot) = 1 - F_n^m(\cdot)$. Note that $F_n(\cdot) = F_n^0(\cdot)$, $\bar{F}_n(\cdot) = \bar{F}_n^0(\cdot)$ and $\bar{F}_n^n(\cdot) = 1_{\{x<0\}}(\cdot)$.

One obtains a recursion regarding $F_n^m(\cdot)$ by looking at the first generation of GBRWs starting from particles at time m. For $n > m$,

$$F_n^m(x) = \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \prod_{i=1}^{k} F_n^{m+1}(x-y_i) d^k G_{m,k}(y_1,\ldots,y_k).$$

Following [16], we consider a recursion for $\bar{F}_n^m(\cdot)$. For $n > m$, the above equation is equivalent to

$$\bar{F}_n^m(x) = 1 - \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right) d^k G_{m,k}(y_1,\ldots,y_k). \tag{3.3}$$

Without loss of generality, for any fixed $n,k > 0$, we assume $G_{n,k}$ has identical marginal distributions (all denoted by $g_{n,k}(x)$), i.e.,

$$g_{n,k}(x) = \int_{\mathbb{R}^{k-1}} d^{k-1} G_{n,k}(y_1,\ldots,y_{i-1},x,y_{i+1},\ldots,y_k) \text{ for any } 1 \leq i \leq k. \tag{3.4}$$

Otherwise, if not all the marginal distributions of $G_{n,k}$ are the same, we can consider $\tilde{G}_{n,k}$ defined by

$$\tilde{G}_{n,k}(x_1,\ldots,x_k) = \frac{1}{k!} \sum_{\pi\in\mathcal{P}_k} G_{n,k}(x_{\pi(1)},\ldots,x_{\pi(k)}),$$

where $\mathcal{P}_k$ denotes all the permutations on $\{1,\ldots,k\}$. Then $\tilde{G}_{n,k}$ has identical marginal distributions and recursion (3.3) is the same for $G_{n,k}$ and $\tilde{G}_{n,k}$.

To apply an approach similar to [16], we introduce two functions

$$Q_{1,k}(u) = 1-(1-u)^k \text{ and } Q_{2,k}(u) = ku \text{ for } \quad 0 \leq u \leq 1. \tag{3.5}$$

We will work with the following recursion inequalities derived from (3.3), instead of (3.3) itself.

Lemma 1. *Assuming $\bar{F}_n^m(x)$ satisfies the recursion* (3.3)*, then the following recursion bounds hold, for $n > m$,*

$$\sum_{k=1}^{\infty} p_{m,k} g_{m,k} * Q_{1,k}(\bar{F}_n^{m+1})(x) \leq \bar{F}_n^m(x) \leq \sum_{k=1}^{\infty} p_{m,k} g_{m,k} * Q_{2,k}(\bar{F}_n^{m+1})(x), \tag{3.6}$$

where $*$ is the convolution defined by $f*g(x) = \int_{-\infty}^{\infty} f(x-y)dg(y)$ for any two functions $f(x)$ and $g(x)$ whenever the integral makes sense.

Proof. We begin by proving the upper bound in (3.6). Rewrite (3.3) as

$$\bar{F}_n^m(x) = \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \left(1 - \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right)\right) d^k G_{m,k}(y_1, \ldots, y_k).$$

Using the inequality that $1 - \prod_{i=1}^k (1-x_i) \leq \sum_{i=1}^k x_i$ for $0 \leq x_i \leq 1$ and the fact that $G_{m,k}(\cdot, \ldots, \cdot)$ has the same marginal distributions $g_{m,k}(\cdot)$, one obtains that the above quantity is at most

$$\sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \sum_{i=1}^{k} \bar{F}_n^{m+1}(x-y_i) d^k G_{m,k}(y_1, \ldots, y_k) = \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}} k \bar{F}_n^{m+1}(x-y) dg_{m,k}(y).$$

Together with the definition of $Q_{2,k}$, c.f. (3.5), one obtains the upper bound in (3.6).

We next prove the lower bound in (3.6). Applying Hölder's inequality to (3.3), one obtains that

$$\bar{F}_n^m(x) \geq 1 - \sum_{k=1}^{\infty} p_{m,k} \prod_{i=1}^{k} \left(\int_{\mathbb{R}^k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right)^k d^k G_{m,k}(y_1, \ldots, y_k) \right)^{1/k}.$$

Again, since $G_{m,k}(\cdot, \ldots, \cdot)$ possesses the same marginal distributions $g_{m,k}(\cdot)$, the right side above equals

$$\begin{aligned} & 1 - \sum_{k=1}^{\infty} p_{m,k} \prod_{i=1}^{k} \left(\int_{\mathbb{R}} \left(1 - \bar{F}_n^{m+1}(x-y)\right)^k dg_{m,k}(y) \right)^{1/k} \\ = \; & 1 - \sum_{k=1}^{\infty} p_{m,k} \left(\int_{\mathbb{R}} \left(1 - \bar{F}_n^{m+1}(x-y)\right)^k dg_{m,k}(y) \right) \\ = \; & \sum_{k=1}^{\infty} p_{m,k} \left(\int_{\mathbb{R}} \left(1 - \left(1 - \bar{F}_n^{m+1}(x-y)\right)^k\right) dg_{m,k}(y) \right). \end{aligned}$$

Together with the definition of $Q_{1,k}$, see (3.5), one obtains the lower bound in (3.6). □

3.2 Assumptions and Statement of Result for Bounded Branching.

In this section, we discuss the tightness property in the case where the offspring number is uniformly bounded. To state our result, we need some assumptions both on the branching and displacement laws. We introduce assumptions concerning the branching mechanism.

(B1) $\{p_{n,k}\}_{n\geq 0}$ possess a uniformly bounded support, i.e., there exists an integer $k_0 > 1$ such that $p_{n,k} = 0$ for all n if $k \notin \{1, \ldots, k_0\}$.

(B2) The mean offspring number is uniformly greater than 1 by some fixed constant, i.e., there exists a real number $m_0 > 1$ such that $\inf_n\{\sum_{k=1}^{k_0} kp_{n,k}\} > m_0$.

We introduce the following assumptions on the displacement laws $G_{n,k}$ for those n and k such that $p_{n,k} \neq 0$.

(MT1) For some fixed $\epsilon_0 < \frac{1}{4}\log m_0 \wedge 1$, there exists an x_0 such that $\bar{g}_{n,k}(x_0) \geq 1 - \epsilon_0$ for all n and k, where $\bar{g}_{n,k}(x) = 1 - g_{n,k}(x)$. By shifting, we may and will assume that $x_0 = 0$, that is, $\bar{g}_{n,k}(0) \geq 1 - \epsilon_0$.

(MT2) There exist $a > 0$ and $M_0 > 0$ such that $\bar{g}_{n,k}(x + M) \leq e^{-aM}\bar{g}_{n,k}(x)$ for all n, k and $M > M_0$, $x \geq 0$.

(GT) For any $\eta_1 > 0$, there exists a $B > 0$ such that $G_{n,k}(B, \ldots, B) \geq 1 - \eta_1$ and $G_{n,k}([-B, \infty)^k) \geq 1 - \eta_1$ for all n and k. (With an abuse of notation, $G_{n,k}$ is also used here as a function on measurable sets defined by $G_{n,k}(A) := \int_A d^k G_{n,k}(x_1, \ldots, x_k)$ for $A \subset \mathbb{R}^k$. See (3.2) for the definition of $G_{n,k}$ as a distribution function on $\mathbb{R}^k$.)

Assumptions (MT1) and (MT2) are about the marginal distributions. (MT1) prevents too much mass drifting to $-\infty$, while (MT2) guarantees that the right tails of the marginals decay at least exponentially. (GT) states the tightness of the joint distribution of the increments. Note that (MT1) is implied by (GT), however, we still state (MT1) separately with notations to be used later in the proofs.

Now we are ready to state our main theorem.

Theorem 4. *Under the above assumptions (B1), (B2), (MT2) and (GT), the family of the recentered maxima distributions* $\{F_n(\cdot - Med(F_n))\}_{n\geq 0}$ *is tight.*

Theorem 4 is proved in Section 3.3, with the proofs of some propositions deferred to Section 3.4. With an analysis of a Lyapunov function, we control the right tails of distributions $F_n(\cdot - Med(F_n))$. Then we use assumption (GT) together with the right tail property to control the behavior of left tails of the distributions. Using a similar approach, we can also prove a variation of Theorem 4 under slightly different assumptions in Section 3.5.

3.3 A Lyapunov Function, Main Induction and Proof of Theorem 4

This section follows [16], with some minor revisions, in introducing a Lyapunov function. Namely, for a choice of ϵ_1, b and M (to be determined later), we define the Lyapunov function $L(\cdot)$ as

$$L(u) = \sup_{\{x:u(x)\in(0,\frac{1}{2}]\}} l(u;x), \tag{3.7}$$

where

$$l(u;x) = \log\left(\frac{1}{u(x)}\right) + \log_b\left(1+\epsilon_1 - \frac{u(x-M)}{u(x)}\right)_+ . \tag{3.8}$$

Here $(x)_+ = x \vee 0$, and we take the convention that $\log 0 = -\infty$.

As in [16], the heart of the proof is contained in the following proposition.

Proposition 1. *Under assumptions (B1), (B2), (MT1) and (MT2), there is a choice of* ϵ_1*,* b *and* M *such that* $\sup_{m\le n} L(\bar{F}_n^m) < C$ *for some finite number* $C > 0$.

The proof of Proposition 1 will take the bulk of the chapter, and is detailed in Section 3.4. Before proving it, we discuss its consequences. As in [16, Corollary 2.8], the same proof, using Proposition 1, yields the following

Corollary 2. *Let the assumptions (B1), (B2), (MT1) and (MT2) hold. Then, there exists* δ_1 *such that, for all* n *and* $m \le n$*,*

$$\bar{F}_n^m(x) \le \delta_1 \;\; \textit{implies} \;\; \bar{F}_n^m(x-M) \ge (1+\frac{\epsilon_1}{2})\bar{F}_n^m(x). \tag{3.9}$$

This corollary gives the desired control over the behavior of the right tail of $\bar{F}_n^m(\cdot)$. We next control the left tail. First, one obtains the following pointwise bounds for the integral (3.3).

Lemma 2. *The assumption (GT) implies that, for any $\eta_1 > 0$, there exists a B such that*

$$Q_m(\bar{F}_n^{m+1})(x+B) - \eta_1 \le \bar{F}_n^m(x) \le Q_m(\bar{F}_n^{m+1})(x-B) + \eta_1, \tag{3.10}$$

where $Q_m(u) = \sum_{k=1}^{\infty} p_{m,k}\left(1-(1-u)^k\right)$.

Proof. For any $\eta_1 > 0$, choose the B as in the assumption (GT). The upper bound is obtained by only considering the integral over $(-\infty, B]^k$ in (3.3).

$$\bar{F}_n^m(x) \le 1 - \sum_{k=1}^{\infty} p_{m,k} \int_{(-\infty,B]^k} \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right) d^k G_{m,k}(y_1, \ldots, y_k).$$

By the monotonicity of $\bar{F}_n^m(\cdot)$, the right side is less than

$$1 - \sum_{k=1}^{\infty} p_{m,k} \left(1 - \bar{F}_n^{m+1}(x-B)\right)^k G_{m,k}(B, \cdots, B).$$

For any η_1, choose B as in assumption (GT). Then $G_{m,k}(B, \cdots, B) \ge 1 - \eta_1$, and the above quantity is less than or equal to

$$\begin{aligned} & 1 - \sum_{k=1}^{\infty} p_{m,k} \left(1 - \bar{F}_n^{m+1}(x-B)\right)^k (1-\eta_1) \\ = \;& Q_m(\bar{F}_n^{m+1})(x-B) + \eta_1 \sum_{k=1}^{\infty} p_{m,k} \left(1 - \bar{F}_n^{m+1}(x-B)\right)^k \\ \le \;& Q_m(\bar{F}_n^{m+1})(x-B) + \eta_1, \end{aligned}$$

proving the upper bound in (3.10). To obtain the lower bound, first rewrite (3.3) as

$$\bar{F}_n^m(x) = \sum_{k=1}^{\infty} p_{m,k} \int_{\mathbb{R}^k} \left(1 - \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right)\right) d^k G_{m,k}(y_1, \ldots, y_k).$$

By restricting the above integral to $[-B, \infty)^k$, one has a lower bound on $\bar{F}_n^m$,

$$\bar{F}_n^m(x) \ge \sum_{k=1}^{\infty} p_{m,k} \int_{[-B,\infty)^k} \left(1 - \prod_{i=1}^{k} \left(1 - \bar{F}_n^{m+1}(x-y_i)\right)\right) d^k G_{m,k}(y_1, \ldots, y_k).$$

Since $\bar{F}_n^{m+1}(x)$ is decreasing in x and $G_{m,k}\left([-B,\infty)^k\right) \ge 1 - \eta_1$ as in assumption (GT), one has

$$\begin{aligned}
\bar{F}_n^m(x) &\geq \sum_{k=1}^{\infty} p_{m,k}\left(1-\left(1-\bar{F}_n^{m+1}(x+B)\right)^k\right) G_{m,k}\left([-B,\infty)^k\right) \\
&\geq \sum_{k=1}^{\infty} p_{m,k}\left(1-\left(1-\bar{F}_n^{m+1}(x+B)\right)^k\right)(1-\eta_1) \\
&= Q_m(\bar{F}_n^{m+1})(x+B)-\eta_1\sum_{k=1}^{\infty} p_{m,k}\left(1-\left(1-\bar{F}_n^{m+1}(x+B)\right)^k\right) \\
&\geq Q_m(\bar{F}_n^{m+1})(x+B)-\eta_1,
\end{aligned}$$

proving the lower bound in (3.10) and completing the proof of Lemma 2. □

Lemma 2 almost verifies [16, Assumption 2.4], except that Q_m depends on m. However, with the assumption (B1), Q_m satisfies [16, T1 and T2 in Definition 2.3] uniformly in m. Namely, the family of strictly increasing functions $Q_m : [0,1] \to [0,1]$, with $Q_m(0) = 0$ and $Q_m(1) = 1$, satisfies the following:

(T1') $Q_m(x) > x$ for all $x \in (0,1)$. For any $\delta > 0$, one can choose $c_\delta = 1 + \frac{m_0-1}{k_0}\delta > 1$ such that $Q_m(x) > c_\delta x$ for all $x \leq 1-\delta$ and all m.

(T2') For each $\delta \in (0,1)$, there exists a nonnegative function $g_\delta(\epsilon) \to 0$ as $\epsilon \searrow 0$ (for example, choose $g_\delta(\epsilon) = \frac{(1-(1-\delta)^{k_0})}{k_0\delta}\left(\frac{1+\epsilon}{\delta+\epsilon}\right)^{k_0-1}\epsilon$) such that, for any m, if $x \geq \delta$ and $Q_m\left((1+g_\delta(\epsilon))\,x\right) \leq \frac{1-\delta}{1+\epsilon}$, then $Q_m\left((1+g_\delta(\epsilon))\,x\right) \geq (1+\epsilon)Q_m(x)$.

