UNIVERSITY OF OXFORD

DEPARTMENT OF STATISTICS

Two studies on branching processes

A thesis submitted in partial fulfilment of the requirements for the degree
Doctorate of Philosophy in Statistics

October 2020

Roman Stasiński
St John's College

Department of Statistics
University of Oxford

Contents

Abstract

We consider two branching processes: a Galton-Watson process and a branching Brownian motion.

In Chapter 2 we study behaviour of a minimax recursion defined on Galton-Watson trees. This recursion corresponds to a two-player combinatorial game. The value associated to the root of a tree truncated at some level $2n$ corresponds to the probability that the first player wins a game played on such a truncated tree. We study convergence of the value at the root as $n \to \infty$, in particular we give distributional limits under suitable rescaling in various cases. We address also a question of endogeny, which can be seen as a property that holds if the game can be played close to optimally for large n by inspecting only the initial section of the tree.

In Chapter 3 we consider a branching Brownian motion in $\mathbb{R}^d$. We prove that there exists a random subset Θ of $\mathbb{S}^{d-1}$ of full measure such that the limit of the derivative martingale exists simultaneously for all directions in Θ almost surely. This allows us to define a random measure on $\mathbb{S}^{d-1}$ whose density is given by the derivative martingale. We argue that this measure should correspond to the distribution of the direction of the furthest particle and we prove this claim in dimension one.

In Chapter 4 we present conclusions and further research directions, in particular we discuss a conjecture on the form of the extremal point process of the multidimensional BBM.

Acknowledgements

I would like to thank my supervisors Professor Julien Berestycki and Professor James Martin for their guidance. I have learnt from them a lot during this exciting journey.

I am grateful for all the feedback and advice I received from Professor Alison Etheridge and Professor Andreas Kyprianou who examined my thesis, and also for numerous discussions I had with Professor Christina Goldschmidt, Michał Przykucki and Bastien Mallein. I would also like to thank Eleonora Kreačić and Mitch Gooding for answering all my questions at the beginning of my DPhil studies, and to all other friends I was fortunate to make during my time in Oxford.

It is also a good moment to look further back and to thank all the people that I met during my mathematical education and without whom I would never be where I am now. In particular I would like to thank Ilona Hajduk, Professor Michał Krych, Professor Agnieszka Bojanowska-Jackowska, Professor Jacek Miękisz and Professor Jacek Gondzio. I am grateful for all the inspiration and for encouraging me to work harder and aim higher.

Last but not least, I am immensely grateful to Emilka Pompe for being there to support me. I am glad that we embarked on this adventure together.

Chapter 1

Introduction

Rooted trees are a class of discrete mathematical objects of fundamental importance in computer science, combinatorics and probability. They can be used to model situations involving fragmentation, splitting etc. A simple example is presented in Figure 1: we start with an object N_1 (the 'root'), which then splits into three objects N_{11}, N_{12} and N_{13}, which then split again.

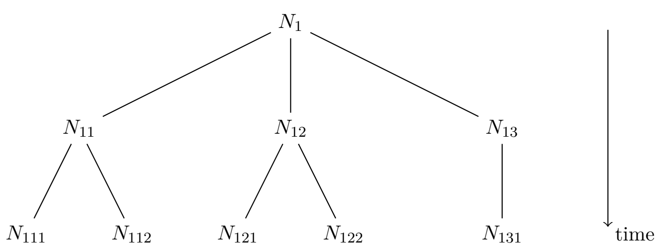

Figure 1.1: A tree rooted at N_1

In this work we shall consider random rooted trees and the basic assumption that we shall make is as follows: numbers of offspring of each node in the tree are i.i.d. random variables supported on $\{0, 1, 2, 3, \ldots\}$. Such an object is called a Galton-Watson tree – one can think of it as a simple example of a discrete time branching process. It is also a basic population model and therefore natural questions that arise are those of population growth and extinction, when time is indexed by successive generations. For example, if μ is the expected number of offspring generated by each individual, then the population goes extinct almost surely if and only if $\mu \leq 1$. We can also answer the question about growth: if Z_n is the number of individuals in generation n, then Z_n/μ^n is a non-negative martingale. Therefore, the martingale convergence theorem tells us that Z_n/μ^n converges almost surely.

Suppose we want to study a spread of population in its habitat. Then in addition to the genealogical structure of the population we need to record the position of each individual. We could model it by adding a spatial location to each node (which would reflect the position of the corresponding member of the population). We could further assume that each node splits into offspring nodes that are located at positions given by i.i.d. displacements with respect to the position of the parent node (see Figure 1.2).

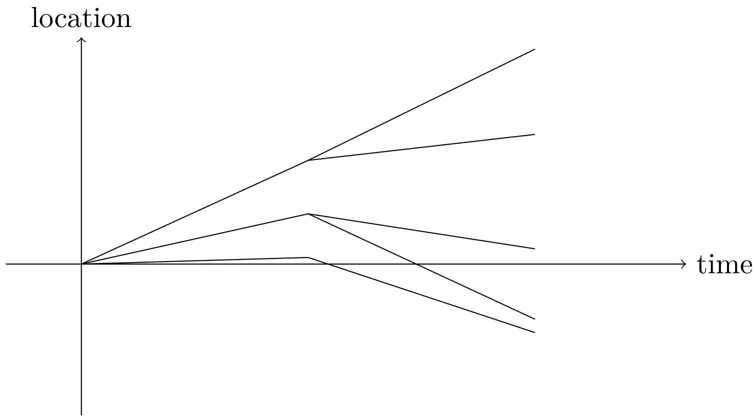

Figure 1.2: A branching random walk.

This model can be extended even further relaxing the assumption that splitting occurs at integer times. We assume instead that each node (now called a *particle*) lives for an exponential amount of time during which it moves according to a Brownian motion (see Figure 1.3). This model is known as a branching Brownian motion (BBM). In its standard version, the BBM is assumed to only have binary splitting, i.e. each particle gives birth to two new particles located at the final position of the trajectory of their parent.

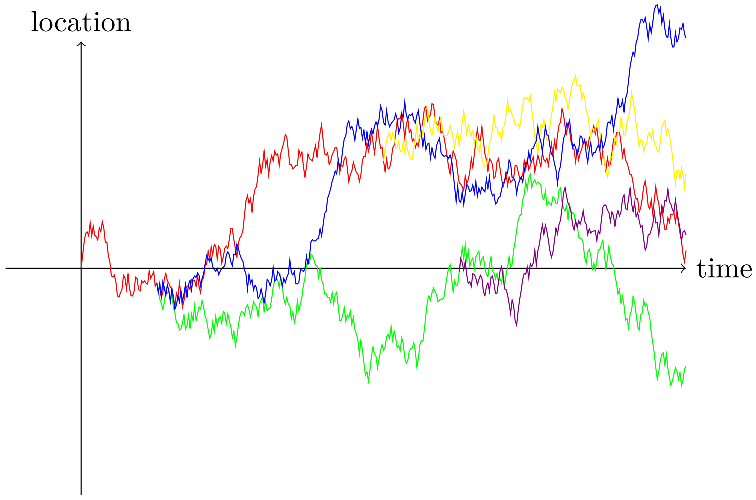

Figure 1.3: A branching Brownian motion.

The rest of this chapter gives a brief overview of the main results in this thesis. There are two main chapters, corresponding to two distinct projects. In Chapter 2 we study a minimax recursion associated with Galton-Watson trees and in Chapter 3 we prove weak convergence of a family of direction-wise derivative martingales in a multidimensional branching Brownian motion. Both chapters are based on submitted articles: Chapter 2 is based on [51] (accepted to *Combinatorics, Probability and Computing*) and Chapter 3 is based on [63] (accepted to *Annales de l'Institut Henri Poincaré*). In Chapter 4 we discuss further research directions.

1.1 Minimax trees and their connection to games on graphs

Random trees are often used as a structure on which a random process might be built, or as an approximation to more complicated random graphs.

In the first case, a process may be built on a tree by assigning random values to each node and imposing some recursive structure upon them. Let $\mathcal{T}^{(k)}$ be the set of nodes of a rooted tree that are located at level k, i.e. are separated from the root by k edges. Fix $n > 0$ and erase all nodes at levels larger than n (we will refer to such an operation as 'truncating at level n'). Denote by $\mathcal{L}^{(n)}$ the set of leaves of the resulting tree, i.e. the set of nodes with no descendants (note that the leaves can be located at all levels $0 < k \leq n$) and assign to each $u \in \mathcal{L}^{(n)}$ an i.i.d. random variable $X^{(u)}$. For each $u \in \mathcal{T}^{(k)} \setminus \mathcal{L}^{(n)}, k < n$ we define

$$X^{(u)} := g(X^{(v)}, v \in \mathcal{T}^{(k+1)}, u \prec v), \tag{1.1}$$

where $u \prec v$ denotes the set of offspring of node u.