To check the above two properties, one uses the strict convexity of $1-(1-x)^k$ and its monotonicity in k. Details are omitted here. From the above (T1') and (T2'), one can deduce the following lemma in exactly the same way as in [16, Lemma 2.10].

Lemma 3. *Suppose that* (3.9) *holds for all* $m \leq n$ *under some choice of* $\delta_1, M, \epsilon_1 > 0$. *Also, suppose that assumption (B1) and* (3.10) *hold. For fixed* $\eta_0 \in (0,1)$, *there exist a constant* $\gamma = \gamma(\eta_0) < 1$ *and a continuous function* $f(t) = f_{\eta_0}(t) : [0,1] \to [0,1]$, *with* $f(t) \to_{t\to 0} 0$, *such that, for any* $\epsilon \in (0, \frac{1-\eta_0}{\eta_0})$, $\eta \in [\delta_1, \eta_0]$ *and large enough* $N_1 = N_1(\epsilon)$, *the following holds. If* $M' > M$ *and, for any* $m < n$, $\bar{F}_n^m(x) \geq \delta_1$,

$$\bar{F}_n^m(x-M') \leq (1+\epsilon)\bar{F}_n^m(x) \;\; \textit{and} \;\; \bar{F}_n^m(x-M') \leq \eta,$$

then

$$\bar{F}_n^{m+1}(x+N_1-M') \leq (1+f(\epsilon))\,\bar{F}_n^{m+1}(x-N_1)$$

and

$$\bar{F}_n^{m+1}(x + N_1 - M') \leq \gamma\eta.$$

By iterating, the above lemma gives a connection between the left and right tail behavior. That is, by applying Corollary 2 and Lemma 3 several times as in [16, Proof of Proposition 2.9], the same contrapositive argument proves: for fixed $\eta_0 \in (0, 1)$, there exist an $\hat{\epsilon}_0 = \hat{\epsilon}_0(\eta_0) > 0$, an n_0 and an $\hat{M}$ such that, if $n > n_0$ and $\bar{F}_n^0(x-\hat{M}) \leq \eta_0$, then $\bar{F}_n^0(x - \hat{M}) \geq (1 + \hat{\epsilon}_0)\bar{F}_n^0(x)$. Recalling that $F_n(\cdot) = F_n^0(\cdot)$, this will yield the following tightness proposition.

Proposition 2. *Suppose that* (3.9) *holds for all* $m \leq n$ *under some choice of* $\delta_1, M, \epsilon_1 > 0$. *Also, suppose that assumption (B1) and* (3.10) *hold. Then, the family of recentered maxima distributions* $\{F_n\left(\cdot - Med(F_n)\right)\}_{n\geq 0}$ *is tight.*

We have proved Theorem 4 under the assumption that Proposition 1 is true. Therefore, it remains to show Proposition 1, which we do in the next section.

3.4 Analysis of Lyapunov Function and Proof of Proposition 1

In this section we focus on proving Proposition 1, which is an analog of [16, Theorem 2.7]. The same idea works here: the exponential decay of $g_{n,k}$ will not bring much mass from far away during a single recursion step. However, the exact approach in [16] does not quite apply here. [16] deals separately with the nonlinearity and convolution in a recursion equality. In our case, the recursion (3.3) does not possess such a nice form. Fortunately, we have the recursion inequalities (3.6). These bounds require one to analyze the nonlinearity and convolution together. Throughout this section, all the sums about k are from 1 to k_0 since assumption (B1) is assumed. We begin with some properties of the two functions in (3.5). $Q_{2,k}(u) = ku$ is simple, and the following simple facts about $Q_{1,k}(u)$ will be used later on.

Lemma 4. *There exists a* $c_1 = c_1(k_0) \geq 1$ *such that, for all* $1 \leq k \leq k_0$ *and* $0 \leq u \leq 1$,

$$Q_{1,k}(u) \geq u \tag{3.11}$$

and

$$ku - c_1u^2 \le Q_{1,k}(u) \le ku = Q_{2,k}(u). \tag{3.12}$$

Next, we state a choice of ϵ_1, b and M in the Lyapunov function under which Proposition 1 is true. Throughout, we fix $k_0, m_0, \epsilon_0, M_0$ and a as in assumptions (B1), (B2), (MT1) and (MT2). Next, we choose $0 < \epsilon_1 < \frac{1}{100}$ small, $b > 1$ close to 1, $M > 100$ big and an auxiliary variable $0 < \kappa < \frac{1}{100}$ small (used later to control the flatness change) such that the following restrictions hold.

$$M > 4M_0 \text{ and } e^{-aM/2} \le (4k_0)^4 e^{-aM/2} \le \frac{1}{100}; \tag{3.13}$$

$$\frac{8(2k_0)^{5/2}\epsilon_1^{1/2\log b - 3/2}}{(1-\epsilon_0)\kappa^{3/2}} < \frac{1}{2c_1}; \tag{3.14}$$

$$c_1\frac{1+\epsilon_1}{1-\epsilon_0}\epsilon_1^{1/\log b} \le \sum_{k=1}^{k_0} kp_{n,k} - m_0 \text{ for all } n; \tag{3.15}$$

$$\frac{\log m_0}{2} \ge 2(\epsilon_1 + \epsilon_0) + \frac{6\kappa}{\log b}; \tag{3.16}$$

$$\frac{aM}{16\log b} \ge 2\left(\epsilon_1 + \epsilon_0 + \log(4k_0)\right) - \frac{\log\kappa}{\log b}; \tag{3.17}$$

$$\frac{a}{16\log b} \ge \frac{2\log(4k_0)}{M}. \tag{3.18}$$

The above conditions are compatible. In fact, thinking of κ as $\beta\log b$, one can choose ϵ_1 and β small enough so that (3.16) holds due to the choice of ϵ_0 in assumption (MT1), then one chooses a b close enough to 1 so that (3.14), (3.15) and (3.18) hold due to the choice of m_0 as in assumption (B2), and finally one chooses M large enough so that (3.13) and (3.17) hold.

With the choice of the above ϵ_1, b, M and κ, we can now prove Proposition 1.

Proof of Proposition 1. Choose $C = \log 2$. The conclusion $\sup_{m\le n} L(\bar{F}_n^m) \le C$ will follow from the claim:

$$L(\bar{F}_n^m) > C \text{ implies that } L(\bar{F}_n^{m+1}) > C \text{ for any } m < n. \tag{3.19}$$

Suppose the conclusion is violated, then $L(\bar{F}_n^m) > C$ for some $m \leq n$. Iterating the claim $n-m$ times, one gets $L(\bar{F}_n^n) > C$. However, $L(\bar{F}_n^n) = -\infty$ because $\bar{F}_n^n(x) = 1_{\{x<0\}}(x)$. This contradiction proves proposition 1, assuming claim (3.19). $\square$

The claim (3.19) follows from the following proposition because of (3.6).

Proposition 3. *Suppose that two non-increasing cadlag functions $u, v : \mathbb{R} \to [0,1]$ satisfy*

$$\sum_{k=1}^{k_0} p_k g_k * Q_{1,k}(u)(x) \leq v(x) \leq \sum_{k=1}^{k_0} p_k g_k * Q_{2,k}(u)(x), \tag{3.20}$$

where p_k and g_k satisfy the assumptions in Section 3.2 as $p_{n,k}$ and $g_{n,k}$, and $Q_{1,k}$ and $Q_{2,k}$ satisfy Lemma 4. Then

$$L(v) > C \ \textit{ implies that } \ L(u) > C. \tag{3.21}$$

In order to prove Proposition 3, a few observations, notation and lemmas are needed. Starting from $L(v) > C$, one obtains, by definition (3.7) of the Lyapunov function, that there exists an $x_1 \in \mathbb{R}$ such that

$$v(x_1) \leq \frac{1}{2} \ \text{ and } \ l(v; x_1) \geq \max\{C, L(v) - \frac{1}{4}\log m_0\}. \tag{3.22}$$

By definition (3.8) of $l(v;x)$, one obtains that v is small and flat at x_1 in the following sense:

$$1 + \epsilon := \frac{v(x_2)}{v(x_1)} < 1 + \epsilon_1 \tag{3.23}$$

and

$$f_0 := v(x_1) < (\epsilon_1 - \epsilon)^{1/\log b} e^{-C} < \frac{1}{2}, \tag{3.24}$$

where $x_2 := x_1 - M$. Using the bounds (3.20) and (3.23), one gets that

$$\sum_{k=1}^{k_0} p_k g_k * Q_{1,k}(u)(x_2) \leq (1+\epsilon) \sum_{k=1}^{k_0} p_k g_k * Q_{2,k}(u)(x_1), \tag{3.25}$$

from which we will search for a flat segment in $u(x)$ where $u(x)$ is also small.

To control the value of $u(x)$, we derive here some preliminary estimates of $u(x)$ at x_1 and x_2, which will be used later to control the value of $u(x)$ at other places. For

$i = 1, 2$, first applying the Chebyshev inequality and then applying (3.20) and the fact $\bar{g}_k(0) \geq 1 - \epsilon_0$ from assumption (MT1), one gets

$$\begin{aligned}\sum_{k=1}^{k_0} p_k Q_{1,k}(u)(x_i) &\leq \sum_{k=1}^{k_0} p_k \frac{1}{\bar{g}_k(0)} \int_{\mathbb{R}} Q_{1,k}(u)(x_i - y) dg_k(y) \\ &\leq \frac{1}{1-\epsilon_0} v(x_i). \end{aligned} \tag{3.26}$$

This, together with the lower bound (3.11) on $Q_{1,k}$, the definition (3.24) of f_0 and the definition (3.23) of ϵ, implies that

$$u(x_1) \leq \frac{f_0}{1-\epsilon_0}, \tag{3.27}$$

and

$$u(x_2) \leq \frac{1+\epsilon}{1-\epsilon_0} f_0. \tag{3.28}$$

A finer estimate of $u(x_2)$ can be obtained and will be needed. First, using (3.26) and the lower bound (3.12) on $Q_{1,k}$, one gets

$$\left(\sum_{k=1}^{k_0} k p_k - c_1 u(x_2) \right) u(x_2) \leq \frac{1+\epsilon}{1-\epsilon_0} f_0.$$

By combining the first estimate (3.28) of $u(x_2)$, the bound (3.24) on f_0 and the restriction (3.15), the coefficient multiplying $u(x_2)$ on the left side of the last inequality is at least

$$\begin{aligned} \sum_{k=1}^{k_0} k p_k - c_1 u(x_2) &\geq \sum_{k=1}^{k_0} k p_k - c_1 \frac{1+\epsilon}{1-\epsilon_0} f_0 \\ &\geq \sum_{k=1}^{k_0} k p_k - c_1 \frac{1+\epsilon_1}{1-\epsilon_0} (\epsilon_1 - \epsilon)^{1/\log b} e^{-C} \\ &\geq \sum_{k=1}^{k_0} k p_k - c_1 \frac{1+\epsilon_1}{1-\epsilon_0} \epsilon_1^{1/\log b} \geq m_0. \end{aligned}$$

Therefore, we conclude that

$$u(x_2) \leq \frac{1+\epsilon}{m_0(1-\epsilon_0)} f_0 = \frac{1+\epsilon}{m_0(1-\epsilon_0)} v(x_1). \tag{3.29}$$

To control the flatness of $u(x)$, we define some more auxiliary variables and then state some lemmas. The constants $\delta = \kappa(\epsilon_1 - \epsilon)$, $\epsilon' = \epsilon + \delta$, $\epsilon'' = \epsilon + 2\delta$ and $\epsilon^{(3)} = \epsilon + 3\delta$

are defined to monitor the flatness change. Note that $\epsilon, \epsilon', \epsilon'', \epsilon^{(3)} < \epsilon_1$ because $\kappa < \frac{1}{100}$. We somewhat simplify the argument in [16]. Set

$$y_0 = \frac{1}{a} \log \frac{2k_0}{\delta f_0}, \tag{3.30}$$

$$q = \inf\{y \geq M/2 : u(x_2 - y) > (4k_0)^2 u(x_1 - y)\} \tag{3.31}$$

and

$$r = y_0 \wedge \begin{cases} q, & \text{if } u(x_2 - q)^- \geq (4k_0)u(x_1 - (q + \frac{M}{2})), \\ q - \frac{M}{2}, & \text{otherwise,} \end{cases} \tag{3.32}$$

where $f(x)^- := \lim_{y \to x-} f(y)$ is the left limit of f at x. Intuitively, q is used to denote the first nonflatness place to the left of x_1. When $r < y_0$, r is used to denote a nonflat interval, namely, it is easy to check that

$$u(x_2 - y) \geq (4k_0)u(x_1 - y) \text{ for all } y \in (r, r + M/2]. \tag{3.33}$$

We can now state the following sequence of lemmas, whose proofs will be discussed in the next subsection. The convention of

$$\int_a^b f(x)dg(x) = \int_{(a,b]} f(x)dg(x)$$

for $a, b \in \mathbb{R}$ will be made throughout the rest of the chapter.

Lemma 5. *Assume that* (3.24) *and* (3.25) *hold. Then,*

$$\sum_{k=1}^{k_0} p_k \int_{-\infty}^{r} Q_{1,k}(u)(x_2 - y)dg_k(y) \leq (1 + \epsilon') \sum_{k=1}^{k_0} p_k \int_{-\infty}^{r} Q_{2,k}(u)(x_1 - y)dg_k(y). \tag{3.34}$$

Lemma 6. *If* (3.24) *and* (3.34) *are satisfied, then there exist some* $1 \leq k \leq k_0$ *and* r' *such that*

$$\int_{-\infty}^{r'} u(x_2 - y)dg_k(y) \leq (1 + \epsilon'') \int_{-\infty}^{r'} u(x_1 - y)dg_k(y), \tag{3.35}$$

where $r' = r$ *when* $r' > M$.

Lemma 7. *Suppose* (3.35) *holds. Then either*

(a) $u(x_2 - y_1) \leq (1 + \epsilon^{(3)})u(x_1 - y_1)$ *for some* $y_1 \leq r' \wedge M$, *or*

(b) $u(x_2 - y_1) \le (1 + \epsilon'' - \delta e^{ay_1/8})u(x_1 - y_1)$ *for some* $y_1 \in (M, r]$.

Lemma 6 and Lemma 7 are analogs of [16, Lemma 3.5, Proposition 3.2], respectively. Equipped with lemma 7, we are ready to prove Proposition 3.

Proof of Proposition 3 assuming Lemma 7. We will compare $L(u)$ and $L(v)$ using (3.29) and Lemma 7. As Lemma 7 suggests, two different cases will be discussed separately.

Case (a). Assume $u(x_2 - y_1) \le (1 + \epsilon^{(3)})u(x_1 - y_1)$ for some $y_1 \le r' \wedge M$. Then, (3.29) implies that

$$u(x_1 - y_1) \le u(x_2) \le \frac{1+\epsilon}{m_0(1-\epsilon_0)} v(x_1).$$

Therefore, it follows from the definition (3.8) of $l(u; x)$ that

$$\begin{aligned} l(u, x_1 - y_1) - l(v, x_1) &\ge \log \frac{v(x_1)}{u(x_1 - y_1)} + \log_b \frac{\epsilon_1 - \epsilon^{(3)}}{\epsilon_1 - \epsilon} \\ &\ge \log \frac{m_0(1-\epsilon_0)}{1+\epsilon} + \log_b(1 - 3\kappa) \\ &\ge \log m_0 - 2(\epsilon_1 + \epsilon_0) - \frac{6\kappa}{\log b} \ge \frac{\log m_0}{2}, \end{aligned}$$

where (3.16) guarantees the last inequality.

Case (b). Assume $u(x_2 - y_1) \le (1 + \epsilon'' - \delta e^{ay_1/8})u(x_1 - y_1)$ for some $y_1 \in (M, r]$. Then, the definition (3.32) of r and (3.29) imply that

$$u(x_1 - y_1) \le (4k_0)^{2y_1/M+2} u(x_1 - M/2) \le (4k_0)^{2y_1/M+2} \frac{1+\epsilon}{m_0(1-\epsilon_0)} v(x_1).$$

Therefore, it follows that

$$\begin{aligned} l(u, x_1 - y_1) - l(v, x_1) &= \log \frac{v(x_1)}{u(x_1 - y_1)} + \log_b \frac{\epsilon_1 - \epsilon'' + \delta e^{ay_1/8}}{\epsilon_1 - \epsilon} \\ &\ge \log \frac{m_0(1-\epsilon_0)}{(1+\epsilon)(4k_0)^{2y_1/M+2}} + \log_b(1 - 2\kappa + \kappa e^{ay_1/8}) \\ &\ge \log m_0 - 2(\epsilon_0 + \epsilon_1) - \frac{2\log(4k_0)}{M} y_1 - 2\log(4k_0) \\ &\quad + \frac{\log \kappa + ay_1/8}{\log b} . \end{aligned}$$

Rewrite the last term $\frac{ay_1}{8\log b}$ as $\frac{ay_1}{16\log b} + \frac{ay_1}{16\log b}$, use $y_1 \ge M$ in one summand and deduce that the above quantity is at least

$$\log m_0 - 2(\epsilon_0 + \epsilon_1 + \log(4k_0)) + \frac{\log \kappa}{\log b} + \frac{aM}{16\log b} + y_1\Big(\frac{a}{16\log b} - \frac{2\log(4k_0)}{M}\Big) \ge \frac{1}{2}\log m_0,$$

where (3.18) and (3.17) guarantee the last inequality.

To wrap the argument up, both cases imply, by (3.8), (3.22) and $C = \log 2$,

$$\log \frac{1}{u(x_1 - y_1)} \geq l(u, x_1 - y_1) \geq C + \frac{1}{2}\log m_0 \geq -\log \frac{1}{2},$$

which implies that $u(x_1 - y_1) \leq \frac{1}{2}$. Therefore, by the definition (3.7) of $L(u)$ and (3.22) again,

$$L(u) \geq l(u, x_1 - y_1) \geq l(v, x_1) + \frac{1}{2}\log m_0 \geq L(v) + \frac{1}{4}\log m_0 \geq L(v),$$

from which (3.21) follows. Thus, the proof of Proposition 3 is complete. □

3.4.1 Proof of Lemmas

With the assumption (MT2), the proof of [16, Proposition 3.2] carries over (with some change of notation) to the proof of Lemma 7 assuming Lemma 6. So we only need to prove Lemma 6 and 5. (For completeness, we give the proof of Lemma 7 in the appendix of this chapter.) The proof of Lemma 6 will be presented first, and then the proof of Lemma 5.

Proof of Lemma 6. When $q > M/2$, we have $u(x_2 - y) \leq (4k_0)^2 u(x_1 - y)$ for $y \in [M/2, q]$. Thus, one obtains that, for any $y \leq r \leq q$,

$$u(x_2 - y) \leq u(x_2 - r) \leq (4k_0)^{2r/M+2} u(x_2).$$

Since $r \leq y_0 = \frac{1}{a}\log\frac{2k_0}{\delta f_0}$, one has, using (3.28), that the above is at most

$$(4k_0)^{2y_0/M+2}\frac{1+\epsilon}{1-\epsilon_0} f_0 < \frac{2(4k_0)^2}{1-\epsilon_0}(4k_0)^{\frac{2}{aM}\log\frac{2k_0}{\delta f_0}} f_0 = \frac{2(4k_0)^2}{1-\epsilon_0}(\frac{2k_0}{\delta f_0})^{\frac{2}{aM}\log(4k_0)} f_0.$$

Note that $\frac{2}{aM}\log(4k_0) < \frac{1}{2}$ from (3.13). Applying the bound (3.24) on f_0, the above quantity is at most

$$\frac{2(4k_0)^2}{1-\epsilon_0}\frac{\sqrt{2k_0} f_0^{1/2}}{\delta^{1/2}} = \frac{8(2k_0)^{5/2} f_0^{1/2}}{(1-\epsilon_0)\delta^{3/2}}\delta < \frac{8(2k_0)^{5/2}(\epsilon_1 - \epsilon)^{1/2\log b - 3/2}}{(1-\epsilon_0)\kappa^{3/2}}\delta.$$

Therefore, it follows from (3.14) that

$$u(x_2 - y) \leq \frac{1}{2c_1}\delta \text{ for any } y \leq r. \tag{3.36}$$

This, combined with (3.12), implies that, for any $1 \le k \le k_0$ and $y \le r_1$,

$$\begin{aligned} & Q_{1,k}(u)(x_2-y) \ge ku(x_2-y) - c_1\left(u(x_2-y)\right)^2 \\ = \; & ku(x_2-y)(1-\frac{c_1}{k}u(x_2-y)) \ge ku(x_2-y)(1-\frac{1}{2}\delta). \end{aligned}$$

Applying the above bound and the definition (3.5) of $Q_{2,k}(u)$ in the first inequality, and (3.34) in the second, one has

$$\begin{aligned} & \frac{\sum_{k=1}^{k_0} kp_k \int_{-\infty}^{r} u(x_2-y)dg_k(y)}{\sum_{k=1}^{k_0} kp_k \int_{-\infty}^{r} u(x_1-y)dg_k(y)} \\ \le \; & \frac{1}{1-\frac{1}{2}\delta} \frac{\sum_{k=1}^{k_0} p_k \int_{-\infty}^{r} Q_{1,k}(u)(x_2-y)dG_k(y)}{\sum_{k=1}^{k_0} p_k \int_{-\infty}^{r} Q_{2,k}(u)(x_1-y)dG_k(y)} \\ \le \; & \frac{1+\epsilon'}{1-\frac{1}{2}\delta} \le 1+\epsilon''. \end{aligned} \tag{3.37}$$

If the conclusion of the lemma does not hold, i.e., for all $1 \le k \le k_0$,

$$\int_{-\infty}^{r} u(x_2-y)dg_k(y) > (1+\epsilon'') \int_{-\infty}^{r} u(x_1-y)dg_k(y),$$

one obtains a contradiction to (3.37). This completes the proof of Lemma 6 in case $q > M/2$.

When $q = M/2$ and $u(x_2-M/2) \le 4k_0u(x_2)$, with (3.28), one still has, for $y \le r \le q$,

$$u(x_2-y) \le u(x_2-r) \le 4k_0u(x_2) \le \frac{8k_0f_0}{(1-\epsilon_0)\delta}\delta. \tag{3.38}$$

Using the bound (3.24) on f_0 and (3.14), the above is at most

$$\frac{8k_0(\epsilon_1-\epsilon)^{1/\log b-1}}{(1-\epsilon_0)\kappa}\delta \le \frac{1}{2c_1}\delta.$$

Thus, (3.36) holds. Repeating the argument below (3.36), one gets Lemma 6 in this case.

When $q = M/2$ but $u(x_2-M/2) > (4k_0)u(x_2)$, we truncate (3.34) before transforming this case to the previous case. Define

$$r' = \inf\{y \ge 0 : u(x_2-y) > 4k_0u(x_2)\}.$$

Then $0 \leq r' < M/2$ and $u(x_2 - r') \leq 4k_0 u(x_2)$. By monotonicity of u, $u(x_2 - y) \geq 4k_0 u(x_1 - y)$ for $y \in (r', r]$. Therefore, for $1 \leq k \leq k_0$,

$$\begin{aligned} & \int_{r'}^{r} Q_{1,k}(u)(x_2 - y) dg_k(y) - (1+\epsilon') \int_{r'}^{r} Q_{2,k}(u)(x_1 - y) dg_k(y) \\ \geq & \int_{r'}^{r} u(x_2 - y) dg_k(y) - 2 \int_{r'}^{r} k_0 u(x_1 - y) dg_k(y) \\ = & \int_{r'}^{r} \left(u(x_2 - y) - 2k_0 u(x_1 - y) \right) dg_k(y) \geq 0, \end{aligned}$$

which, together with (3.34), yields the truncated inequality

$$\sum_{k=1}^{k_0} p_k \int_{-\infty}^{r'} Q_{1,k}(u)(x_2 - y) dg_k(y) \leq (1+\epsilon') \sum_{k=1}^{k_0} p_k \int_{-\infty}^{r'} Q_{2,k}(u)(x_1 - y) dg_k(y).$$

This is a analog of (3.34) with r replaced by r', and $u(x_2 - r') \leq 4k_0 u(x_2)$. Replacing r by r' in the argument starting from (3.38), one concludes the proof of Lemma 6 in all cases. □

Proof of Lemma 5. The purpose of this lemma is to justify the flatness of the truncated integral. That is, we want to prove that mass from faraway does not affect the value of the integral in a significant way. This is almost guaranteed by the exponential decay of $g_{n,k}(\cdot)$. However, we need to control the difference between $Q_{1,k}(u)(x_2 - y)$ and $Q_{2,k}(u)(x_1 - y)$, using the lower bound (3.11) on $Q_{1,k}(u)$ and the definition (3.5) of $Q_{2,k}(u)$. Two different cases will be presented separately.

Case (i). When $r < y_0$, (3.33) holds. Because of (3.25) and $\epsilon < \epsilon'$, (3.34) will follow from

$$\int_{r}^{\infty} Q_{1,k}(u)(x_2 - y) dg_k(y) - (1+\epsilon') \int_{r}^{\infty} Q_{2,k}(u)(x_1 - y) dg_k(y) \geq 0. \tag{3.39}$$

To prove (3.39), because of (3.11), it suffices to show that

$$\int_{r}^{\infty} u(x_2 - y) dg_k(y) - 2k_0 \int_{r}^{\infty} u(x_1 - y) dg_k(y) \geq 0. \tag{3.39$'$}$$

We break the left side into 3 pieces. First, by (3.33),

$$\begin{aligned} & \frac{1}{2} \int_{r}^{r+M/2} u(x_2 - y) dg_k(y) - 2k_0 \int_{r}^{r+M/2} u(x_1 - y) dg_k(y) \\ = & \int_{r}^{r+M/2} \left(\frac{1}{2} u(x_2 - y) - 2k_0 u(x_1 - y) \right) dg_k(y) \geq 0. \end{aligned} \tag{3.40}$$

Second, because of assumption (MT2) (rapid decay of $\bar{g}_k(\cdot)$) and (3.13), one has

$$\begin{aligned} & \frac{1}{2}\int_r^{r+M/2} u(x_2-y)dg_k(y) - 2k_0\int_{r+M/2}^{r+M} u(x_1-y)dg_k(y) \\ \geq\ & \frac{1}{4}u(x_2-r)\bar{g}_k(r) - 2k_0 u(x_2-r)\bar{g}_k(r+M/2) \qquad (3.41) \\ \geq\ & (\frac{1}{4} - 2k_0 e^{-aM/2})u(x_2-r)\bar{g}_k(r) \geq 0. \end{aligned}$$

Third, again because of assumption (MT2) (rapid decay of $\bar{g}_k(\cdot)$) and (3.13), one has

$$\begin{aligned} & \int_{r+M/2}^{\infty} u(x_2-y)dg_k(y) - 2k_0\int_{r+M}^{\infty} u(x_1-y)dg_k(y) \\ \geq\ & \int_{r+M/2}^{\infty} u(x_2-(y-M/2))dg_k(y) - 2k_0\int_r^{\infty} u(x_2-y)dg_k(y+M) \\ =\ & \int_r^{\infty} u(x_2-y)dg_k(y+M/2) - 2k_0\int_r^{\infty} u(x_2-y)dg_k(y+M) \qquad (3.42) \\ \geq\ & (1-2k_0e^{-aM/2})\int_r^{\infty} u(x_2-y)dg_k(y+M/2) \geq 0 \end{aligned}$$

Summing (3.40), (3.41) and (3.42), one gets (3.39'). Thus, (3.39) is verified in case (i), and (3.34) holds.

Case (ii). When $r = y_0$, (3.33) may not be true. However, the difference between the two sides of (3.34) is

$$\begin{aligned} & \sum_{k=1}^{k_0} p_k \int_{-\infty}^{r} Q_{1,k}(u)(x_2-y)dg_k(y) - (1+\epsilon')\sum_{k=0}^{k_0} p_k\int_{-\infty}^{r} Q_{2,k}(u)(x_1-y)dg_k(y) \\ \leq\ & \sum_{k=1}^{k_0} p_k g_k * Q_{1,k}(u)(x_2) - (1+\epsilon')\sum_{k=1}^{k_0} p_k\int_{-\infty}^{r} Q_{2,k}(u)(x_1-y)dg_k(y). \end{aligned}$$

Recall that $\epsilon' = \epsilon + \delta$. (3.25) implies that the above quantity is less than or equal to

$$\begin{aligned} & (1+\epsilon)\sum_{k=1}^{k_0} p_k g_k * Q_{2,k}(u)(x_1) - (1+\epsilon')\sum_{k=1}^{k_0} p_k\int_{-\infty}^{r} Q_{2,k}(u)(x_1-y)dg_k(y) \\ =\ & (1+\epsilon')\sum_{k=1}^{k_0} p_k\int_r^{\infty} Q_{2,k}(u)(x_1-y)dg_k(y) - \delta\sum_{k=1}^{k_0} p_k g_k * Q_{2,k}(u)(x_1). \end{aligned}$$

Since $r = y_0 = \frac{1}{a}\log\frac{2k_0}{\delta f_0}$, the assumption (MT2) implies $\bar{g}_k(r) \leq e^{-ay_0} = \frac{\delta f_0}{2k_0}$. $Q_{2,k}(u) \leq k_0$, (3.20) and (3.24) yield that the above quantity again does not exceed

$$(1+\epsilon')k_0\frac{\delta f_0}{2k_0} - \delta f_0 \leq 0.$$

So (3.34) is proved in case (ii). This completes the proof of the lemma. □

3.5 Tightness for Identical Marginals

In this section, we discuss the tightness problem in the case when all the marginal distributions at the same level are the same, i.e., $g_{n,k}(\cdot) = g_n(\cdot)$ does not depend on the number of offspring. Compared with the assumptions made in Section 3.2, we relax the bounded support assumption (B1) on $p_{n,k}$s, at the price of a uniform marginal assumption on $G_{n,k}$ (see (MT0') below). Namely, we assume

(B1') There exist positive real numbers m_0 and m_1 such that $\inf_n\{\sum_{k=1}^{\infty} kp_{n,k}\} > m_0 > 1$ and $\sup_n \sum_{k=1}^{\infty} k^2 p_{n,k} < m_1$.