This allows us to calculate recursively, starting from the leaves and backtracking to the root, all the values in the tree. Note that if the tree we are considering is a Galton-Watson process, then for each $k \leq n$ we obtain that $(X^{(u)})_{u \in \mathcal{T}^{(k)}}$ are i.i.d. random variables. This is a starting point for many questions. For example, let X_n be the random variable corresponding to the value at the root of the tree truncated at level n. Does then X_n converge in distribution as $n \to \infty$? Note that if $X_n \xrightarrow{d} X$, then X must satisfy a fixed point equation:

$$X = g(X^{(i)}, i \leq N), \tag{1.2}$$

where N is generated according to the offspring distribution of the Galton-Watson tree and $X^{(i)}$ are i.i.d. copies of X. A thorough survey of such models where the function $g(\cdot)$ is essentially a maximum or minimum was given by Aldous and Bandyopadhyay [4]. Other studied models in this class include the following ones (the list is non-exhaustive). Collet, Eckmann, Glaser and Martin [28] considered a binary tree and a function $g(x_1, x_2) = \max(x_1 + x_2 - 1, 0)$. A generalisation of this model to Galton-Watson trees was studied by Chen, Dagard, Derrida, Hu, Lifshits and Shi in [27] and [26]. Recursions where the function g takes values in a finite set and the tree is a critical Galton-Watson tree conditioned on size tending to infinity were investigated by Broutin, Devroye and Freiman [20].

In Chapter 2 we study minimax trees, which use the following propagation rule: at even levels the value at the node is equal to the minimum of the values at their children, at odd ones it is equal to the maximum. We shall assume that the random variables put at the leaves of the tree are distributed uniformly on $[0, 1]$ (note that there is nothing particularly special about uniform boundary conditions; by a simple rescaling we can map between this case and the case of i.i.d. boundary values from any other continuous distribution). If by W_{2n} we denote the value at the root of a Galton-Watson tree with offspring distribution concentrated on positive integers and truncated at level $2n$, then we obtain the following distributional recursion:

$$W_{2n} \overset{d}{=} \min_{1 \le i \le M} \max_{1 \le j \le M_i} W_{2n-2}^{(i,j)}. \tag{1.3}$$

Here M and $M_1, M_2, M_3, \ldots$ are i.i.d. draws from the offspring distribution, and $W_{2n-2}^{(i,j)}$, $i, j \in \mathbb{N}$, are i.i.d. copies of the random variable W_{2n-2}, independent of M and $\{M_i\}$. For an illustration of the structure of a minimax tree see Figure 1.4.

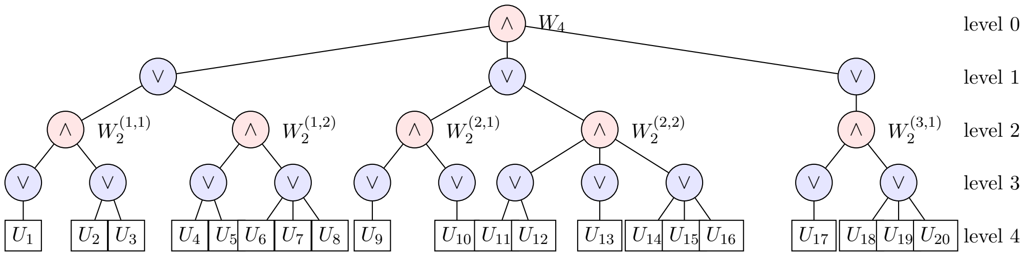

Figure 1.4: An example of a minimax tree with 4 levels. $U_1, \ldots, U_{20}$ are i.i.d. random variables which were assigned to the leaves of the tree. All non-leaf nodes have 1, 2 or 3 children. $\wedge$ and $\vee$ correspond to taking a minimum and maximum respectively.

This minimax procedure summarised in equation (1.3) has the following natural interpretation in terms of a two-player game on finite rooted trees with values assigned to the leaves of the tree. Two players alternate turns: a token starts at the root of the tree and each move consists in moving the token along an edge of the tree that has not been used yet (i.e. each move increases the distance between the token and the root). The game stops when the token reaches one of the leaves, and the outcome of the game is given by the value corresponding to that leaf. Player 1 is trying to minimise the outcome, whereas player 2 is trying to maximise it. W_{2n} corresponds to the outcome of the game under 'optimal play'.

One can observe (see equation (2.3)) that the cumulative distribution function of W_{2n} satisfies the following identity:

$$\mathbb{P}(W_{2n} \le x) = f(\mathbb{P}(W_{2n-2} \le x)), \tag{1.4}$$

with $f(s) = R(R(s))$, $R(s) = 1 - G(s)$ and $G(s)$ the probability generating function of the offspring distribution.

In Chapter 2 we describe possible limits of W_{2n}, generalising results of Pearl [57]. We say that a fixed point q of f is *unstable from the right* if $q < 1$ and $\lim_{\epsilon\to 0}\lim_{n\to\infty} f^n(q+\epsilon) > q$; similarly *unstable from the left* if $q > 0$ and $\lim_{\epsilon\to 0}\lim_{n\to\infty} f^n(q-\epsilon) < q$. We say that a fixed point is *stable* (from one or both sides) if it is not unstable (from the corresponding sides). We obtain the following theorem:

Theorem 2.2. *$W_{2n} \xrightarrow{d} W$ as $n \to \infty$, for some random variable W. There are two cases.*

(a) If f is the identity function, then $W_{2n} \sim U[0,1]$ for all n.

(b) Otherwise, let Q be the set of fixed points of f that are unstable from at least one side. Then W is discrete and has atoms precisely at the elements of Q.

For $q \in Q$, define

$$q_- = \begin{cases} \sup\{x : x < q, x = f(x)\}, & \text{if } q > 0 \text{ and } q \text{ is unstable from the left} \\ q & \text{otherwise} \end{cases}$$
$$q_+ = \begin{cases} \inf\{x : x > q, x = f(x)\}, & \text{if } q < 1 \text{ and } q \text{ is unstable from the right} \\ q & \text{otherwise} \end{cases}.$$

Then $\mathbb{P}(W = q) = q_+ - q_-$.

Note that when each node has almost surely exactly one child, then f is the identity function. Perhaps surprisingly, this is not the only such situation – another example is a geometric offspring distribution (see Section 2.2.2). The question of describing the class of distributions for which f is the identity is one of open problems mentioned in Chapter 2.

We consider also fluctuations around atoms of limiting distributions obtained in Theorem 2.2. This generalises the results of Ali Khan, Devroye and Neininger [5] from the case of regular trees to the case of Galton-Watson trees. Denote the probability mass function of the offspring distribution of the considered tree by $p_1, p_2, p_3, \dots$ (recall that we assumed that each node has at least one child almost surely). We obtain the following result:

Theorem 2.3. *Assume that f is not the identity function and let Q be the set of fixed points of f unstable from at least one side.*

Let $q \in Q$. Define q_- and q_+ as in Theorem 2.2, and let $\xi = f'(q)$. Then:

(a) If $1 < \xi < \infty$, then

$$\mathcal{L}\left(\xi^n (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}(V) \text{ as } n \to \infty,$$

where V is a random variable with a continuous distribution function.

(b) Suppose $\xi = 1$, and $k \geq 2$ is such that $f^{(r)}(q) = 0$ for $1 < r < k$ and $f^{(k)}(q) \neq 0$. Then

$$\mathcal{L}\left(n^{\frac{1}{k-1}} (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}(V),$$

where for $a = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}}$ we have $V = \begin{cases} a & w.p.\ \frac{q_+ - q}{q_+ - q_-} \\ -a & w.p.\ \frac{q - q_-}{q_+ - q_-} \end{cases}.$

(c) If $\xi = \infty$, then $q \in \{0, 1\}$. Assume now that

$$\mathbb{E}(M \mathbb{I}_{M \leq n}) = \sum_{k=1}^{n} k p_k \sim c n^{\rho} \text{ as } n \to \infty \tag{1.5}$$

for some $c > 0$ and $\rho \in (0, 1)$, where M is distributed according to the offspring distribution of the Galton-Watson tree, and let $K = \min\{i : p_i \neq 0\}$. Then $K < 1/(1-\rho)$, and $|f(t) - q| \sim C|t - q|^{K(1-\rho)}$ as $t \to q$ (i.e. $t \downarrow 0$ if $q = 0$, and $t \uparrow 1$ if $q = 1$) for some $C > 0$. Moreover,

$$\mathcal{L}(-[K(1-\rho)]^n \log |W_{2n} - q| \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}(Y),$$

where Y is a random variable such that $\mathbb{P}(Y \in (0, \infty)) = 1$.

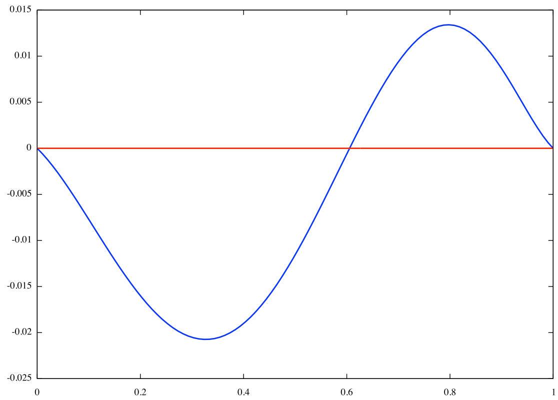

Figure 1.5: An example of a situation described by Theorem 2.2(b) and Theorem 2.3(a). The plotted function $f(x) - x$ for $x \in [0, 1]$ corresponds to an offspring distribution supported on $\{1, 3\}$ with $p_1 = 0.45 = 1 - p_3$. The point 0 is stable from the right and the point 1 is stable from the left. The set Q consist of one point q, which is unstable from both sides, located in $(0, 1)$ with $q_- = 0$ and $q_+ = 1$. Moreover, $f'(q) > 1$, hence an exponential scaling from Theorem 2.3 applies.