(MT0') $g_{n,k}(\cdot) = g_n(\cdot)$ for all $k \geq 1$.

(MT1') For some fixed $\epsilon_0 < \frac{1}{4}\log m_0 \wedge 1$, there exists an x_0 such that $\bar{g}_n(x_0) \geq 1 - \epsilon_0$ for all n, where $\bar{g}_n(x) = 1 - g_n(x)$. By shifting, we will assume that $x_0 = 0$, that is, $\bar{g}_n(0) \geq 1 - \epsilon_0$.

(MT2') There exist $a > 0$ and $M_0 > 0$ such that $\bar{g}_n(x + M) \leq e^{-aM}\bar{g}_n(x)$ for all n and $M > M_0$, $x \geq 0$.

(GT') For any $\eta_1 > 0$, there exists a $B > 0$ such that $G_{n,k}(B, \ldots, B) \geq 1 - \eta_1$ and $\bar{g}_n(-B) \geq 1 - \eta_1$ for all n and k.

Then we still have the following tightness result.

Theorem 5. *Under assumptions (B1'), (MT0'), (MT1'), (MT2') and (GT'), the family of the recentered maxima distribution $\{F_n(\cdot - Med(F_n))\}$ is tight.*

Since the proof is similar to the proof of Theorem 4, we only give a sketch. The argument is based on the following recursion inequality, another form of (3.6) under the assumption (MT0'),

$$g_m * \left(\sum_{k=1}^{\infty} p_{m,k} Q_{1,k}(\bar{F}_n^{m+1})\right)(x) \leq \bar{F}_n^m(x) \leq g_m * \left(\sum_{k=1}^{\infty} p_{m,k} Q_{2,k}(\bar{F}_n^{m+1})\right)(x), \quad (3.43)$$

where $Q_{1,k}$ and $Q_{2,k}$ are defined as (3.5). Set

$$Q_{m,(1)}(u) = \sum_{k=1}^{\infty} p_{m,k} Q_{1,k}(u), \tag{3.44}$$

and

$$Q_{m,(2)}(u) = \sum_{k=1}^{\infty} p_{m,k} Q_{2,k}(u). \tag{3.45}$$

Although the difference between $Q_{1,k}$ and $Q_{2,k}$ increases as k increases, the weighted functions $Q_{m,(1)}$ and $Q_{m,(2)}$ still behave nicely and possess an analog of Lemma 4.

Lemma 8. *Letting $Q_{m,(1)}$ and $Q_{m,(2)}$ be defined as in* (3.44) *and* (3.45)*, respectively, then it follows from assumption (B') that*

$$Q_{m,(1)}(u) > u, \tag{3.46}$$

and

$$Q_{m,(2)}(u) - c_2 u^2 \leq Q_{m,(1)}(u) \leq Q_{m,(2)}(u) \leq \sqrt{m_1} u. \tag{3.47}$$

Lemma 5 relies on the facts that $Q_{1,k}(u) \geq u$ and $Q_{2,k}(u) \leq k_0 u$, and Lemma 6 relies on the fact that $Q_{1,k}(u) \geq Q_{2,k}(u) - c_1 u^2$. Therefore, with a modification of q and r, we can prove analogs of those two lemmas due to the bounds in Lemma 8. An analog of Proposition 3 then follows. Proposition 1 and Corollary 2 hold under the new assumptions in this section.

Assumption (GT') plays a role as (GT) in connecting the left and right tail behavior. Specifically, it guarantees Lemma 2, Lemma 3 and Proposition 2 under the new settings. Theorem 5 follows immediately as Theorem 4.

Appendix

This proof of Lemma 7 is almost the same as the proof of [16, Proposition 3.2] except for different notations, so we give it here for completeness.

proof of Lemma 7 assuming Lemma 6. The lemma will be proved by contradiction. Assume that neither (a) nor (b) in lemma 7 holds, i.e.,

$$u(x_2 - y) > (1 + \epsilon^{(3)}) u(x_1 - y) \text{ for all } y \leq r' \wedge M, \tag{$\bar{a}$}$$

and

$$u(x_2 - y) > (1 + \epsilon'' - \delta e^{ay/8}) u(x_1 - y) \text{ for all } y \in (M, r']. \tag{$\bar{b}$}$$

If $r' \le M$, then only ($\bar{a}$) holds and it implies that

$$\int_{-\infty}^{r'} u(x_2 - y) dg_k(y) > (1 + \epsilon^{(3)}) \int_{-\infty}^{r'} u(x_1 - y) dg_k(y).$$

Since $\epsilon^{(3)} = \epsilon'' + \delta > \epsilon''$, this is a contradiction to (3.34). So we are done.

If $r' > M$, then $r' = r$ and ($\bar{a}$) and ($\bar{b}$) imply that

$$\int_{-\infty}^{M} u(x_2 - y) dg_k(y) > (1 + \epsilon^{(3)}) \int_{-\infty}^{M} u(x_1 - y) dg_k(y)$$

and

$$\int_{M}^{r} u(x_2 - y) dg_k(y) > \int_{M}^{r} (1 + \epsilon'' - \delta e^{ay/8}) u(x_1 - y) dg_k(y).$$

Summing the above two inequality, one gets that

$$\begin{aligned}\int_{-\infty}^{r} u(x_2 - y) dg_k(y) &> (1 + \epsilon'') \int_{-\infty}^{r} u(x_1 - y) dg_k(y) \\ &+\delta \left[\int_{-\infty}^{M} u(x_1 - y) dg_k(y) - \int_{M}^{r} e^{ay/8} u(x_1 - y) dg_k(y) \right].\end{aligned}$$

We claim that

$$\int_{-\infty}^{M} u(x_1 - y) dg_k(y) - \int_{M}^{r} e^{ay/8} u(x_1 - y) dg_k(y) \ge 0, \tag{3.48}$$

which will imply a contradiction of (3.34) and complete the proof.

It thus remains to prove the claim (3.48). The second integral on the left side of (3.48) can be written as

$$\int_{M}^{r} e^{ay/8} u(x_1 - y) dg_k(y) = \sum_{l=1}^{\infty} \int_{lM}^{lM+M} e^{ay/8} u(x_1 - y) 1_{\{y \le r\}} dg_k(y).$$

Since $q \ge r > M/2$, one has $u(x_2 - y) \le (4k_0)^2 u(x_1 - y)$ for all $y \in [M/2, r]$, the last quantity in the above display is less than or equal to

$$\begin{aligned}&\sum_{l=1}^{\infty} \int_{lM}^{lM+M} e^{alM/8 + aM/8} (4k_0)^{2l+2} u(x_1 - M/2) dg_k(y) \\ \le\ &\sum_{l=1}^{\infty} e^{alM/8 + aM/8} (4k_0)^{2l+2} u(x_1 - M/2) \bar{g}_k(lM).\end{aligned}$$

Using Assumption (MT2) and (3.13), the above is again less than or equal to

$$\begin{aligned} & \sum_{l=1}^{\infty} e^{alM/8+aM/8}(4k_0)^{2l+2}u(x_1-M/2)e^{-alM+aM/2}\bar{g}_k(M/2) \\ \leq \quad & \frac{1}{4}u(x_1-M/2)\bar{g}_k(M/2). \end{aligned}$$

But the last term does not exceed

$$\int_{M/2}^{M} u(x_1-y)dg_k(y) \leq \int_{-\infty}^{M} u(x_1-y)dg_k(y).$$

So the proof of (3.48) is complete. We are done with proving Lemma (7). □

Chapter 4

Branching Random Walks in Time Inhomogeneous Environments

4.1 Introduction

Branching random walks and their maxima have been studied mostly in space-time homogeneous environments (deterministic or random). For work on the deterministic homogeneous case of relevance to our study we refer to [14] and the recent [1] and [2]. For the random environment case, a sample of relevant papers is [36, 38, 46, 49, 58, 62, 69]. As is well documented in these references, under reasonable hypotheses, in the homogeneous case the maximum grows linearly, with a logarithmic correction, and is tight around its median.

Branching random walks are also studied under some space inhomogeneous environments. A sample of those papers are [10, 23, 28, 37, 43, 45, 53].

Recently, Bramson and Zeitouni [16] and Fang [30] showed that the maxima of branching random walks, recentered around their median, are still tight in time inhomogeneous environments satisfying certain uniform regularity assumptions, in particular, the laws of the increments can vary with respect to time and the walks may have some local dependence. A natural question is to ask, in that situation, what is the asymptotic behavior of the maxima. Similar questions were discussed in the context of branching Brownian motion using PDE techniques, see e.g. Nolen and Ryzhik [72], using the fact that the distributions of the maxima satisfy the KPP equation whose solution exhibits a traveling wave phenomenon.

In all these models, while the linear traveling speed of the maxima is a relatively easy consequence of the large deviation principle, the evaluation of the second order correction term, like the ones in Bramson [14] and Addario-Berry and Reed [1], is more involved and requires a detailed analysis of the walks; to our knowledge, it has so far only been performed in the time homogeneous case.

Our goal is to start exploring the time inhomogeneous setup. As we will detail below, the situation, even in the simplest setting, is complex and, for example, the order in which inhomogeneity presents itself matters, both in the leading term and in the correction term.

In this chapter, in order to best describe the phenomenon discussed above, we focus on the simplest case of binary branching random walks where the diffusivity of the particles takes two distinct values as a function of time. It is a modified version of Definition 1 in Chapter 1.

We now describe the setup in detail. For $\sigma > 0$, let $N(0, \sigma^2)$ denote the normal distributions with mean zero and variance σ^2. Let n be an integer, and let $\sigma_1^2, \sigma_2^2 > 0$ be given. We start the system with one particle at location 0 at time 0. Suppose that v is a particle at location S_v at time k. Then v dies at time $k+1$ and gives birth to two particles $v1$ and $v2$, and each of the two offspring ($\{vi, i = 1, 2\}$) moves independently to a new location S_{vi} with the increment $S_{vi} - S_v$ independent of S_v and distributed as $N(0, \sigma_1^2)$ if $k < n/2$ and as $N(0, \sigma_2^2)$ if $n/2 \leq k < n$. Let $\mathbb{D}_n$ denote the collection of all particles at time n. For a particle $v \in \mathbb{D}_n$ and $i < n$, we let v^i denote the ith level ancestor of v, that is the unique element of $\mathbb{D}_i$ on the geodesic connecting v and the root. We study the maximal displacement $M_n = \max_{v \in \mathbb{D}_n} S_v$ at time n, for n large. [1]

It will be clear that the analysis extends to a wide class of inhomogeneities with finitely many values and 'macroscopic' change (similar to the description in the previous paragraph), and to the Galton-Watson setup. A universal result that will allow for

[1] Since one can understand a branching random walk as a 'competition' between branching and random walk, one may get similar results by fixing the variance and changing the branching rate with respect to time.

continuous change of the variances is more complicated, is expected to present different correction terms, and is the subject of further study.

In order to describe the results in a concise way, we recall the notation $O_P(1)$ for stochastically boundedness. That is, if a sequence of random variables R_n satisfies $R_n = O_P(1)$, then, for any $\epsilon > 0$, there exists an M such that $P(|R_n| > M) < \epsilon$ for all n.

An interesting feature of M_n is that the asymptotic behavior depends on the order relation between σ_1^2 and σ_2^2. That is, while

$$M_n = \left(\sqrt{2\log 2}\,\sigma_{\text{eff}}\right) n - \beta \frac{\sigma_{\text{eff}}}{\sqrt{2\log 2}} \log n + O_P(1) \tag{4.1}$$

is true for some choice of σ_{eff} and β, σ_{eff} and β take different expressions for different ordering of σ_1 and σ_2. Note that (4.1) is equivalent to that the sequence $\{M_n - Med(M_n)\}_n$ is tight and

$$Med(M_n) = \left(\sqrt{2\log 2}\,\sigma_{\text{eff}}\right) n - \beta \frac{\sigma_{\text{eff}}}{\sqrt{2\log 2}} \log n + O(1),$$

where the median of a random variable X is denoted by $Med(X) = \sup\{x : P(X \leq x) \leq \frac{1}{2}\}$. In the following, we will use superscripts to distinguish different cases, see (4.2), (4.3) and (4.4) below.

A special and well-known case is when $\sigma_1 = \sigma_2 = \sigma$, i.e., all the increments are i.i.d.. In that case, the maximal displacement is described as follows:

$$M_n^{=} = \left(\sqrt{2\log 2}\,\sigma\right) n - \frac{3}{2}\frac{\sigma}{\sqrt{2\log 2}} \log n + O_P(1); \tag{4.2}$$

the proof can be found in [1], and its analog for branching Brownian motion can be found in [14] using probabilistic techniques and [57] using PDE techniques. This homogeneous case corresponds to (4.1) with $\sigma_{\text{eff}} = \sigma$ and $\beta = \frac{3}{2}$. In this chapter, we deal with the extension to the inhomogeneous case. The main results are the following two theorems.

Theorem 6. *When $\sigma_1^2 < \sigma_2^2$ (increasing variances), the maximal displacement is*

$$M_n^{\uparrow} = \left(\sqrt{(\sigma_1^2 + \sigma_2^2)\log 2}\right) n - \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{4\sqrt{\log 2}} \log n + O_P(1), \tag{4.3}$$

which is of the form (4.1) with $\sigma_{\text{eff}} = \sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}$ and $\beta = \frac{1}{2}$.

Theorem 7. *When $\sigma_1^2 > \sigma_2^2$ (decreasing variances), the maximal displacement is*

$$M_n^{\downarrow} = \frac{\sqrt{2\log 2}(\sigma_1+\sigma_2)}{2} n - \frac{3(\sigma_1+\sigma_2)}{2\sqrt{2\log 2}} \log n + O_P(1), \tag{4.4}$$

which is of the form (4.1) *with* $\sigma_{\text{eff}} = \frac{\sigma_1+\sigma_2}{2}$ *and* $\beta = 3$.

For comparison purpose, it is useful to introduce the model of 2^n independent (inhomogeneous) random walks with centered independent Gaussian variables, with variance profile as above. Denote by M_n^{ind} the maximal displacement at time n in this model. Because of the complete independence, it can be easily shown that

$$M_n^{\text{ind}} = \left(\sqrt{(\sigma_1^2+\sigma_2^2)\log 2}\right) n - \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{4\sqrt{\log 2}} \log n + O_P(1) \tag{4.5}$$

for all choices of σ_1^2 and σ_2^2. Thus, in this case, $\sigma_{\text{eff}} = \sqrt{(\sigma_1^2+\sigma_2^2)/2}$ and $\beta = 1/2$. Thus, the difference between $M_n^{=}$ and M_n^{ind} when $\sigma_1^2 = \sigma_2^2$ lies in the logarithmic correction. As commented (for branching Brownian motion) in [14], the different correction is due to the intrinsic dependence between particles coming from the branching structure in branching random walks.

Another related quantity is the sub-maximum obtained by a greedy algorithm, which only considers the maximum over all decendents of the maximal particle at time $n/2$. Applying (4.2), we find that the output of such algorithm is

$$\begin{aligned} &\left(\sqrt{2\log 2}\sigma_1 \frac{n}{2} - \frac{3}{2}\frac{\sigma_1}{\sqrt{2\log 2}} \log\frac{n}{2}\right) + \left(\sqrt{2\log 2}\sigma_2 \frac{n}{2} - \frac{3}{2}\frac{\sigma_2}{\sqrt{2\log 2}} \log\frac{n}{2}\right) + O_P(1) \\ &= \frac{\sqrt{2\log 2}(\sigma_1+\sigma_2)}{2} n - \frac{3(\sigma_1+\sigma_2)}{2\sqrt{2\log 2}} \log n + O_P(1). \end{aligned} \tag{4.6}$$

Comparing (4.6) with the theorems, we see that this algorithm yields the maximum up to an $O_P(1)$ error in the case of decreasing variances (compare with (4.4)) but not in the case of increasing variances (compare with (4.3)) or of homogeneous increments (compare with (4.2)).

A few comparisons are now in order.

1. When the variances are increasing, $M_n^{\uparrow}$ is asymptotically (up to $O_P(1)$ error) the same as M_n^{ind}, which is exactly the same as the maximum of independent homogeneous random walks with effective variance $\frac{\sigma_1^2+\sigma_2^2}{2}$.

2. When the variances are decreasing, $M_n^{\downarrow}$ shares the same asymptotic behavior with the sub-maximum (4.6). In this case, a greedy strategy yields the approximate maximum.

3. With the same set of diffusivity constants $\{\sigma_1^2, \sigma_2^2\}$ but different order, $M_n^{\uparrow}$ is greater than $M_n^{\downarrow}$.

4. While the leading order terms in (4.2), (4.3) and (4.4) are continuous in σ_1 and σ_2 (they coincide upon setting $\sigma_1 = \sigma_2$), the logarithmic corrections exhibit a phase transition phenomenon (they are not the same when we let $\sigma_1 = \sigma_2$).

We will prove Theorem 6 in Section 4.2 and Theorem 7 in Section 4.3. Before proving the theorems, we state a tightness result.

Lemma 9. *The sequences* $\{M_n^{\uparrow} - Med(M_n^{\uparrow})\}_n$ *and* $\{M_n^{\downarrow} - Med(M_n^{\downarrow})\}_n$ *are tight.*

This lemma follows from Bramson and Zeitouni [16] or Fang [30]. One can write down a similar recursion for the distribution of M_n to the one in those two papers, except for different subscripts and superscripts. Since the argument there depends only on one step of the recursion, it applies here directly without any change and leads to the tightness result in the lemma.

A note on notation: throughout, we use C to denote a generic positive constant, possibly depending on σ_1 and σ_2, that may change from line to line.

4.2 Increasing Variances: $\sigma_1^2 < \sigma_2^2$

In this section, we prove Theorem 6. We begin in Subsection 4.2.1 with a result on the fluctuation of an inhomogeneous random walk. In the short Subsection 4.2.2 we provide large-deviations based heuristics for our results. While it is not used in the actual proof, it explains the leading term of the maximal displacement and gives hints about the derivation of the logarithmic correction term. The actual proof of Theorem 6 is provided in subsection 4.2.3.

4.2.1 Fluctuation of an Inhomogeneous Random Walk

Let

$$S_n = \sum_{i=1}^{n/2} X_i + \sum_{i=n/2+1}^{n} Y_i \tag{4.7}$$

be an inhomogeneous random walk, where $X_i \sim N(0, \sigma_1^2)$, $Y_i \sim N(0, \sigma_2^2)$, and X_i and Y_i are independent. Define

$$s_{k,n}(x) = \begin{cases} \dfrac{\sigma_1^2 k}{(\sigma_1^2 + \sigma_2^2)\frac{n}{2}} x, & 0 \le k \le \dfrac{n}{2}, \\ \dfrac{\sigma_1^2 \frac{n}{2} + \sigma_2^2 (k - \frac{n}{2})}{(\sigma_1^2 + \sigma_2^2)\frac{n}{2}} x, & \dfrac{n}{2} \le k \le n, \end{cases} \tag{4.8}$$

and

$$f_{k,n} = \begin{cases} c_f k^{2/3}, & k \le n/2, \\ c_f (n-k)^{2/3}, & n/2 < k \le n. \end{cases} \tag{4.9}$$

As the following lemma says, conditioned on $\{S_n = x\}$, the path of the walk S_n follows $s_{k,n}(x)$ with fluctuation less than or equal to $f_{k,n}$ at level $k \le n$.

Lemma 10. *There exists a constant $C > 0$ (independent of n) such that*

$$P(S_n(k) \in [s_{k,n}(S_n) - f_{k,n}, s_{k,n}(S_n) + f_{k,n}] \ \textit{for all} \ 0 \le k \le n | S_n) \ge C,$$

where $S_n(k)$ is the sum of the first k summands of S_n, i.e.,

$$S_n(k) = \begin{cases} \displaystyle\sum_{k=1}^{k} X_k, & k \le n/2, \\ \displaystyle\sum_{k=1}^{n/2} X_k + \sum_{k=n/2+1}^{k} Y_k, & n/2 < k \le n. \end{cases}$$

Proof. Let $\tilde{S}_{k,n} = S_n(k) - s_{k,n}(S_n)$. Then, similar to Brownian bridge, one can check that $\tilde{S}_{k,n}$ are independent of S_n. To see this, first note that the covariance between $\tilde{S}_{k,n}$ and S_n is

$$Cov(\tilde{S}_{k,n}, S_n) = E\tilde{S}_{k,n} S_n - E\tilde{S}_{k,n} E S_n = E\tilde{S}_{k,n} S_n,$$

since $ES_n = 0$ and $E\tilde{S}_{k,n} = 0$.

For $k \leq n/2$,

$$\tilde{S}_{k,n} = \left(1 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)\sum_{i=1}^{k} X_i - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=k+1}^{n/2} X_i - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=n/2+1}^{n} Y_i.$$

Expand $\tilde{S}_{k,n}S_n$, take expectation, and then all terms vanish except for those containing X_i^2 and Y_i^2. Taking into account that $EX_i^2 = \sigma_1^2$ and $EY_i^2 = \sigma_2^2$, one has

$$\begin{aligned}
& Cov(\tilde{S}_{k,n}, S_n) = E\tilde{S}_{k,n}S_n \\
= & \left(1 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)\sum_{i=1}^{k} EX_i^2 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=k+1}^{n/2} EX_i^2 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=n/2+1}^{n} EY_i^2 \\
= & \left(1 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)k\sigma_1^2 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}(n/2-k)\sigma_1^2 - \frac{\sigma_1^2 k}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}(n/2)\sigma_2^2 \\
= & \ 0.
\end{aligned} \tag{4.10}$$

For $n/2 < k \leq n$, one can calculate $Cov(\tilde{S}_{k,n}, S_n) = 0$ similarly as follows. First,

$$\tilde{S}_{k,n} = \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=1}^{n/2} X_i + \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=n/2+1}^{k} Y_i - \left(1 - \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)\sum_{i=k+1}^{n} Y_i.$$

Then, expanding $\tilde{S}_{k,n}S_n$ and taking expectation, one has

$$\begin{aligned}
& Cov(\tilde{S}_{k,n}, S_n) = E\tilde{S}_{k,n}S_n \\
= & \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=1}^{n/2} EX_i^2 + \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\sum_{i=n/2+1}^{k} EY_i^2 - \left(1 - \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)\sum_{i=k+1}^{n} EY_i^2 \\
= & \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}(n/2)\sigma_1^2 + \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}(k-n/2)\sigma_2^2 - \left(1 - \frac{\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)\frac{n}{2}}\right)(n-k)\sigma_2^2 \\
= & \ 0
\end{aligned}$$

Therefore, $\tilde{S}_{k,n}$ are independent of S_n since they are Gaussian. Using this independence,

$$\begin{aligned}
& P(S_n(k) \in [s_{k,n}(S_n) - f_{k,n}, s_{k,n}(S_n) + f_{k,n}] \ \text{ for all } \ 0 \leq k \leq n | S_n) \\
= \ & P(\tilde{S}_{k,n} \in [-f_{k,n}, f_{k,n}] \ \text{ for all } \ 0 \leq k \leq n | S_n) \\
= \ & P(\tilde{S}_{k,n} \in [-f_{k,n}, f_{k,n}] \ \text{ for all } \ 0 \leq k \leq n).
\end{aligned}$$

By calculation similar to (4.10), $\tilde{S}_{k,n}$ is a Gaussian sequence with mean zero and variance $k\sigma_1^2\frac{\left((\sigma_1^2+\sigma_2^2)n-2\sigma_1^2k\right)}{(\sigma_1^2+\sigma_2^2)n}$ for $k \le n/2$ and $(n-k)\sigma_2^2\frac{\left((\sigma_1^2+\sigma_2)n-2\sigma_2^2(n-k)\right)}{(\sigma_1^2+\sigma_2^2)n}$ for $n/2 < k \le n$. The above quantity is

$$1 - P(|\tilde{S}_{k,n}| > f_{k,n}, \text{ for some } 0 \le k \le n) \ge 1 - \sum_{k=1}^{n} P(|\tilde{S}_{k,n}| > f_{k,n}).$$

Using a standard Gaussian estimate, e.g. [25, Theorem 1.4], the above quantity is at least,

$$1 - \sum_{k=1}^{n} \frac{c_0}{\sqrt{k}} e^{-\frac{f_{k,n}^2}{k}c_1} \ge 1 - 2\sum_{k=1}^{\infty} \frac{c_0}{\sqrt{k}} e^{-c_f^2 c_1 k^{1/3}} := C > 0$$

where c_0, c_1 are constants depending on σ_1 and σ_2, and $C > 0$ can be realized by choosing the constant c_f large. This proves the lemma. □

4.2.2 Sample Path Large Deviation Heuristics

We explain (without giving a proof) what we expect for the order n term of $M_n\uparrow$, by giving a large deviation argument. The exact proof will be postponed to the next subsection. Consider the same S_n as defined in (4.7) and a function $\phi(t)$ defined on $[0,1]$ with $\phi(0) = 0$. A sample path large deviation result, see [21, Theorem 5.1.2], tells us that the probability for $S_{\lfloor rn \rfloor}$ to be roughly $\phi(r)n$ for $0 \le r \le s \le 1$ is roughly $e^{-nI_s(\phi)}$, where

$$I_s(\phi) = \int_0^s \Lambda_r^*(\dot{\phi}(r))dr, \tag{4.11}$$

$\dot{\phi}(r) = \frac{d}{dr}\phi(r)$, and $\Lambda_r^*(x) = \frac{x^2}{2\sigma_1^2}$, for $0 \le r \le 1/2$, and $\frac{x^2}{2\sigma_1^2}$, for $1/2 < r \le 1$. A first moment argument would yield a necessary condition for a walk that roughly follows the path $\phi(r)n$ to exist among the branching random walks,

$$I_s(\phi) \le s\log 2, \quad \text{for all } \ 0 \le s \le 1. \tag{4.12}$$

This is equivalent to

$$\begin{cases} \displaystyle\int_0^s \frac{\dot{\phi}^2(r)}{2\sigma_1^2}dr \le s\log 2, & 0 \le s \le \dfrac{1}{2}, \\ \displaystyle\int_0^{\frac{1}{2}} \frac{\dot{\phi}^2(r)}{2\sigma_1^2}dr + \int_{\frac{1}{2}}^s \frac{\dot{\phi}^2(r)}{2\sigma_2^2}dr \le s\log 2, & \dfrac{1}{2} \le s \le 1. \end{cases} \tag{4.13}$$

Otherwise, if (4.12) is violated for some s_0, i.e., $I_{s_0}(\phi) > s_0 \log 2$, there will be no path following $\phi(r)n$ to $\phi(s_0)n$, since the expected number of such paths is $2^{sn}e^{-nI_s(\phi)} = e^{-(I_s(\phi)-s\log 2)n}$, which decreases exponentially.

Our goal is then to maximize $\phi(1)$ under the constraints (4.13). By Jensen's inequality and convexity, one can prove that it is equivalent to maximizing $\phi(1)$ subject to

$$\frac{\phi^2(1/2)}{\sigma_1^2} \le \frac{1}{2}\log 2, \ \frac{\phi^2(1/2)}{\sigma_1^2} + \frac{(\phi(1)-\phi(1/2))^2}{\sigma_2^2} \le \log 2. \tag{4.14}$$

Note that the above argument does not necessarily require $\sigma_1^2 < \sigma_2^2$.

Under the assumption that $\sigma_1^2 < \sigma_2^2$, we can solve the optimization problem with the optimal curve

$$\phi(s) = \begin{cases} \dfrac{2\sigma_1^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}s, & 0 \le s \le \dfrac{1}{2}, \\ \dfrac{2\sigma_1^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}\dfrac{1}{2} + \dfrac{2\sigma_2^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}}(s-\dfrac{1}{2}), & \dfrac{1}{2} \le s \le 1. \end{cases} \tag{4.15}$$

If we plot this optimal curve and the suboptimal curve leading to (4.6) as in Figure 4.1, it is easy to see that the ancestor at time $n/2$ of the actual maximum at time n is not a maximum at time $n/2$, since $\frac{2\sigma_1^2\sqrt{\log 2}}{\sqrt{(\sigma_1^2+\sigma_2^2)}} < \sqrt{2\sigma_1^2\log 2}$. A further rigorous calculation as in the next subsection shows that, along the optimal curve (4.15), the branching random walks have an exponential decay of correlation. Thus a fluctuation between $n^{1/2}$ and n that is larger than the typical fluctuation of a random walk is admissible. This is consistent with the naive observation from Figure 4.1. This kind of behavior also occurs in the independent random walks model, explaining why $M_n^{\uparrow}$ and M_n^{ind} have the same asymptotical expansion up to an $O(1)$ error, see (4.3) and (4.5).

4.2.3 Proof of Theorem 6

With Lemma 10 and the observation from Section 4.2.2, we can now provide a proof of Theorem 6, applying the first and second moments method to the appropriate sets. In the proof, we use S_n to denote the walk defined by (4.7) and S_k to denote the sum of the first k summand in S_n.