We also discuss a concept of *endogeny*. If instead of putting uniform random variables at the leaves we put Bernoulli$(1 - x^*)$ random variables, where x^* is a fixed point of f, then this distribution is preserved when propagated in the tree, i.e. $W_{2k} \stackrel{d}{=} W_{2n}$ for all k, n. To see this assume that $\mathbb{P}(W_{2n} = 0) = x^*$, which clearly holds for $n = 0$. From (1.4) we obtain that

$$\mathbb{P}(W_{2n+2} = 0) = \mathbb{P}(W_{2n+2} \le x^*) = f(\mathbb{P}(W_{2n} \le x^*)) = f(\mathbb{P}(W_{2n} = 0)) = f(x^*) = x^*. \quad (1.6)$$

Consider a tree of depth $2n$ (given by the Galton-Watson tree corresponding to the function f, truncated at level $2n$) with the terminal values drawn independently from Bernoulli$(1-x^*)$. Then (1.6) implies that the distribution of the value at the root is also Bernoulli$(1 - x^*)$. More generally, consider the structure of the first k levels of the tree: the distribution of these first k levels is the same for any n (such that $k \le 2n$).

As a result, by Kolmogorov's extension theorem we may construct an infinite tree with Bernoulli$(1-x^*)$ marginals at each node. The question one may ask now is when the Bernoulli random variable at the root is measurable with respect to the tree structure – if this is the case, we call the process 'endogenous'. Endogeny can be interpreted as follows: if one observes the structure of a large neighbourhood of the root, but not the values at the nodes, then the value at the root is still a random variable. If one can predict it with high probability then endogeny holds. In the case of our game-on-tree interpretation, endogeny means that the first player can play (on average) close to optimally by inspecting just the structure of the first few levels of the tree. Therefore, studying such properties (i.e. relative importance of the local tree structure) can be used to understand which tree-search algorithms converge quickly towards good lines of play.

We give a simple characterisation of endogeny in our model.

Theorem 2.5. *Let $x^* \in (0,1)$ be a fixed point of f, and consider the stationary recursive tree process with Bernoulli($1 - x^*$) marginals for the values at even levels. The process is endogenous if and only if $f'(x^*) \leq 1$.*

As we mentioned before, random trees can be used as an approximation to random graphs in certain situations. This is because many random graph models exhibit a local tree-like structure, i.e. when the graph is large enough, an observer placed in one of its vertices with high probability is going to see a tree as their neighbourhood. One can in general consider a sequence of graphs and use the notions of *local weak convergence* to describe the limiting behaviour of the sequence. A great introduction to this topic with many insightful examples was given by Aldous and Steele [3].

Let us recall the definition of local weak convergence. By m-neighbourhood of a vertex v in a graph G we shall denote a subgraph of G, where all vertices that cannot be connected to v by a path of at most m edges are removed. Denote by $\mathcal{G}$ the set of all locally finite connected rooted graphs (by a *rooted graph* we mean a graph with one distinguished node) considered up to a rooted isomorphism (i.e. a graph isomorphism preserving the root). We say that in $\mathcal{G}$ a sequence of rooted graphs (G_n, v_n) *converges locally* to a rooted graph (G, v), if for every m there exists n_m such that for all $n > n_m$ the m-neighbourhoods of v_n in G_n, and v in G are isomorphic. One can equip $\mathcal{G}$ with a metric inducing this converge, and which turns $\mathcal{G}$ into a complete separable metric space. We can now use the standard notion of weak convergence on Polish spaces. We denote the space of probability measures on $\mathcal{G}$ by $\mathcal{P}(\mathcal{G})$, and we associate with a finite graph G a probability measure $\mathcal{U}(G) \in \mathcal{P}(\mathcal{G})$ corresponding to choosing a vertex v uniformly at random to be the root of the graph, and restraining G to the connected component of v. We say that a sequence of finite graphs $(G_n)_{n\in\mathbb{N}}$ admits a *local weak limit* $\rho \in \mathcal{P}(\mathcal{G})$, if $\mathcal{U}(G_n)$ converges weakly to ρ.

Recall that a *matching* in a graph is a set of edges such that no two edges share a common vertex. A *maximum matching* is any matching of the maximum cardinality. It is known that in some cases asymptotic properties of a graph sequence may be deduced directly from the graph limit. One of the examples is the size of a maximum matching: Bordenave, Lelarge and Salez [15] proved that if a sequence of graphs admits a local weak limit ρ, then the proportion of vertices covered by a maximum matching in those graphs converges to a quantity that can be determined directly from ρ.

Minimax trees are related to limiting properties of a game called 'Trap' played on graphs [10] and further to the number of vertices belonging to every maximum matching of the graph on which the game is played. Before we make this vague statement precise, let us introduce the game itself.

Let $G = (V, E)$ be a simple finite graph, choose $v \in V$, and consider the following two-player game: a token starts at v and two players take turns moving it. Each move consists of moving the token along an edge of the graph and deleting the previous vertex and all neighbouring edges from the graph. A player looses when they cannot make any further moves (i.e. they are 'trapped'). Since the graph is finite, the game must end, and so one of the two players has a winning strategy. Let $\mathcal{W}$ be the set of starting vertices for which the Trap is a win for the first player. The following characterization of $\mathcal{W}$ (see for example Basu, Holroyd, Martin and Wästlund [10], Proposition 4) is known:

Theorem 1.1. *Let $G = (V, E)$ be a finite, connected, simple graph. 'Trap' on G starting from $v \in V$ is a win for the first player if and only if v is contained in all maximum matchings of G.*

Consider now a configuration model with a degree distribution $(b_k)_{k\geq 2}$, i.e. a random graph with n nodes generated as follows. First assign to each node a random number of half-edges following the graph degree distribution and then randomly match the half-edges to create full edges (possibly creating loops and multi-edges). It is known (see e.g. Montanari [53]) that as $n \to \infty$, such a graph sequence converges almost surely in the sense of local weak convergence to a *unimodular Galton-Watson tree*, i.e. an infinite tree where the root has the offspring distribution $(b_k)_{k\geq 2}$ and all other nodes have the offspring distribution given by the size-biased distribution shifted by 1:

$$\hat{b}_{k-1} := \frac{kb_k}{\sum_{i=2}^{\infty} kb_k}.$$

Note that we assumed that there are no vertices of degree one, thus the limiting tree is an infinite tree with no leaves. Now let us get back to the 'Trap' game played on a board being a realisation of such a configuration model with n vertices. Choose uniformly at random a starting vertex v. Consider removing v from the graph with all pendant edges, and then consider games started from each of the neighbours of v. Local weak convergence means that for each neighbour w of v, after removing v, the m-neighbourhood of w behaves like a Galton-Watson tree with offspring distribution $(\hat{b}_k)_{k\geq 1}$. Thus

$$\mathbb{P}(w \text{ wins in } G) = 1 - \mathbb{P}(\text{ all neighbours of } w \text{ win }) \approx \hat{R}(\mathbb{P}(\text{ neighbour of } w \text{ wins in } G)),$$

where $\hat{R}(x) := 1 - \sum_{k=1}^{\infty} \hat{b}_k x^k$.

If the outcome of the game is indeed decided locally (which corresponds to the endogenous situation), then by recursion the winning probability when starting from w should be close to x^*, where x^* is a stable fixed point of $\hat{f}$ and $\hat{f}(x) = \hat{R}(\hat{R}(x))$. Therefore, the 'Trap' game started from w may be seen approximately as an instance of the minimax tree with appropriate values at leaves – independent Bernoulli random variables with parameter x^*.

Finally, with high probability these m-neighbourhoods do not intersect, hence the winning probability starting from v should be approximately equal to $R(x^*)$, where $R(x) = 1 - \sum_{k=1}^{\infty} b_k x^k$ corresponds to the distribution of the offspring of the root in the limiting unimodular Galton-Watson tree (which is different from the distribution at higher levels).

If $\hat{f}$ has a single stable fixed point, this suggests that the proportion of vertices from which 'Trap' is a win for the first player (this proportion being equal to $|\mathcal{W}_n|/n$, where $|\mathcal{W}_n|$ denotes the size of the set $\mathcal{W}_n$) should almost surely converge to $R(x^*)$ as n tends to infinity. If f has multiple stable points, it may happen on the other hand that there is some global information (for example the parity of the largest component) that determines the distribution of win/loss events. This would lead to almost sure divergence, in the sense that almost surely the sequence of proportions of vertices appearing in every maximum matching would not converge as n tends to infinity.