Proof of Theorem 6. *Upper bound.* Let $a_n = \sqrt{(\sigma_1^2+\sigma_2^2)\log 2}n - \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{4\sqrt{\log 2}}\log n$. Let $N_{1,n} = \sum_{v\in\mathbb{D}_n} 1_{\{S_v > a_n+y\}}$ be the number of particles at time n whose displacements are

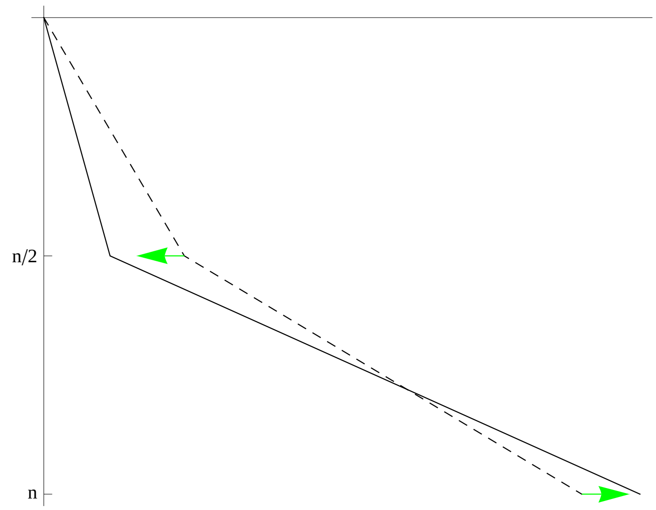


Figure 4.1: Dashed: maximum at time n of BRW starting from maximum at time $n/2$. Solid: maximum at time n of BRW starting from time 0.

greater than $a_n + y$. Then

$$EN_{1,n} = 2^n P(S_n \geq a_n + y) \leq c_2 e^{-c_3 y}$$

where c_2 and c_3 are constants independent of n and the last inequality is due to the fact that $S_n \sim N(0, \frac{\sigma_1^2+\sigma_2^2}{2} n)$. So we have, by the Chebyshev's inequality,

$$P(M_n^{\uparrow} > a_n + y) = P(N_1 \geq 1) \leq EN_{1,n} \leq c_2 e^{-c_3 y}. \tag{4.16}$$

Therefore, this probability can be made as small as we wish by choosing a large y.

Lower bound. Consider the walks which are at $s_n \in I_n = [a_n, a_n + 1]$ at time n and follow $s_{k,n}(s_n)$, defined by (4.8), at intermediate times with fluctuation bounded by $f_{k,n}$, defined by (4.9). Let $I_{k,n}(x) = [s_{k,n}(x) - f_{k,n}, s_{k,n}(x) + f_{k,n}]$ be the 'admissible' interval at time k given $S_n = x$, and let

$$N_{2,n} = \sum_{v \in \mathbb{D}_n} 1_{\{S_v \in I_n, S_{v^k} \in I_{k,n}(S_v) \text{ for all } 0 \leq k \leq n\}}$$

be the number of such walks. By Lemma 10,

$$\begin{aligned} EN_{2,n} &= 2^n P(S_n \in I_n, S_n(k) \in I_{k,n}(S_n) \text{ for all } 0 \leq k \leq n) \\ &= 2^n E(1_{\{S_n \in I_n\}} P(S_n(k) \in I_{k,n}(S_n) \text{ for all } 0 \leq k \leq n | S_n)) \\ &\geq 2^n C P(S_n \in I_n) \geq c_4. \end{aligned} \tag{4.17}$$

Next, we bound the second moment $EN_{2,n}^2$. By considering the location of any pair $v_1, v_2 \in \mathbb{D}_n$ of particles at time n and at their common ancestor $v_1 \wedge v_2$, we have

$$\begin{aligned} EN_{2,n}^2 &= E \sum_{v_1,v_2\in\mathbb{D}_n} 1_{\{S_{v_i}\in I_n,\, S_{(v_i)^j}\in I_{j,n}(S_{(v_i)^j}) \text{ for all } 0\le j\le n, i=1,2\}} \\ &= \sum_{k=0}^{n} \sum_{\substack{v_1,v_2\in\mathbb{D}_n \\ v_1\wedge v_2\in\mathbb{D}_k}} E1_{\{S_{v_i}\in I_n,\, S_{(v_i)^j}\in I_{j,n}(S_{(v_i)^j}) \text{ for all } 0\le j\le n, i=1,2\}} \\ &\le \sum_{k=0}^{n} \sum_{\substack{v_1,v_2\in\mathbb{D}_n \\ v_1\wedge v_2\in\mathbb{D}_k}} P(S_{v_1}\in I_n,\ S_{(v_1)^j}\in I_{j,n}(S_{(v_1)^j}) \text{ for all } 0\le j\le n) \\ &\qquad \cdot P(S_{v_2}-S_{v_1\wedge v_2}\in[x-s_{k,n}(x)-f_{k,n}, x-s_{k,n}(x)+f_{k,n}], x\in I_n), \end{aligned}$$

where we use the independence between $S_{v_2}-S_{v_1\wedge v_2}$ and $S_{(v_1)^j}$ in the last inequality. And the last expression (double sum) in the above display is

$$\begin{aligned} &\sum_{k=0}^{n} 2^{2n-k} P(S_n\in I_n, S_n(j)\in I_{j,n}(S_n) \text{ for all } 0\le j\le n) \\ &\qquad \cdot P(S_n-S_n(k)\in[x-s_{k,n}(x)-f_{k,n}, x-s_{k,n}(x)+f_{k,n}], x\in I_n) \\ \le\ & EN_{2,n}\sum_{k=0}^{n} 2^{n-k} P(S_n-S_n(k)\in[x-s_{k,n}(x)-f_{k,n}, x-s_{k,n}(x)+f_{k,n}], x\in I_n). \end{aligned}$$

The above probabilities can be estimated separately when $k\le n/2$ and $n/2<k\le n$. For $k\le n/2$, $S_n-S_n(k)\sim N(0,\frac{n}{2}(\sigma_1^2+\sigma_2^2)-k\sigma_1^2)$. Thus,

$$\begin{aligned} &P(S_n-S_n(k)\in[x-s_{k,n}(x)-f_{k,n}, x-s_{k,n}(x)+f_{k,n}], x\in I_n) \\ \le\ & 2f_{k,n}\frac{1}{\sqrt{\pi((\sigma_1^2+\sigma_2^2)n-2k\sigma_1^2)}}\exp\left(-\frac{\left((1-\frac{2\sigma_1^2k}{(\sigma_1^2+\sigma_2^2)n})a_n-f_{k,n}\right)^2}{(\sigma_1^2+\sigma_2^2)n-2k\sigma_1^2}\right) \\ \le\ & 2^{-n+\frac{2\sigma_1^2}{\sigma_1^2+\sigma_2^2}k+o(k)}. \end{aligned}$$

For $n/2<k\le n$, $S_n-S_n(k)\sim N(0,(n-k)\sigma_2^2)$. Thus,

$$\begin{aligned} &P(S_n-S_n(k)\in[x-s_{k,n}(x)-f_{k,n}, x-s_{k,n}(x)+f_{k,n}], x\in I_n) \\ \le\ & 2f_{k,n}\frac{1}{\sqrt{2\pi(n-k)\sigma_2^2}}\exp\left(-\frac{\left(\frac{2\sigma_2^2(n-k)}{(\sigma_1^2+\sigma_2^2)n}a_n-f_{k,n}\right)^2}{2(n-k)\sigma_2^2}\right) \\ \le\ & 2^{-\frac{2\sigma_2^2}{\sigma_1^2+\sigma_2^2}(n-k)+o(n-k)}. \end{aligned}$$

Therefore,

$$EN_{2,n}^2 \le EN_{2,n}\left(\sum_{k=0}^{n/2} 2^{\frac{\sigma_1^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2}k+o(k)} + \sum_{k=n/2+1}^{n} 2^{\frac{\sigma_1^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2}(n-k)+o(n-k)}\right) \le c_5 EN_{2,n}, \quad (4.18)$$

where $c_5 = 2\sum_{k=0}^{\infty} 2^{\frac{\sigma_1^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2}k+o(k)}$. By the Cauchy-Schwartz inequality,

$$P(M_n^{\uparrow} \ge a_n) \ge P(N_{2,n} > 0) \ge \frac{(EN_{2,n})^2}{EN_{2,n}^2} \ge c_4/c_5 > 0. \quad (4.19)$$

The upper bound (4.16) and lower bound (4.19) imply that there exists a large enough constant y_0 such that

$$P(M_n^{\uparrow} \in [a_n, a_n + y_0]) \ge \frac{c_4}{2c_5} > 0.$$

Lemma 9 tells us that the sequence $\{M_n^{\uparrow} - \text{Med}(M_n^{\uparrow})\}_n$ is tight, so $M_n^{\uparrow} = a_n + O(1)$ a.s.. That completes the proof. □

4.3 Decreasing Variances: $\sigma_1^2 > \sigma_2^2$

We will again separate the proof of Theorem 7 into two parts, the lower bound and the upper bound. Fortunately, we can apply (4.2) directly to get a lower bound so that we can avoid repeating the second moment argument. However, we do need to reproduce (the first moment argument) part of the proof of (4.2) in order to get an upper bound.

4.3.1 An Estimate for Brownian Bridge

Toward this end, we need the following analog of Bramson [14, Proposition 1']. The original proof in Bramson's used the Gaussian density and reflection principle of continuous time Brownian motion, which also hold for the discrete time version. The proof extends without much effort to yield the following estimate for the Brownian bridge $B_k - \frac{k}{n}B_n$, where B_n is a random walk with standard normal increments.

Lemma 11. *Let*

$$L(k) = \begin{cases} 0 & \textit{if } s = 0, n, \\ 100\log k & \textit{if } k = 1, \ldots, n/2, \\ 100\log(n-k) & \textit{if } k = n/2, \ldots, n-1. \end{cases}$$

Then, there exists a constant C such that, for all $y > 0$,

$$P(B_n - \frac{k}{n}B_n \leq L(k) + y \text{ for } 0 \leq k \leq n) \leq \frac{C(1+y)^2}{n}.$$

The coefficient 100 before log is chosen large enough to be suitable for later use, and is not crucial in Lemma 11.

4.3.2 Proof of Theorem 7

Before proving the theorem, we discuss the equivalent optimization problems (4.13) and (4.14) under our current setting $\sigma_1^2 > \sigma_2^2$. It can be solved by employing the optimal curve

$$\phi(s) = \begin{cases} \sqrt{2\log 2}\sigma_1 s, & 0 \leq s \leq \frac{1}{2}, \\ \sqrt{2\log 2}\sigma_1 \frac{1}{2} + \sqrt{2\log 2}\sigma_2 (s - \frac{1}{2}), & \frac{1}{2} \leq s \leq 1. \end{cases} \tag{4.20}$$

If we plot the curve $\phi(s)$ and the suboptimal curve leading to (4.6) as in Figure 4.2, these two curves coincide with each other up to order n. Figure 4.2 seems to indicate that the maximum at time n for the branching random walk starting from time 0 comes from the maximum at time $n/2$. As will be shown rigorously, if a particle at time $n/2$ is left significantly behind the maximum, its descendents will not be able to catch up by time n. The difference between Figure 4.1 and Figure 4.2 explains the difference in the logarithmic correction between $M_n^{\uparrow}$ and $M_n^{\downarrow}$.

Proof of Theorem 7. Lower Bound. For each $i = 1, 2$, the formula (4.2) implies that there exist y_i (possibly negative) such that, for branching random walk at time $n/2$ with variance σ_i^2,

$$P\left(M_{n/2} > \sqrt{2\log 2}\sigma_i \frac{n}{2} - \frac{3\sigma_i}{2\sqrt{2\log 2}} \log \frac{n}{2} + y_i\right) \geq \frac{1}{2}.$$

By considering a branching random walk starting from a particle at time $n/2$, whose location is greater than $\sqrt{2\log 2}\sigma_1 \frac{n}{2} - \frac{3\sigma_1}{2\sqrt{2\log 2}} \log \frac{n}{2} + y_1$, and applying the above display with $i = 1$ and 2,we know that

$$P\left(M_n^{\downarrow} > \frac{\sqrt{2\log 2}(\sigma_1 + \sigma_2)}{2} n - \frac{3(\sigma_1 + \sigma_2)}{2\sqrt{2\log 2}} \log \frac{n}{2} + y_1 + y_2\right) \geq \frac{1}{4}. \tag{4.21}$$

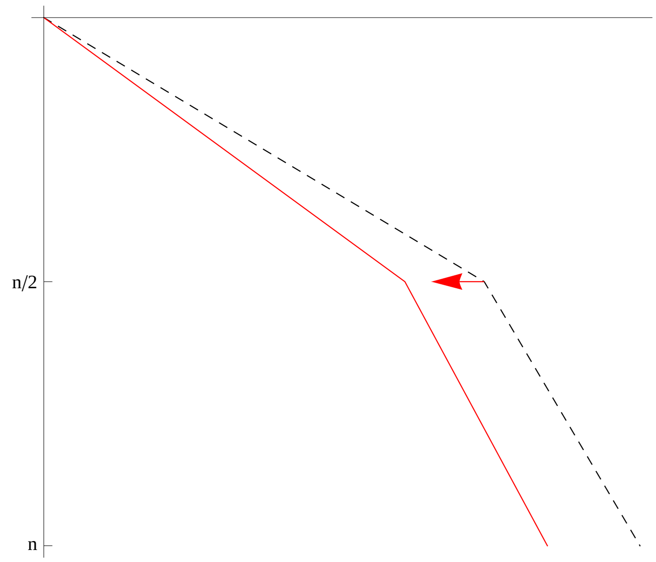


Figure 4.2: Dash: both the optimal path to the maximum at time n and the path leading to the maximum of BRW starting from the maximum at time $n/2$. Solid: the path to the maximal (rightmost) descendent of a particle at time $n/2$ that is significantly behind the maximum then.

Upper Bound. We will use a first moment argument to prove that there exists a constant y (large enough) such that

$$P\left(M_n^{\downarrow} > \frac{\sqrt{2\log 2}(\sigma_1+\sigma_2)}{2}n - \frac{3(\sigma_1+\sigma_2)}{2\sqrt{2\log 2}}\log\frac{n}{2} + y\right) < \frac{1}{10}. \tag{4.22}$$

Similarly to the last argument in the proof of Theorem 6, the upper bound (4.22) and the lower bound (4.21), together with the tightness result from Lemma 9, prove Theorem 7. So it remains to show (4.22).

Toward this end, we define a polygonal line (piecewise linear curve) leading to $\frac{\sqrt{2\log 2}(\sigma_1+\sigma_2)}{2}n - \frac{3(\sigma_1+\sigma_2)}{2\sqrt{2\log 2}}\log\frac{n}{2}$ as follows: for $1 \le k \le n/2$,

$$M(k) = \frac{k}{n/2}(\sqrt{2\log 2}\sigma_1\frac{n}{2} - \frac{3\sigma_1}{2\sqrt{2\log 2}}\log\frac{n}{2});$$

and for $n/2+1 \le k \le n$,

$$M(k) = M(n/2) + \frac{k-n/2}{n/2}(\sqrt{2\log 2}\sigma_2\frac{n}{2} - \frac{3\sigma_2}{2\sqrt{2\log 2}}\log\frac{n}{2}).$$

Note that $\frac{k}{n}\log n \le \log k$ for $k \le n$. Also define

$$f(k) = \begin{cases} y & k = 0, \frac{n}{2}, n, \\ y + \frac{5\sigma_1}{2\sqrt{2\log 2}}\log k & 1 \le k \le n/4, \\ y + \frac{5\sigma_1}{2\sqrt{2\log 2}}\log(\frac{n}{2} - k) & \frac{n}{4} \le k \le \frac{n}{2} - 1, \\ y + \frac{5\sigma_2}{2\sqrt{2\log 2}}\log(k - \frac{n}{2}) & \frac{n}{2} + 1 \le k \le \frac{3n}{4}, \\ y + \frac{5\sigma_2}{2\sqrt{2\log 2}}\log(n - k) & \frac{3n}{4} \le k \le n - 1. \end{cases}$$

We will use $f(k)$ to denote the allowed offset (deviation) from $M(k)$ in the following argument.