1.2 Branching Brownian motion

Branching Brownian motion (BBM) has been studied using both probabilistic and analytic methods. It is closely related to the Fisher-Kolmogorov-Petrovsky-Piscounov (F-KPP) equation:

$$\begin{cases} \partial_t u = \frac{1}{2}\partial_{xx} u + g(u), \\ g(x) = x(x-1), \\ u(0,x) = \mathbb{I}_{x\geq 0}, \end{cases} \tag{1.7}$$

where $u(t,x) : \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$. This PDE was introduced and first studied by Fisher [32]. In the same year Kolmogorov, Petrovsky and Piscounov [42] studied a version of equation (1.7) with a more general form of the 'forcing term' $g(x)$. They showed that once appropriately centred by $m(t)$, the solution of equation (1.7) settles down to a 'travelling wave', i.e. there exists a function $\omega(x)$ such that

$$u\left(t, m(t) + x\right) \to \omega(x) \text{ as } t \to \infty \tag{1.8}$$

uniformly in x. The limiting shape $\omega(x)$ is the (unique up to a translation) solution of the ODE

$$0 = \frac{1}{2}\omega'' + \sqrt{2}\omega' + \omega(\omega - 1).$$

Moreover, they observed that the centering term $m(t)$ has to grow almost linearly, i.e. $m'(t) \to \sqrt{2}$ as $t \to \infty$.

The connection between the F-KPP equation and the BBM is often attributed to McKean [52], who observed the remarkable fact that the solution $u(t,x)$ of the F-KPP equation may be expressed in terms of the branching Brownian motion, but it was actually first proved by Skorohod [62] for a more general class of branching models. They derived the following Feynman-Kac representation:

$$u(t,x) = \mathbb{P}\left[\max_{u\in\mathcal{N}_t} X_t(u) \leq x\right],$$

where $\mathcal{N}_t$ is the set of particles alive at time t and $X_t(u)$ is the position of a particle u

at time t, when the BBM starts from one particle located at the origin. This observation, together with convergence (1.8), implies that $\max_{u\in\mathcal{N}_t} X_t(u)-m(t)$ converges in distribution. Moreover, to the first order of accuracy, the position of the rightmost particle of the BBM grows like $\sqrt{2}t$. Note that it does not imply that there is one single line of particles that is the right-most at all times. On the contrary, each particle is going to have one of their descendants in the lead at some point!

A more precise form of the centering term $m(t)$ (often referred to as the location of the front of the BBM) was given by Bramson [18] who established that for (1.8) to hold, $m(t)$ has to be of the form

$$m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \log t + b(t) + C, \tag{1.9}$$

where C is any constant and $b(t) = O(1)$. In the subsequent work [19] he showed that in fact $b(t) = o(1)$. Bramson also determined the right tail of the travelling wave solution ω of (1.7). He proved that there exists a constant $c_\star$ such that

$$1 - \omega(x) \sim c_\star x e^{-\sqrt{2}x}. \tag{1.10}$$

Bramson's proofs are rather analytical and later on alternative proofs of existence, asymptotics and uniqueness of the travelling wave solution to (1.7), using probabilistic methods alone, were obtained by Harris [36] and Kyprianou [43]. Further work on the fluctuations was carried out, among others, by Nolen, Roquejoffre and Ryzhik [55] who showed analytically that $b(t)$ in the equation (1.9) is of the form $-3\sqrt{\pi}/\sqrt{t} + O(t^{\gamma-1})$ for any $\gamma > 0$.

One of the key ideas behind the probabilistic analysis of the BBM was to study two particular martingales. The first one, called 'exponential additive martingale'

$$W_t^\lambda := \sum_{u\in\mathcal{N}_t} e^{\lambda X_t(u) - \frac{1}{2}\lambda^2 t - t}, \tag{1.11}$$

appears in work of Doney [29], Kingman [41] and McKean [52], but the first systematic study of this martingale is due to Biggins [12]. Clearly, W_t^λ is a non-negative martingale and hence it converges almost surely, but Biggins determined when the limit W_∞^λ of W_t^λ is non-trivial; he proved in particular that the critical martingale (with $\lambda = \sqrt{2}$) converges almost surely to 0. This implies that

$$0 = \lim_{t\to\infty} \sum_{u\in\mathcal{N}_t} e^{\sqrt{2}(X_t(u)-\sqrt{2}t)} \geq \limsup_{t\to\infty} e^{\sqrt{2}(\max_{u\in\mathcal{N}_t} X_t(u)-\sqrt{2}t)},$$

which means that almost surely for all large times all particles are below the line $\sqrt{2}t$. Therefore, one could kill particles at level $\sqrt{2}t + A$, and for A sufficiently large with high probability none of the particles would hit this barrier. This procedure is known as 'shaving' and an example of its application will be given later.

In [13] Biggins showed that in a suitable parameter set the convergence of the additive martingale is in fact uniform, and the limit is an analytic function of the parameter. Later on Lyons [45] gave a short proof of Biggins' theorem using a 'spine decomposition', while Hardy and Harris [35] considered the L^p convergence of W_t^λ.

The second important martingale for the BBM is the so-called 'derivative martingale' (the name comes from the fact that it is a derivative of W_t^λ with respect to λ taken at the critical value $\lambda = \sqrt{2}$):

$$Z_t := \sum_{u\in\mathcal{N}_t} (\sqrt{2}t - X_t(u))e^{\sqrt{2}(X_t(u)-\sqrt{2}t)}.$$

It was first studied by Neveu [54] and Lalley and Sellke [44] who proved that its limit Z_∞ exists and is positive almost surely. Another way of showing that Z_t converges was described by Kyprianou [43], who applied the aforementioned 'shaving' technique as follows. Kyprianou considered the process

$$Z_t^A := \sum_{u\in\mathcal{N}_t} (\sqrt{2}t - X_t(u) + A)e^{\sqrt{2}(X_t(u)-\sqrt{2}t)}\mathbb{I}_{\forall s\in[0,t], X_s(u)\leq\sqrt{2}s+A}.$$

He observed that for each A, Z_t^A is a non-negative martingale, hence converges almost surely to some Z_∞^A. On the other hand, for large A with high probability none of the particles ever reach the barrier $\sqrt{2}t + A$ and on this event $Z_t^A = Z_t + AW_t$. Since W_t converges almost surely to 0, on this event Z_t converges to Z_∞^A which proves that Z_t converges with probability one.

Recently fluctuations of Z_t around its limit Z_∞ were described by Maillard and Pain [49].

Further interesting results were obtained for the discrete analogue of the BBM: a branching random walk. This is a branching process characterised by two random variables: M and

D corresponding to the offspring and displacement distributions. It starts with one particle at the origin and is defined recursively as follows: a particle located at time n at position x, at time $n+1$ gives birth to M particles located at positions $x + D_1, x + D_2, \dots, x + D_M$, and all random variables used in the construction are independent (recall Figure 1.2). Hu and Shi [40], and Aïdékon and Shi [2] considered $W_n^{\sqrt{2}}$ – an analogue of the additive martingale for the branching random walk, defined as in (1.11). They proved that after an appropriate deterministic rescaling (found by Aïdékon and Shi to be $\sqrt{n}$), it converges in probability to $\sqrt{2/\pi} Z_\infty$, but the convergence does not hold in the almost sure sense, as $\limsup_{n\to\infty} \sqrt{n} W_n^{\sqrt{2}} = +\infty$ almost surely. The picture was completed by Hu [39], who proved that $\liminf_{n\to\infty} \sqrt{n} W_n^{\sqrt{2}} = \sqrt{2/\pi} Z_\infty$ a.s. Another connection between W_n^λ and Z_t was found by Madaule [48] who recovered Z_∞ as the limit of $W_\infty^\lambda / (\sqrt{2} - \lambda)$ with $\lambda \nearrow \sqrt{2}$.

Along the way several important methods applicable to general branching processes have been developed. In [54] Neveu introduced the idea of stopping lines, which generalise stopping times (see also [23] for more details). Another important tool, the spine decomposition, appeared in multiple articles, for example in [24] in the context of the branching Brownian motion or in [31] for measure-valued Markov processes, but had not received much attention until the seminal paper by Lyons, Pemantle and Peres [46]. It was subsequently extensively used, among others in [45], [43] and [37]. The main idea relies on the following observation: if one uses some non-negative mean one additive martingale M_t (for example W_t introduced earlier) to define a new probability measure $\mathbb{Q}$ by $\frac{d\mathbb{Q}_{|\mathcal{F}_t}}{d\mathbb{P}_{|\mathcal{F}_t}} = M_t$, where $(\mathcal{F}_n)$ is the filtration of the underlying BBM, then under the measure $\mathbb{Q}$ the BBM behaves as a process with one line of particles moving and branching differently to the rest – this line of particles is referred to as a 'spine'. A canonical application of this technique relies on the following fact (see e.g. Theorem 5.3.3. in [30] stated for discrete time): let $\mathcal{F}_\infty$ be the smallest σ-field containing all $\mathcal{F}_t$. Let $\mathbb{P}, \mathbb{Q}$ be two probability measures on $(\Omega, \mathcal{F}_\infty)$. Assume that for any t, $\mathbb{Q}_{|\mathcal{F}_t} \ll \mathbb{P}_{|\mathcal{F}_t}$ and let $X_t := \frac{d\mathbb{Q}_{|\mathcal{F}_t}}{d\mathbb{P}_{|\mathcal{F}_t}}$ and $X := \limsup_{t\to\infty} X_t$ which is $\mathbb{P}$-a.s. finite. Then

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(X \mathbb{I}_{\{A\}}) + \mathbb{Q}(A \cap \{X = \infty\}), \quad \forall A \in \mathcal{F}_\infty,$$

where $\mathbb{E}_{\mathbb{P}}$ denotes the expectation under $\mathbb{P}$. As a result, to prove uniform convergence of X_t under $\mathbb{P}$ it's enough to prove that X_t converges almost surely under $\mathbb{Q}$.

All the techniques mentioned above have been applied to describe more precisely the behaviour of the rightmost particle of the BBM. Lalley and Selke [44] proved that

$$\max_{u\in\mathcal{N}_t} X_t(u) - m_t - \frac{1}{\sqrt{2}}\log Z_\infty$$

converges in distribution towards a Gumbel random variable. They also conjectured that the empirical (time-averaged) distribution of the maximal displacement converges – this was confirmed by Arguin, Bovier and Kistler [8], who described the limit as the Gumbel distribution shifted by $\frac{1}{\sqrt{2}}\log Z_\infty$.

The next step was to describe the fluctuations of the position of the rightmost particle around the expected location. This question was first addressed in the case of the branching random walk: Hu and Shi [40], Aïdékon and Shi [2], and Hu [39] proved that

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{\log n}\left(\max_{u\in\mathcal{N}(n)} X_n(u) - \sqrt{2}n - \frac{3}{2\sqrt{2}}\log n\right) &= 0,\\
\limsup_{n\to\infty} \frac{1}{\log n}\left(\max_{u\in\mathcal{N}(n)} X_n(u) - \sqrt{2}n - \frac{3}{2\sqrt{2}}\log n\right) &= \frac{1}{\sqrt{2}}.
\end{aligned} \tag{1.12}$$

A simple proof of (1.12) in the BBM case was later given by Roberts [60].

Further research was done on the broader picture emerging around the position of the maximal particle. In particular, it was of interest to find out what could be said about the point measure of the properly centred positions of the extremal particles, i.e.

$$\sum_{u\in\mathcal{N}_t} \delta_{X_t(u)-m_t}.$$

It was recently shown in a series of papers by Arguin, Bovier and Kistler [6, 7, 9] and by Aïdékon, Berestycki, Brunet and Shi [1] that this centred point process converges in distribution towards a decorated Poisson point process (a Poisson point process where each atom is replaced by some point process). The underlying Poisson point process of this decorated Poisson point process has (random) intensity $c_\star Z_\infty e^{-\sqrt{2}x}\mathrm{d}x$, where $c_\star$ is the constant appearing in (1.10). The decoration is constructed as the limiting process of the BBM conditioned on the event that the maximal displacement exceeds $\sqrt{2}t$. These results were further enhanced by Bovier and Hartung [17], who proved the extended version of the convergence containing information about locations of particles in the underlying Galton-Watson tree.

Significant effort has been put into understanding of the one-dimensional BBM, but so far not much attention has been given to the multidimensional case. Currently only the approximate position of the furthest particle is known: Mallein [50] proved that

$$\left(\max_{u\in\mathcal{N}_t} ||X_t(u)|| - r_t, t \geq 0\right)$$

is tight, where $r_t = \sqrt{2}t + \frac{d-4}{2\sqrt{2}} \log t$. Note that $r_t = m_t + \frac{d-1}{2\sqrt{2t}} \log t + C + o(1)$ (see (1.9)). One could interpret this as follows: in dimension $d > 1$ at time t large enough we expect to see a cloud of points of radius $\sqrt{2}t - \frac{3}{2\sqrt{2}} \log t$ with spikes of height $\frac{d-1}{2\sqrt{2t}} \log t$ tossed uniformly at random. This result shows that the dimension has direct impact on the typical behaviour of the maximal displacement. It also shows that in dimensions $d > 4$ at time t we expect to see particles further away from the origin than $\sqrt{2}t$, which means that the standard 'shaving' procedure where we kill particles reaching level $\sqrt{2}t + A$ cannot be applied directly.

Note that a projection of the multidimensional BBM onto any line (i.e. the process $(X_t(u) \cdot \theta, u \in \mathcal{N}_t)$) is a one-dimensional BBM. Since the derivative martingale Z_t plays an important role in the description of the behaviour of the right-most particle in the one-dimensional case, it's therefore interesting to understand its direction-wise behaviour in the multi-dimensional case. This turns out to be a non-trivial problem: even though the limit of the derivative martingale of the projection of the BBM onto each direction θ, which we denote by $Z_t(\theta)$, exists almost surely, it does not imply that the process $(Z_t(\theta))_{\theta\in\mathbb{S}^{d-1}}$ seen as a process on the sphere converges. It nevertheless turns out to be true and this is stated as our main theorem of Chapter 3. For two measurable functions $f, g : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ we define

$$\langle f, g\rangle := \int_{\mathbb{S}^{d-1}} f(\theta)g(\theta)\sigma(\mathrm{d}\theta),$$

where $\sigma(\mathrm{d}\theta)$ stands for the surface measure of $\mathbb{S}^{d-1}$. We obtain the following result:

Theorem 3.3. *Almost surely there exists a measurable subset* Θ *of* $\mathbb{S}^{d-1}$ *of full measure (i.e.* $\sigma(\Theta) = \sigma(\mathbb{S}^{d-1})$*), such that* $Z_\infty(\theta) := \lim_{t\to\infty} Z_t(\theta)$ *exists for* $\theta \in \Theta$*, and for any bounded measurable function* f

$$\lim_{t\to\infty} \langle Z_t, f\rangle = \langle Z_\infty, f\rangle \ a.s., \tag{1.13}$$

writing $Z_\infty(\theta) = 0$ *for all* $\theta \notin \Theta$*. Additionally,* $0 < \lim_{t\to\infty}\langle Z_t, 1\rangle < \infty$ *almost surely.*

We conjecture that if $D(t)$ is the direction of the maximal particle at time t, then for a measurable set $B \subset \mathbb{S}^{d-1}$

$$\lim_{s\to\infty}\lim_{t\to\infty}\mathbb{P}(D(t)\in B \mid \mathcal{G}_s) = \int_B Z_\infty(\theta)\sigma(\mathrm{d}\theta) \Big/ \int_{\mathbb{S}^{d-1}} Z_\infty(\theta)\sigma(\mathrm{d}\theta)$$

We prove this claim in dimension one, with M_t^+, M_t^- denoting the position of the maximal and minimal particles at time t, and Z_∞^+, Z_∞^- being limits of the derivative martingales of the processes $(X_t(u), u \in \mathcal{N}_t)$ and $(-X_t(u), u \in \mathcal{N}_t)$ respectively. We state this results below:

Corollary 3.2. *We have*

$$\lim_{s\to\infty}\lim_{t\to\infty}\mathbb{P}\left(M_t^+ > -M_t^- \;\middle|\; \mathcal{G}_s\right) = \frac{Z_\infty^+}{Z_\infty^- + Z_\infty^+} \quad a.s.$$

Chapter 2

Minimax functions on Galton-Watson trees

Based on joint work with James Martin [51]

Abstract

We consider the behaviour of minimax recursions defined on random trees. Such recursions give the value of a general class of two-player combinatorial games. We examine in particular the case where the tree is given by a Galton-Watson branching process, truncated at some depth $2n$, and the terminal values of the level-$2n$ nodes are drawn independently from some common distribution. The case of a regular tree was previously considered by Pearl [57], who showed that as $n \to \infty$ the value of the game converges to a constant, and by Ali Khan, Devroye and Neininger [5], who obtained a distributional limit under a suitable rescaling.

For a general offspring distribution, there is a surprisingly rich variety of behaviour: the (unrescaled) value of the game may converge to a constant, or to a discrete limit with several atoms, or to a continuous distribution. We also give distributional limits under suitable rescalings in various cases.

We also address questions of *endogeny*. Suppose the game is played on a tree with many levels, so that the terminal values are far from the root. To be confident of playing a good first move, do we need to see the whole tree and its terminal values, or can we play close to optimally by inspecting just

the first few levels of the tree? The answers again depend in an interesting way on the offspring distribution.

We also mention several open questions.

2.1 Introduction

In this chapter we consider the behaviour of minimax recursions defined on random trees.

Consider a finite rooted tree with depth m. We will call the root 'level 0', the children of the root 'level 1', and so on. Suppose every node at levels $0, 1, \ldots, m-1$ has at least one child; the nodes at level m are all leaves. Suppose every leaf node (i.e. every node at level m) has some real value associated to it. Then recursively propagate the values towards the root in a minimax way: each node at an odd level gets a value which is the max of the values of its children, and each node at an even level gets a value which is the min of the values of its children.