The probability on the left side of (4.22) is equal to

$$P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n) + y).$$

For each $v \in \mathbb{D}_n$, we define $\tau_v = \inf\{k : S_{v^k} > M(k) + f(k)\}$; then (4.22) is implied by

$$\sum_{k=1}^{n} P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n) + y, \tau_v = k) < 1/10. \tag{4.23}$$

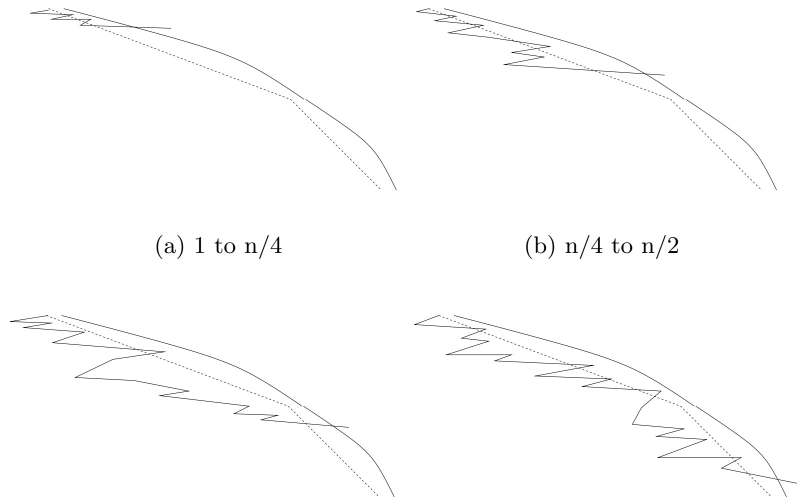
We will split the sum into four regimes: $[1, n/4]$, $[n/4, n/2]$, $[n/2, 3n/4]$ and $[3n/4, n]$, corresponding to the four parts of the definition of $f(k)$. The sum over each regime, corresponding to the events in the four pictures in Figure 4.3, can be made small. The first two are the discrete analog of the upper bound argument in Bramson [14]. We will present a complete proof for the first two cases, since the argument is not too long and the argument (not only the result) is used in the latter two cases.

(i). When $1 \le k \le n/4$, we have, by the Chebyshev's inequality,

$$\begin{aligned} & P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n) + y, \tau_v = k) \\ \le\ & P(\exists v \in \mathbb{D}_k, \text{ such that } S_v > M(k) + f(k)) \le E \sum_{v \in \mathbb{D}_k} 1_{\{S_v > M(k) + f(k)\}}. \end{aligned}$$

The above expectation is less than or equal to

$$\begin{aligned} \frac{C2^k}{\sqrt{k}} e^{-\frac{(M(k)+f(k))^2}{2\sigma_1^2}} &\le \frac{C2^k}{\sqrt{k}} \exp\left(-\frac{\left(\sqrt{2\log 2}\sigma_1 k + \frac{\sigma_1}{\sqrt{2\log 2}}\log k + y\right)^2}{2k\sigma_1^2} \right) \\ &\le Ck^{-3/2} e^{-\frac{\sqrt{2\log 2}}{\sigma_1} y}. \end{aligned} \tag{4.24}$$



(a) 1 to n/4 (b) n/4 to n/2 (c) n/2 to 3n/4 (d) 3n/4 to n

Figure 4.3: Four small probability events. Dash line: $M(k)$. Solid curve: $M(k)+f(k)$. Polygonal line: a random walk.

Summing these upper bounds over $k \in [1, n/4]$, we obtain that

$$\sum_{k=1}^{n/4} P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \le Ce^{-\frac{\sqrt{2\log 2}}{\sigma_1}y} \sum_{k=1}^{\infty} k^{-3/2}. \quad (4.25)$$

The right side of the above inequality can be made as small as we wish, say at most $\frac{1}{100}$, by choosing y large enough.

(ii). When $n/4 \le k \le n/2$, we again have, by Chebyshev's inequality,

$$\begin{aligned} & P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \\ \le\ & P(\exists v \in \mathbb{D}_k, \text{ such that } S_v > M(k)+f(k), \text{ and } S_{v^i} \le M(i)+f(i) \text{ for } 1 \le i \le k) \\ \le\ & E \sum_{v \in \mathbb{D}_k} 1_{\{S_v > M(k)+f(k), \text{ and } S_{v^i} \le M(i)+f(i) \text{ for } 1 \le i < k\}}. \end{aligned}$$

Letting S_k be a copy of the random walks before time $n/2$, then the above expectation is equal to

$$\begin{aligned} & 2^k P(S_k > M(k)+f(k), \text{ and } S_i \le M(i)+f(i) \text{ for } 1 \le i < k) \\ \le\ & 2^k P(S_k > M(k)+f(k), \text{ and } \frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k) \le \frac{1}{\sigma_1}(f(i) - \frac{i}{k}f(k)) \text{ for } 1 \le i \le k). \end{aligned} \quad (4.26)$$

$\frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k)$ is a discrete Brownian bridge and is independent of S_k. Because of this independence, the above quantity is less than or equal to

$$2^k P(S_k > M(k) + f(k)) \cdot P(\frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k) \leq \frac{1}{\sigma_1}(f(i) - \frac{i}{k}f(k)) \text{ for } 1 \leq i < k).$$

The first probability can be estimated similarly to (4.24),

$$\begin{aligned} & P(S_k > M(k) + f(k)) \\ \leq \;& \frac{C}{\sqrt{k}} \exp\left(- \frac{\left(\sqrt{2\log 2}\sigma_1 k - \frac{3\sigma_1}{2\sqrt{2\log 2}}\log k + \frac{5\sigma_1}{2\sqrt{2\log 2}}\log(\frac{n}{2} - k) + y\right)^2}{2k\sigma_1^2} \right) \\ \leq \;& C2^{-k}k(\frac{n}{2} - k)^{-5/2} e^{-\frac{\sqrt{2\log 2}}{\sigma_1}y}. \end{aligned} \tag{4.27}$$

To estimate the second probability, we first estimate $\frac{1}{\sigma_1}(f(i) - \frac{i}{k}f(k))$. It is less than or equal to $\frac{1}{\sigma_1}f(i) = \frac{y}{\sigma_1} + \frac{5}{2\sqrt{2\log 2}}\log i$ for $i \leq k/2 < n/4$, and, for $k/2 \leq i < k$, it is less than or equal to

$$\begin{aligned} & \frac{5}{2\sqrt{2\log 2}}\log(n/2 - i) - \frac{i}{k}\frac{5}{2\sqrt{2\log 2}}\log(n/2 - k) + \frac{y}{\sigma_1}(1 - \frac{i}{k}) \\ = \;& \frac{5}{2\sqrt{2\log 2}}\left(\log(n/2 - i) - \log(n/2 - k) + \frac{k-i}{k}\log(n/2 - k)\right) + \frac{y}{\sigma_1}(1 - \frac{i}{k}) \\ \leq \;& \frac{5}{2\sqrt{2\log 2}}\left(\log(k - i) + \frac{k-i}{k}\log k\right) + \frac{y}{\sigma_1} \leq 100\log(k - i) + \frac{y}{\sigma_1}. \end{aligned}$$

Therefore, applying Lemma 11, we have

$$\begin{aligned} & P\left(\frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k) \leq \frac{1}{\sigma_1}(f(i) - \frac{i}{k}f(k)) \text{ for } 1 \leq i \leq k\right) \\ \leq \;& P\left(\frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k) \leq 100\log i + \frac{y}{\sigma_1} \text{ for } 1 \leq i \leq k/2, \text{ and } \frac{1}{\sigma_1}(S_i - \frac{i}{k}S_k) \leq \right. \\ & \left. 100\log(k - i) + \frac{y}{\sigma_1} \text{ for } k/2 \leq i \leq k\right) \leq C(1 + y)^2/k, \end{aligned} \tag{4.28}$$

where C is independent of n, k and y.

By all the above estimates (4.26), (4.27) and (4.28),

$$\sum_{k=n/4}^{n/2} P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n) + y, \tau_v = k) \leq C(1 + y)^2 e^{-\frac{\sqrt{2\log 2}}{\sigma_1}y} \sum_{k=1}^{\infty} k^{-5/2}. \tag{4.29}$$

This can again be made as small as we wish, say at most $\frac{1}{100}$, by choosing y large enough.

(iii). When $n/2 \le k \le 3n/4$, we have

$$\begin{aligned} & P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \\ \le\ & P(\exists v \in \mathbb{D}_k \text{ such that } S_v > M(k)+f(k) \text{ and } S_{v^i} \le M(i)+f(i) \text{ for } 1 \le i \le n/2) \\ \le\ & E \sum_{v \in \mathbb{D}_k} 1_{\{S_v > M(k)+f(k), \text{ and } S_{v^i} \le M(i)+f(i) \text{ for } 1 \le i < n/2\}}. \end{aligned}$$

The above expectation is, by conditioning on $\{S_{v^{n/2}} = M(n)+x\}$,

$$\begin{aligned} 2^k \int_{-\infty}^{y} & P(S'_{k-n/2} > M(k) - M(n/2) + f(k) - x) \cdot \\ & \cdot P(S_i - \frac{i}{n/2} S_{n/2} \le f(i) - \frac{i}{k} x \text{ for } 1 \le i < n/2) \cdot \\ & \cdot p_{S_{n/2}}(M(n/2)+x)dx, \end{aligned} \tag{4.30}$$

where S and S' are two copies of the random walks before and after time $n/2$, respectively, and $p_{S_{n/2}}(x)$ is the density of $S_{n/2} \sim N(0, \frac{\sigma_1^2 n}{2})$.

We then estimate the three factors of the integrand separately. The first one, which is similar to (4.24), is bounded above by

$$\begin{aligned} & P(S'_{k-n/2} > M(k) - M(n/2) + f(k) - x) \le \frac{C}{\sqrt{k-n/2}} e^{-\frac{(M(k)-M(n/2)+f(k)-x)^2}{2(k-n/2)\sigma_2^2}} \\ \le\ & C 2^{-(k-n/2)} (k - \frac{n}{2})^{-3/2} e^{-\frac{\sqrt{2\log 2}}{\sigma_2}(y-x)}. \end{aligned}$$

The second one, which is similar to (4.28), is estimated using Lemma 11,

$$P(S_i - \frac{i}{n/2} S_{n/2} \le f(i) - \frac{i}{k} x \text{ for } 1 \le i < n/2) \le C(1+2y-x)^2/n. \tag{4.31}$$

The third one is simply the normal density

$$p_{S_{n/2}}(M(n/2)+x) = \frac{C}{\sqrt{n}} e^{-\frac{(M(n/2)+x)^2}{n\sigma_1^2}} \le C 2^{-n/2} n e^{-\frac{\sqrt{2\log 2}}{\sigma_1} x}. \tag{4.32}$$

Therefore, the integral term (4.30) is no more than

$$C(k-n/2)^{-3/2} e^{-\frac{\sqrt{2\log 2}}{\sigma_2} y} \int_{-\infty}^{y} (1+2y-x)^2 e^{(\frac{\sqrt{2\log 2}}{\sigma_2} - \frac{\sqrt{2\log 2}}{\sigma_1})x} dx,$$

which is less than or equal to $C(1+y)^2 e^{-\frac{\sqrt{2\log 2}}{\sigma_1}y}(k-n/2)^{-3/2}$ since $\sigma_2 < \sigma_1$.

Summing these upper bounds together, we obtain that

$$\sum_{k=n/2}^{3n/4} P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \leq C(1+y)^2 e^{-\frac{\sqrt{2\log 2}}{\sigma_1}y} \sum_{k=1}^{\infty} k^{-3/2}. \tag{4.33}$$

This can again be made as small as we wish, say at most $\frac{1}{100}$, by choosing y large enough.

(iv). When $3n/4 < k \leq n$, we have

$$\begin{aligned} & P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \\ \leq \;& P(\exists v \in \mathbb{D}_k \text{ such that } S_v > M(k)+f(k), \text{ and } S_{v^i} \leq M(i)+f(i) \text{ for } 1 \leq i < k) \\ \leq \;& E \sum_{v \in \mathbb{D}_k} 1_{\{S_v > M(k)+f(k), \text{ and } S_{v^i} \leq M(i)+f(i), \text{ for } 1 \leq i < k\}}. \end{aligned}$$

The above expectation is, by conditioning on $\{S_{v^{n/2}} = M(n)+x\}$,

$$\begin{aligned} 2^k \int_{-\infty}^{y} & P(S'_{k-n/2} > M(k) - M(n/2) + f(k) - x, \\ & \qquad S'_i < M(i) - M(n/2) + f(i) - x, \text{ for } n/2 < i \leq k) \\ & \cdot P(S_i - \frac{i}{n/2} S_{n/2} \leq f(i) - \frac{i}{k} x \text{ for } 1 \leq i < n/2) \cdot p_{S_{n/2}}(M(n/2)+x) dx \end{aligned}$$

where S and S' are copies of the random walks before and after time $n/2$, respectively.

The second and third probabilities in the integral are already estimated in (4.31) and (4.32). It remains to bound the first probability. Similar to (4.26), it is bounded above by

$$\begin{aligned} P\Big(& S'_{k-n/2} > M(k) - M(n/2) + f(k) - x, S'_i < M(i) - M(n/2) + f(i) - x, \\ & \text{for } n/2 < i \leq k\Big) \leq C(1+2y-x)^2 e^{-\frac{\sqrt{2\log 2}}{\sigma_2}(2y-x)} (n-k)^{-5/2}. \end{aligned}$$

With these estimates, we obtain in this case, in the same way as in (iii), that

$$\sum_{k=3n/4}^{n} P(\exists v \in \mathbb{D}_n \text{ such that } S_v > M(n)+y, \tau_v = k) \leq C(1+y)^2 e^{-\frac{\sqrt{2\log 2}}{\sigma_1}y} \sum_{k=1}^{\infty} k^{-5/2}. \tag{4.34}$$

This can again be made as small as we wish, say at most $\frac{1}{100}$, by choosing y large enough.

Summing (4.25), (4.29), (4.33) and (4.34), then (4.23) and thus (4.22) follow. This concludes the proof of Theorem 7. □

4.4 Further Remarks

We state several immediate generalization and open questions related to binary branching random walks in time inhomogeneous environments where the diffusivity of the particles takes more than two distinct values as a function of time and changes macroscopically.