This minimax procedure has a natural interpretation in terms of a two player game. Two players alternate turns; a token starts at the root, and a move of the game consists of moving the token from its current node to one of the children of that node. The leaf nodes are terminal positions; the outcome of the game is the value associated to the leaf node where the game ends. Player 1 is trying to minimise this outcome, and player 2 is trying to maximise it. The outcome of the game with 'optimal play' is the value associated to the root.

Suppose the terminal values are random, drawn independently from some common distribution.

Pearl [57] considered the case where the tree is regular (every non-leaf node has d children for some $d \geq 2$) and the terminal values are independent and uniformly distributed on the interval $[0, 1]$. For simplicity assume that the depth of the tree is even; write W_{2n} for a random variable representing the value at the root of a tree of depth $2n$. Pearl showed that W_{2n} converges in distribution to a constant as $n \to \infty$. This result was refined by Ali Khan, Devroye and Neininger [5], who derived an asymptotic distribution for W_{2n} after appropriate rescaling.

In this chapter we consider the case where the tree is given by a Galton-Watson branching process, truncated at level $2n$. This generalisation leads to a surprisingly rich variety of behaviour, depending on the offspring distribution of the branching process. For example, the limiting distribution of W_{2n} may be concentrated at a single point (as in the regular case), or may now have several atoms, or may even be continuous.

There is also a rich interplay between the two sources of randomness now present in the model (the tree itself, and the terminal values at the leaves). Suppose we play the game on a tree with many levels, so that the terminal values are far from the root. In order to be confident of playing a good first move, do we need to see the whole tree and terminal values, or can we play close to optimally by inspecting just the structure of the first few levels of the tree? Such questions can be formulated precisely in terms of the *endogeny* property for certain recursive tree processes, as introduced by [4]. The answers again depend in an interesting way on the offspring distribution.

Such questions concerning the relative importance of local tree structure and terminal values are of considerable interest in understanding the effectiveness of certain tree-search algorithms such as *Monte Carlo tree search* (MCTS) – see [22] for a survey. MCTS has famously been applied in recent years to games such as go, where it provided a considerable increase in playing strength [33] even before being allied with powerful deep learning techniques [61]. For some games, simple versions of these algorithms, without local evaluation functions, and with only very crude input from the terminal values (given for example by 'random rollouts' through unexplored parts of the tree), are nonetheless able to converge quickly towards good lines of play. Understanding which aspects of a game's structure make such convergence possible is an interesting challenge both in theoretical and in practical terms.

Our main results concerning distributional limits are presented in the next section. In Section 2.2 we discuss a range of examples and mention some open problems. The results about endogeny are given in Section 2.3. The main proofs are given in Section 2.4 and Section 2.5.

Before that we mention some recent related work. Broutin, Devroye and Fraiman [20] consider recursive distributional equations (including those of minimax type) defined on Galton-Watson trees conditioned to have a given total size n. Holroyd and Martin [38] consider minimax-type games (and various misère and asymmetric variants) defined on (perhaps infinite) Galton-Watson trees, with particular emphasis on the nature of phase transitions for the outcomes of the game as the underlying offspring distribution varies (see Section 2.2

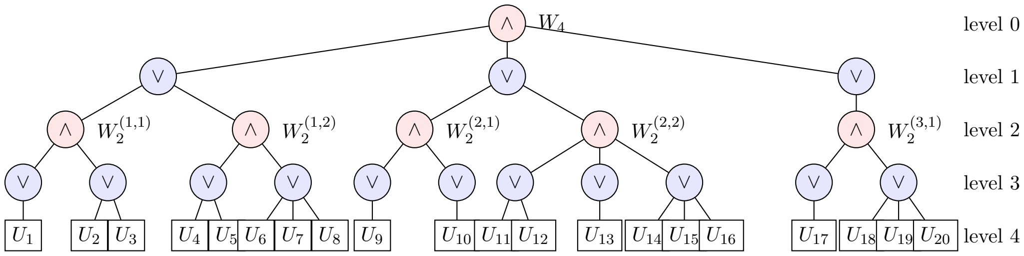

Figure 2.1: An example of a minimax tree, with 4 levels. Here all non-leaf nodes have 1, 2 or 3 children. $U_1, \dots, U_{20}$ are i.i.d. random variables uniform on $[0, 1]$.

for further comments). Note that in both [20] and [38], unlike in the case of this chapter, the offspring distribution puts positive weight at 0, so that there are leaves close to the root.

Similar questions arise in the context of random AND/OR trees and random Boolean functions. For example the model of Pemantle and Ward [58] involves a regular tree in which each node independently is a max or a min with equal probability; see Section 2.2.2 for comments on the relation to a particular case of our model. See for example Broutin and Mailler [21] for a variety of recent results in a more general setting, and many relevant references.

2.1.1 Main results

Consider a Galton-Watson tree with an offspring distribution with mass function $p_1, p_2, p_3, \dots$ on $\{1, 2, 3, \dots\}$ (note that every individual has at least one child). Let $G(x) = \sum_{k=0}^{\infty} p_k x^k$ be the probability generating function of the offspring distribution (which is a strictly increasing function mapping $[0, 1]$ to $[0, 1]$ bijectively). We will also write throughout

$$R(x) = 1 - G(x)$$

and

$$f(x) = R(R(x)). \tag{2.1}$$

Truncate the tree at level $2n$, so that all the vertices at level $2n$ are leaves. Let the terminal values associated to the leaves be i.i.d. uniform on $[0, 1]$ (independently of the structure of the tree). Recursively, assign values to the internal nodes of the tree (in particular, to the root) using the minimax procedure defined above. See Figure 2.1 for an illustration.

(Note that there is nothing particularly special about uniform boundary conditions. By a simple rescaling we can map between this case and the case of i.i.d. boundary values from any other continuous distribution. Later we will also consider discrete boundary values, for example those taking only values 0 and 1, where we can interpret 0 as a win for the first player, and 1 as a win for the second player).

We denote by W_{2n} the random variable associated with the root of a tree of depth $2n$. Then we have a distributional recursion:

$$W_{2n} \overset{d}{=} \min_{1\leq i\leq M} \max_{1\leq j\leq M_i} W_{2n-2}^{(i,j)}, \tag{2.2}$$

where M and $M_1, M_2, M_3, \ldots$ are i.i.d. draws from the offspring distribution, and $W_{2n-2}^{(i,j)}$, $i, j \in \mathbb{N}$, are i.i.d. copies of the random variable W_{2n-2}, independent of M and $\{M_i\}$.

Note that for any given i,

$$\begin{aligned}
\mathbb{P}\left(\max_{1\leq j\leq M_i} W_{2n-2}^{(i,j)} > x\right) &= 1 - \mathbb{P}\left(W_{2n-2}^{(i,j)} \leq x \text{ for } j = 1, \ldots, M_i\right) \\
&= 1 - \sum_m p_m \mathbb{P}\left(W_{2n-2} \leq x\right)^m \\
&= R\big(\mathbb{P}(W_{2n-2} \leq x)\big).
\end{aligned}$$

So from (2.2) we have

$$\begin{aligned}
\mathbb{P}\left(W_{2n} \leq x\right) &= \mathbb{P}\left(\min_{1\leq i\leq M} \max_{1\leq j\leq M_i} W_{2n-2}^{(i,j)} \leq x\right) \\
&= 1 - \mathbb{P}\left(\max_{1\leq j\leq M_i} W_{2n-2}^{(i,j)} > x \text{ for all } i = 1, \ldots, M\right) \\
&= 1 - \sum_m p_m \left[R\left(\mathbb{P}\left(W_{2n-2} \leq x\right)\right)\right]^m \qquad (2.3) \\
&= R\big(R\big(\mathbb{P}(W_{2n-2} \leq x)\big)\big) \\
&= f\big(\mathbb{P}(W_{2n-2} \leq x)\big).
\end{aligned}$$

Equation (2.3) implies that to describe the behaviour of W_{2n} for large n, we will be interested in the function f and in particular its fixed points.

We begin with the results for the case of a regular tree.

Theorem 2.1. *Suppose $p_d = 1$ for some $d \geq 2$.*

(a) *(Pearl [57])*

$$W_{2n} \xrightarrow{d} q \text{ as } n \to \infty,$$

where q is the unique fixed point in $(0,1)$ of the function $f_{d\text{-reg}}$ defined by

$$f_{d\text{-reg}}(x) = 1 - \left(1 - x^d\right)^d.$$

(b) *(Ali Khan, Devroye and Neininger [5]) Let $\xi = f'_{d\text{-reg}}(q)$. Then*

$$\xi^n \left(W_{2n} - q\right) \xrightarrow{d} W \text{ as } n \to \infty,$$

where W has a continuous distribution function F_W which satisfies

$$F_W(x) = f_{d\text{-reg}}(F_W(x/\xi)).$$

Now we will consider general offspring distributions. Since G is increasing and bijective as a function from $[0,1]$ to $[0,1]$, we have that $R = 1 - G$ is decreasing and bijective, and $f = R \circ R$ is again increasing and bijective. Also G is analytic on $[0,1)$, so that f is analytic on $(0,1)$.

We'll be particularly interested in fixed points of the function f. The function R itself has a single fixed point, which is obviously also a fixed point of f. Otherwise the fixed points of f come in pairs: if q is one then so is $R(q)$. One such pair are the points 0 and 1. We will say that a fixed point q of f is *unstable from the right* if $q < 1$ and $\lim_{\epsilon \to 0} \lim_{n \to \infty} f^n(q + \epsilon) > q$; similarly *unstable from the left* if $q > 0$ and $\lim_{\epsilon \to 0} \lim_{n \to \infty} f^n(q - \epsilon) < q$.

For a regular tree, Theorem 2.1 tells us that the distribution of W_{2n} converges to a constant. For general distributions, we still have convergence in distribution, but now we may have a 'genuinely random outcome' in the limit as the tree becomes large; the limiting distribution may have more than one atom (and in some surprising cases, the distribution of W_{2n} can simply be the same uniform distribution for all n).

Theorem 2.2. *$W_{2n} \xrightarrow{d} W$ as $n \to \infty$, for some random variable W. There are two cases.*

(a) *If f is the identity function, then $W_{2n} \sim U[0,1]$ for all n.*

(b) *Otherwise, let Q be the set of fixed points of f that are unstable from at least one side. Then W is discrete and has atoms precisely at the elements of Q.*

For $q \in Q$, define

$$q_- = \begin{cases} \sup\{x : x < q, x = f(x)\}, & \text{if } q > 0 \text{ and } q \text{ is unstable from the left} \\ q & \text{otherwise} \end{cases}$$

$$q_+ = \begin{cases} \inf\{x : x > q, x = f(x)\}, & \text{if } q < 1 \text{ and } q \text{ is unstable from the right} \\ q & \text{otherwise} \end{cases} . \tag{2.4}$$

Then $\mathbb{P}(W = q) = q_+ - q_-$.

It's not hard to show that $x \in Q$ if and only if $R(x) \in Q$. Hence again the atoms of the distributional limit W come in pairs, with the possible exception of the fixed point of R. In Section 2.2.1, we comment in particular on the case where W has atoms at 0 and 1.

For $q \in (0,1)$, we may write (2.4) alternatively as $q_- = \lim_{\epsilon \to 0} \lim_{n \to \infty} f^n(q - \epsilon)$ and $q_+ = \lim_{\epsilon \to 0} \lim_{n \to \infty} f^n(q + \epsilon)$ (this follows straightforwardly from the monotonicity and continuity of f).

In the next results we consider fluctuations around the atoms of the limiting distributions obtained in Theorem 2.2(b). The appropriate rescaling around a point $q \in Q$, where Q is defined in the previous theorem, depends on the derivative of $\xi = f'(q)$. If $q \in Q$ then we must have $\xi \geq 1$.

Theorem 2.3. *Consider the model defined by (2.2). Assume that f is not the identity function and let Q be the set of fixed points of f unstable from at least one side.*

Let $q \in Q$. Define q_- and q_+ as at (2.4), and let $\xi = f'(q)$. Then:

(a) If $1 < \xi < \infty$, then

$$\mathcal{L}\left(\xi^n (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}(V) \text{ as } n \to \infty,$$

where V is a random variable with a continuous distribution function.

(b) Suppose $\xi = 1$, and $k \geq 2$ is such that $f^{(r)}(q) = 0$ for $1 < r < k$ and $f^{(k)}(q) \neq 0$. Then

$$\mathcal{L}\left(n^{\frac{1}{k-1}} (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}(V),$$

where for $a = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}}$ *we have* $V = \begin{cases} a & w.p.\ \frac{q_+ - q}{q_+ - q_-} \\ -a & w.p.\ \frac{q - q_-}{q_+ - q_-} \end{cases}.$

(c) If $\xi = \infty$, then $q \in \{0, 1\}$. Assume now that

$$\mathbb{E}(M\mathbb{I}_{M\le n}) = \sum_{k=1}^{n} kp_k \sim cn^{\rho} \text{ as } n \to \infty \tag{2.5}$$

for some $c > 0$ and $\rho \in (0, 1)$, where M is distributed according to the offspring distribution of the Galton-Watson tree, and let $K = \min\{i : p_i \neq 0\}$. Then $K < 1/(1-\rho)$, and $|f(t) - q| \sim C|t - q|^{K(1-\rho)}$ as $t \to q$ (i.e. $t \downarrow 0$ if $q = 0$, and $t \uparrow 1$ if $q = 1$) for some $C > 0$. Moreover,

$$\mathcal{L}(-[K(1-\rho)]^n \log |W_{2n} - q| \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}(Y),$$

where Y is a random variable such that $\mathbb{P}(Y \in (0, \infty)) = 1$.

The scaling limits in part (a) are the closest ones to the result for the regular tree from Theorem 2.1. Note that when q is an endpoint of the interval, the limiting distribution V is now one-sided, supported on $(0, \infty)$ when $q = 0$ and on $(-\infty, 0)$ when $q = 1$.

For part (b), recall that f is analytic on $(0, 1)$ so certainly if $q \in (0, 1)$, such a k exists. Conceivably, there might be no such k in some cases where $q = 0$ or $q = 1$ (although we know of no example where analyticity fails at 0 or 1 except when the derivative is infinite).

On the other hand, many cases with $\xi = \infty$ are not covered by part (c). It seems challenging to describe all possible asymptotics; however, the assumption (2.5) is satisfied for an important class of power-law distributions with infinite mean, satisfying $\mathbb{P}(X > x) \sim x^{1-\alpha}$ with $\alpha \in (1, 2)$.

2.2 Examples, discussion and open questions

Our final main results, concerning the endogeny property, will be stated in Section 2.3. Before that, we discuss a variety of examples illustrating the results of Theorems 2.2 and 2.3.

First consider a case where each node has 1 or 3 children. This simple family already displays an interesting range of behaviours. Let $p_1 = p$ and $p_3 = 1 - p$, for $p \in [0, 1]$. In Figure 2.2, we plot the function $f(x) - x$ for $x \in [0, 1]$, for a variety of values of p. Fixed points of f correspond to zeros of the curve. A crossing from negative to positive corresponds to an unstable fixed point.

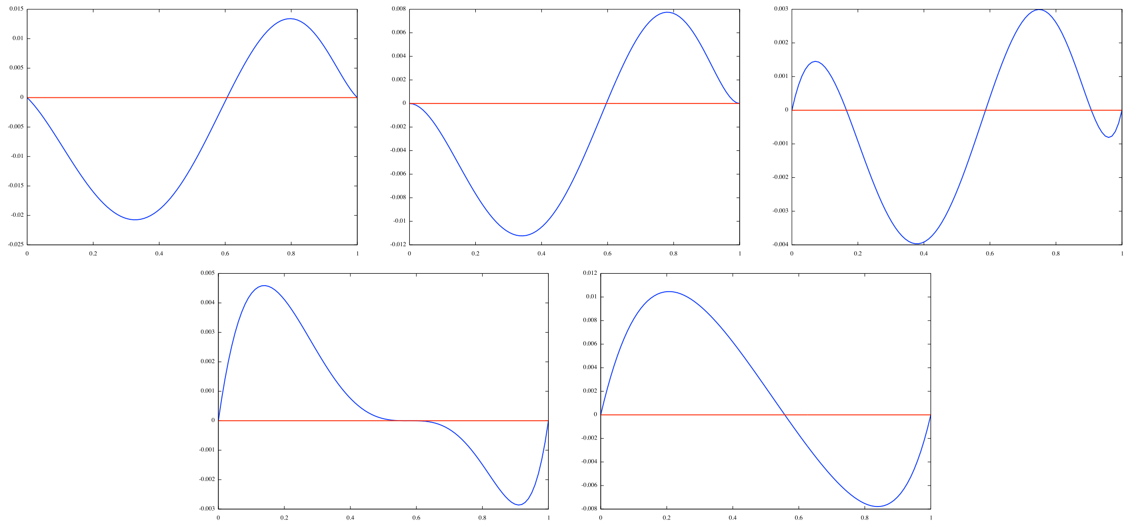

Figure 2.2: The function $f(x) - x$ for $x \in [0, 1]$ for the family of distributions with $p_1 = p$ and $p_3 = 1 - p$, with (a) $p = 0.45$, (b) $p = 0.5$, (c) $p = 0.55$, (d) $p = 0.598$, and (e) $p = 0.7$.

When $p < 0.5$, the points 0 and 1 are stable and there is a unique unstable fixed point in $(0, 1)$, just as in the case of a regular tree; W_{2n} converges to a constant. At $p = 0.5$, we have $f'(0) = f'(1) = 1$; the slope of $f'(x) - x$ at 0 and 1 is 0, but the points are still stable. For $p > 0.5$, the points 0 and 1 are unstable, and the limiting distribution of W in Theorem 2.2 puts positive mass at 0 and 1. At first, there is also positive mass at another fixed point in $(0, 1)$. However above a critical point at roughly $p = 0.598$, two of the fixed points disappear, leaving only a stable fixed point in $(0, 1)$, and the limiting distribution is concentrated only on the points 0 and 1.