Results involving finitely many monotone variances can be obtained similarly to the results on two variances in the previous sections. Specifically, let $k \geq 2$ (constant) be the number of inhomogeneities, $\{\sigma_i^2 > 0 : i = 1, \ldots, k\}$ be the set of variances and $\{t_i > 0 : i = 1, \ldots, k\}$, satisfying $\sum_{i=1}^k t_i = 1$, denote the portions of time when σ_i^2 governs the diffusivity. Consider binary branching random walk up to time n, where the increments over the time interval $[\sum_{i=1}^{j-1} t_i n, \sum_{i=1}^{j} t_i n)$ are $N(0, \sigma_j^2)$ for $1 \leq j \leq k$. When $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_k^2$ are strictly increasing, by an argument similar to that in Section 4.2, the maximal displacement at time n, which behaves asymptotically like the maximum for independent random walks with effective variance $\sum_{i=1}^k t_i \sigma_i^2$, is

$$\sqrt{2(\log 2) \sum_{i=1}^{k} t_i \sigma_i^2 n} - \frac{1}{2} \frac{\sqrt{\sum_{i=1}^k t_i \sigma_i^2}}{\sqrt{2 \log 2}} \log n + O_P(1).$$

When $\sigma_1^2 > \sigma_2^2 > \cdots > \sigma_k^2$ are strictly decreasing, by an argument similar to that in Section 4.3, the maximal displacement at time n, which behaves like the sub-maximum chosen by the previous greedy strategy (see (4.6)), is

$$\sqrt{2 \log 2} (\sum_{i=1}^{k} t_i \sigma_i) n - \frac{3}{2} (\sum_{i=1}^{k} \frac{\sigma_i}{\sqrt{2 \log 2}}) \log n + O_P(1).$$

Results on other inhomogeneous environments are open and are subjects of further study. We only discuss some of the non rigorous intuition in the rest of this section.

In the finitely many variances case, when $\{\sigma_i^2 : i = 1, \ldots, k\}$ are not monotone in i, the analysis of maximal displacement could be case-by-case and a mixture of the previous monotone cases. The leading order term is surely a result of the optimization problem (4.12) from the large deviation. But, the second order term may depend on the fluctuation constraints of the path leading to the maximum, as in the monotone case. One could probably find hints on the fluctuation from the optimal curve solving (4.12). In some segments, the path may behave like Brownian bridge (as in the decreasing variances case), and in some segments, the path may behave like a random walk (as in the increasing variances case).

In the case where the number of different variances increases as the time n increases, analysis seems more challenging. A special case is when the variances are decreasing, for example, at time $0 \le i \le n$ the increment of the walk is $N(0, \sigma_{i,n}^2)$ with $\sigma_{i,n}^2 = 2 - i/n$. The heuristics (from the finitely many decreasing variances case) seem to indicate that the path leading to the maximum at time n cannot be left 'significantly' behind the maxima at all intermediate levels. This path is a 'rightmost' path. From the intuition of [31] and Chapter 2, if the allowed fluctuation is of order n^α ($\alpha < 1/2$), then the correction term is of order $n^{1-2\alpha}$, instead of $\log n$ in (4.1). However, the allowed fluctuation from the intermediate maxima, implicitly imposed by the variances, becomes complicated as the difference between the consecutive variances decreases to zero. A good understanding of this fluctuation may be a key to finding the correction term.

References

[1] Louigi Addario-Berry and Bruce Reed. Minima in branching random walks. *Ann. Probab.*, 37(3):1044–1079, 2009.

[2] Elie Aïdékon. Convergence in law of the minimum of a branching random walk. *arXiv:1101.1810v1*, 2010.

[3] Elie Aïdékon and Zhan Shi. Weak convergence for the minimal position in a branching random walk: a simple proof. *Period. Math. Hungar.*, 61(1-2):43–54, 2010.

[4] David Aldous. Greedy search on the binary tree with random edge-weights. *Combin. Probab. Comput.*, 1(4):281–293, 1992.

[5] David Aldous and Jim Jill. *Reversible Markov Chains and Random Walks on Graphs.* University of California, Berkeley, 2002.

[6] Noga Alon and Joel H. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization, third edition, 2008.

[7] Markus Bachmann. Limit theorems for the minimal position in a branching random walk with independent logconcave displacements. *Adv. in Appl. Probab.*, 32(1):159–176, 2000.

[8] Henri Berestycki, François Hamel, and Nikolai Nadirashvili. The speed of propagation for KPP type problems. II. General domains. *J. Amer. Math. Soc.*, 23(1):1–34, 2010.

[9] Henri Berestycki, Grégoire Nadin, Benoit Perthame, and Lenya Ryzhik. The non-local Fisher-KPP equation: travelling waves and steady states. *Nonlinearity*, 22(12):2813–2844, 2009.

[10] J. Berestycki, É. Brunet, J. W. Harris, and S. C. Harris. The almost-sure population growth rate in branching brownian motion with a quadratic breeding potential. *Statist. Probab. Lett.*, 80(17-18):1442–1446, 2010.

[11] J. D. Biggins. The first-and-last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, 8(3):446–459, 1976.

[12] J. D. Biggins and A. E. Kyprianou. Seneta-heyde norming in the branching random walk. *Ann. Probab.*, 25(1):337–360, 1997.

[13] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Adv. in Appl. Probab.*, 36(2):544–581, 2004.

[14] Maury Bramson. Maximal displacement of branching brownian motion. *Comm. Pure Appl. Math.*, 31(5):531–581, 1978.

[15] Maury Bramson. Minimal displacement of branching random walk. *Z. Wahrsch. Verw. Gebiete*, 45(2):89–108, 1978.

[16] Maury Bramson and Ofer Zeitouni. Tightness for a family of recursion equations. *Ann. Probab.*, 37(2):615–653, 2009.

[17] Maury Bramson and Ofer Zeitouni. Tightness of the recentered maximum of the two-dimensional discrete gaussian free field. *Comm. Pure Appl. Math.*, 65(1):1–20, 2012.

[18] Brigitte Chauvin and Alain Rouault. Kpp equation and supercritical branching brownian motion in the subcritical speed area. *Probab. Theory Related Fields*, 80(2), 1988.

[19] Brigitte Chauvin and Alain Rouault. Supercritical branching brownian motion and k-p-p equation in the critical speed area. *Math. Nachr.*, 149, 1990.

[20] F. M. Dekking and B. Host. Limit distributions for minimal displacement of branching random walks. *Probab. Theory Related Fields*, 90(3):403–426, 1991.

[21] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications.* Applications of Mathematics, second edition, 1998.

[22] Luc Devroye and Bruce Reed. On the variance of the height of random binary search trees. *SIAM J. Comput.*, 24(6), 1995.

[23] Leif Doering and Matthew Roberts. Catalytic branching processes via spine techniques and renewal theory. *arXiv:1106.5428v3*, 2011.

[24] Peter G. Doyle and J. Laurie Snell. *Random walks and electric networks.* Carus Mathematical Monographs, 22., 1984.

[25] R. Durrett. *Probability: Theory and Examples.* Thomson Brooks/Cole, third edition, 2005.

[26] Richard Durrett. Maxima of branching random walks. *Z. Wahrsch. Verw. Gebiete*, 62(2), 1983.

[27] János Engländer. The center of mass for spatial branching processes and an application for self-interaction. *Electron. J. Probab.*, 15(63), 2010.

[28] János Engländer, Simon C. Harris, and Andreas E. Kyprianou. Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):279–298, 2010.

[29] János Engländer and Andreas E. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.*, 32(1A):78–99, 2004.

[30] Ming Fang. Tightness for maxima of generalized branching random walks. *arXiv:1012.0826v1*, 2011.

[31] Ming Fang and Ofer Zeitouni. Consistent minimal displacement of branching random walks. *Electron. Commun. Probab.*, 15:106–118, 2010.

[32] Gabriel Faraud, Yueyun Hu, and Shi Zhan. An almost sure convergence for stochastically biased random walks on trees. *arXiv:1003.5505v5*, 2010.

[33] Bert Fristedt and Lawrence Gray. *A modern approach to probability theory.* Probability and its Applications, 1997.

[34] Thierry Gallay and Arnd Scheel. Diffusive stability of oscillations in reaction-diffusion systems. *Trans. Amer. Math. Soc.*, 363(5):2571–2598, 2011.

[35] Nina Gantert, Yueyun Hu, and Zhan Shi. Asymptotics for the survival probability in a killed branching random walk. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):111–129, 2011.

[36] Nina Gantert, Sebastian Müller, Serguei Popov, and Marina Vachkovskaia. Survival of branching random walks in random environment. *J. Theoret. Probab.*, 23(4), 2010.

[37] Y. Git, J. W. Harris, and S. C. Harris. Exponential growth rates in a typed branching diffusion. *Ann. Appl. Probab.*, 17(2):609–653, 2007.

[38] Andreas Greven and Frank den Hollander. Branching random walk in random environment: phase transition for local and global growth rate. *Probab. Theory Related Fields*, 91(2):195–249, 1992.

[39] François Hamel and Lenya Ryzhik. Traveling fronts for the thermo-diffusive system with arbitrary Lewis numbers. *Arch. Ration. Mech. Anal.*, 195(3):923–952, 2010.

[40] J. M. Hammersley. Postulates for subadditive processes. *Ann. Probability*, 2:652–680, 1974.

[41] Robert Hardy and Simon C. Harris. A spine approach to branching diffusions with applications to $\mathcal{L}^p$-convergence of martingales. *Séminaire de Probabilités XLII*, 1979:281–330, 2009. Lecture Notes in Math., Springer, Berlin.

[42] J. W. Harris and S. C. Harris. Survival probabilities for branching brownian motion with absorption. *Electron. Comm. Probab.*, 12:81–92, 2007.

[43] J. W. Harris and S. C. Harris. Branching brownian motion with an inhomogeneous breeding potential. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):793–801, 2009.

[44] S. C. Harris and Matthew I. Roberts. The many-to-few lemma and multiple spines. *arXiv:1106.4761v2*, 2011.

[45] S. C. Harris and D. Williams. Large deviations and martingales for a typed branching diffusion. i. *Astrisque*, (236):133–154, 1996.

[46] Hadrain Heil, Makoto Nakashima, and Nobuo Yoshida. Branching random walks in random environment are diffusive in the regular growth phase. *Electron. J. Probab.*, 16(48):1316–1340, 2011.

[47] Yueyun Hu and Zhan Shi. Slow movement of random walk in random environment on a regular tree. *Ann. Probab.*, 35(5):1978–1997, 2007.

[48] Yueyun Hu and Zhan Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2), 2009.

[49] Yueyun Hu and Nobuo Yoshida. Localization for branching random walks in random environment. *Stochastic Process. Appl.*, 119(5):1632–1651, 2009.

[50] Torrey Johnson and Edward C. Waymire. Tree polymers in the infinite volume limit at critical strong disorder. *J. Appl. Probab.*, 48(3), 2011.

[51] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probability*, 3(5), 1975.

[52] A. Kolmogorov, I. Petrovsky, and N. Piscounov. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Moscou Universitet. Bull. Math.*, (1), 1937.

[53] Leonid Koralov. Branching diffusion in inhomogeneous media. *arXiv:1107.1159v2*, 2011.

[54] A. E. Kyprianou. A note on branching lvy processes. *Stochastic Process. Appl.*, 82(1):1–14, 1999.

[55] A. E. Kyprianou. Martingale convergence and the stopped branching random walk. *Probab. Theory Related Fields*, 116(3):405–419, 2000.

[56] A. E. Kyprianou. Asymptotic radial speed of the support of supercritical branching brownian motion and super-brownian motion in ^{d}. *Markov Process. Related Fields*, 11(1):145–156, 2005.

[57] Ka-Sing Lau. On the nonlinear diffusion equation of Kolmogorov, Petrovsky, and Piscounov. *J. Differential Equations*, 59(1):44–70, 1985.

[58] Quansheng Liu. Branching random walks in random environment. *ICCM*, 2:702–719, 2007.

[59] Yuan Lou, Wei-Ming Ni, and Linlin Su. An indefinite nonlinear diffusion problem in population genetics. II. Stability and multiplicity. *Discrete Contin. Dyn. Syst.*, 27(2):643–655, 2010.

[60] Roger Lui. A nonlinear integral operator arising from a model in population genetics i. monotone initial data. *Siam J. Math. Anal.*, 13(6), 1982.

[61] R. Lyons and Y. Peres. *Probability on Trees and Networks.* Cambridge University Press, 2011. In preparation. Current version available at http://mypage.iu.edu/˜rdlyons/.

[62] F. P. Machado and S. Yu. Popov. One-dimensional branching random walks in a markovian random environment. *J. Appl. Probab.*, 37(4):1157–1163, 2000.

[63] Colin McDiarmid. Minimal positions in a branching random walk. *Ann. Appl. Probab.*, 5(1), 1995.

[64] H. P. McKean. Application of brownian motion to the equation of kolmogorov-petrovskii-piskunov. *Comm. Pure Appl. Math.*, 28(3):323–331, 1975.

[65] Antoine Mellet, James Nolen, Jean-Michel Roquejoffre, and Lenya Ryzhik. Stability of generalized transition fronts. *Comm. Partial Differential Equations*, 34(4-6):521–552, 2009.

[66] A.A. Mogul'skii. Small deviations in a space of trajectories. *Theory Probab. Appl.*, 19(4):726–736, 1974.

[67] Gregoire Nadin, Benoît Perthame, and Lenya Ryzhik. Traveling waves for the Keller-Segel system with Fisher birth terms. *Interfaces Free Bound.*, 10(4):517–538, 2008.

[68] Kimie Nakashima, Wei-Ming Ni, and Linlin Su. An indefinite nonlinear diffusion problem in population genetics. I. Existence and limiting profiles. *Discrete Contin. Dyn. Syst.*, 27(2):617–641, 2010.

[69] Makoto Nakashima. Almost sure central limit theorem for branching random walks in random environment. *Ann. Appl. Probab.*, 21(1):351–373, 2011.

[70] James Nolen. A central limit theorem for pulled fronts in a random medium. *Netw. Heterog. Media*, 6(2):167–194, 2011.

[71] James Nolen. An invariance principle for random traveling waves in one dimension. *SIAM J. Math. Anal.*, 43(1):153–188, 2011.

[72] James Nolen and Lenya Ryzhik. Traveling waves in a one-dimensional heterogeneous medium. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(3):1021–1047, 2009.

[73] James Nolen and Jack Xin. A variational principle for KPP front speeds in temporally random shear flows. *Comm. Math. Phys.*, 269(2):493–532, 2007.

[74] Robin Pemantle. Search cost for a nearly optimal path in a binary tree. *Ann. Appl. Probab.*, 19(4):1273–1291, 2009.

[75] Robin Pemantle and Alan M. Stacey. The branching random walk and contact process on galton-watson and nonhomogeneous trees. *Ann. Probab.*, 29(4):1563–1590, 2001.

[76] Bruce Reed. The height of a random binary search tree. *J. ACM*, 50(3):306–332, 2003.

[77] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion.* Grundlehren der Mathematischen Wissenschaften, 293, third edition, 1999.

[78] Matthew I. Roberts. A simple path to asymptotics for the frontier of a branching brownian motion. *arXiv:1106.4771v3*, 2011.

[79] A. N. Shiryayev. *Probability.* Graduate Texts in Mathematics, 95., 1984.

[80] Frank Spitzer. *Principles of random walks.* Graduate Texts in Mathematics, 34, second edition, 1976.

[81] Edward C. Waymire and Stanley C. Williams. T-martingales, size biasing, and tree polymer cascades. *Appl. Numer. Harmon. Anal.*, pages 353–380, 2010.