Some further illustrative examples are shown in Figure 2.3. For these distributions (the

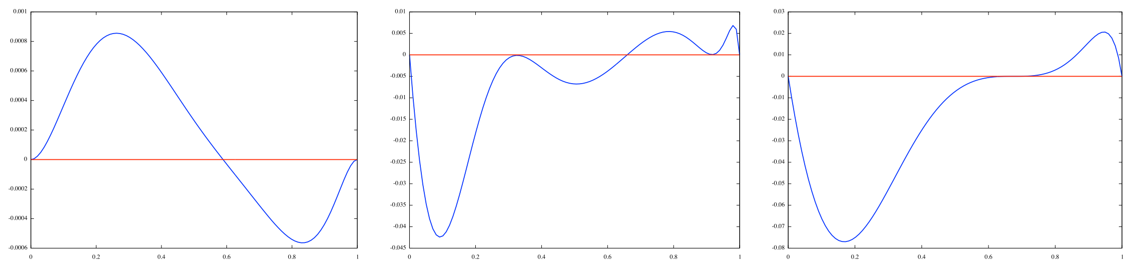

Figure 2.3: The function $f(x)-x$ for $x \in [0, 1]$ in three further cases: (a) $p_1 = 0.5$, $p_2 = 0.25$, $p_4 = 0.25$; (b) $p_2 = 0.783$, $p_{30} = 0.217$, and (c) $p_2 = 0.705$ and $p_{12} = 0.295$.

second and third are only approximate), we see points $q \in Q$ with $f'(q) = 1$, and so the rescalings of Theorem 2.3(b), which are polynomial rather than exponential, apply. In Figure 2.3(a) the relevant fixed points are at 0 and 1, and in Figure 2.3(a), they appear as 'touchpoints' in the graph of $f(x) - x$ and so are unstable from one side only; in these cases only one of the points a and $-a$ in Theorem 2.3(b) receives positive mass. By contrast, in Figure 2.3(c), the point of inflection gives a fixed point that is unstable on both sides.

In passing, we mention briefly another interesting aspect of some of the above examples, concerning phase transitions. As we vary the offspring distribution, we see points where, for example, the number of atoms of the limiting distribution W changes. Often, the transition can be *continuous* (in the topology given by the L^1 distance on the space of probabilistic measures on $\mathbb{N}$): as the offspring distribution is varied, an existing atom may split into several new atoms (as may happen when a point of inflection occurs as in Figure 2.2(d) or Figure 2.3(c)), or new atoms may appear whose weight grows continuously from 0 (such as happens at the points 0 and 1 in Figure 2.2(b)). On the other hand, we can also see *discontinuous* transitions in cases such as Figure 2.3(b); one can perturb the offspring distribution in an arbitrarily small way to remove the 'touchpoints' seen there, so that the atoms of W inside $(0, 1)$ disappear and all there mass jumps to the endpoints 0 and 1. Such ideas, expressed only vaguely here, are studied in a closely related context in [38].

2.2.1 Atoms at endpoints

The limiting distribution W in Theorem 2.2 may have atoms at 0 and 1. We note the following simple criterion:

Proposition 2.4. *Let μ be the mean of the offspring distribution.*

(i) If $p_1\mu < 1$ then $\mathbb{P}(W = 0) = \mathbb{P}(W = 1) = 0$.

(ii) If $p_1\mu > 1$ (including the case $p_1 > 0$ and $\mu = \infty$) then $\mathbb{P}(W = 0) > 0$ and $\mathbb{P}(W = 1) > 0$.

If $p_1\mu = 1$, or if $p_1 = 0$ and $\mu = \infty$, either case is possible. The proof of the result is straightforward. Since $f = R \circ R$ we have $f'(x) = R'(R(x))R'(x)$. Then since $R(0) = 1$ and $R(1) = 0$, and since $R' = -G'$, we have $f'(0) = f'(1) = G'(0)G'(1) = \mu p_1$ (assuming $\mu < \infty$); and we know that a fixed point q of f is an atom of W if $f'(q) > 1$, and not if $f'(q) < 1$.

There is a rather direct interpretation of the condition $p_1\mu > 1$ in terms of the Galton-Watson tree and the play of the game. Consider the set of paths in the tree, starting at the root, with the following property: every vertex along the path at an odd level has only one child. The union of these paths gives a subtree containing the root. For a vertex at an even level (such as the root), the expected number of grandchildren in the subtree is $p_1\mu$, since the vertex itself has an average of μ children, and each of those has precisely one child with probability p_1. Considering only even levels, this then gives a branching process with mean offspring $p_1\mu$; if $p_1\mu > 1$, then this branching process is supercritical and survives for ever with positive probability. In that case, by keeping the game within this tree, the first player can ensure that the second player never has any choice at all; all the second player's moves are forced. For the game truncated at level $2n$, the first player can choose between all the nodes at level $2n$ that are within the subtree; from this it can be shown that $\mathbb{P}(W = 0)$ is at least as big as the probability that this branching process survives.

2.2.2 The case $f(x) \equiv x$, and related open questions

Suppose the offspring distribution is such that f is the identity function. Then from (2.3), if we put independent values at the leaves from any given distribution, then the value at the root has that same distribution (hence the statement in Theorem 2.2(a)). Perhaps surprisingly, this property is not restricted to the trivial case where $p_1 = 1$.

Here are some families of examples where $f = R \circ R = (1 - G) \circ (1 - G)$ is the identity (i.e. R is an *involution*):

(a) Any geometric distribution. If $p_k = p(1-p)^{k-1}$ for $p \in (0, 1)$, then $G(x) = \frac{px}{1-(1-p)x}$ and so $R(x) = \frac{1-x}{1-(1-p)x}$, and one can check $f(x) = x$.

(b) Let $G(x) = \left[1 - (1-x)^{1/n}\right]^n$, for $n = 1, 2, 3, \ldots$. Via a binomial expansion, one can express G as a power series expansion with non-negative coefficients, and $G(1) = 1$, so G is indeed a probability generating function. The coefficient of x^k is non-zero for $k \geq n$.

(c) Let $G(x) = 1 - (1 - x^n)^{1/n}$, for $n = 1, 2, 3, \ldots$. Again G has a power series expansion with non-negative coefficients summing to 1. The coefficient of x^k is non-zero when k is a multiple of n.

These are far from the only cases. For a general source of examples, consider a function $S(x, y)$ from $[0, 1]^2$ to $[0, 1]$ that is symmetric, increasing in each coordinate, and has $S(1, 0) = S(0, 1) = 0$. If we define a function R by setting $S(x, y) = 0$ and writing $y = R(x)$ (which by the symmetry of S gives that $x = R(y)$), then R is indeed an involution: for any $x \in [0, 1]$,

$R(R(x)) = R(y) = x$. Some such functions R have power series expansions, and in some of those cases $G = 1 - R$ has all coefficients positive, as needed for a probability generating function. For example, $S(x, y) = y^2 + y + x^2 + x - 2 = 0$ gives $R(x) = [\sqrt{9 - 4x - 4x^2} - 1]/2$, in which case one can obtain straightforwardly that $G = 1 - R$ is a generating function.

We note several questions that it might be interesting to understand further:

(1) Can one describe in some nice way the class of all distributions for which f is the identity? For the class of examples described in the previous paragraph, can one describe nicely which functions $S(x, y)$ lead to functions R which have power series expansions, and then which of those yield a generating function G?

(2) Are the geometric distributions in example (a) above the only such distributions with finite mean? More generally, what types of tail decay are possible? For (a), the tail $\sum_{r=k}^{\infty} p_r$ of course decays exponentially in k, while for (b) and (c) it decays as $k^{-1/n}$.

To see the polynomial decay in cases (b) and (c) note that it's enough to show that $H(x) = 1 - (1 - x)^{1/n}$ is a generating function with a polynomially decaying tail, as then both $H(x^n)$ and $[H(x)]^n$ have this property. Now $H(x) = \sum_{i=1}^{n} p_j x^j$, where

$$
\begin{aligned}
p_j &= \frac{1}{nj} \frac{(1 - \frac{1}{n})(2 - \frac{1}{n}) \dots (j - 1 - \frac{1}{n})}{(j - 1)!} \\
&= \frac{1}{nj} \left(1 - \frac{1}{n}\right) \left(1 - \frac{1}{2n}\right) \dots \left(1 - \frac{1}{(j - 1)n}\right) \\
&= \frac{1}{nj} \exp\left[-\frac{1}{n}\left(1 + \frac{1}{2} + \dots + \frac{1}{j - 1}\right) + C_1 + o(1)\right] \\
&= \frac{1}{nj} \exp\left[-\frac{1}{n}(\log j + C_2) + o(1)\right] \\
&\sim C_3 j^{-1-\frac{1}{n}}
\end{aligned}
$$

for some C_3. Finally, $\sum_{j=k}^{\infty} p_j \sim C_4 k^{-\frac{1}{n}}$ for some C_4.

(3) Are there direct probabilistic arguments explaining the fact that f becomes the identity in these cases, in terms of the underlying process on the tree?

One case where it's possible to make such an argument is the $n = 2$ case in (b) above. Here p_k is the probability that the cluster containing the origin has size k for critical percolation on the binary tree (these coefficients are closely related to the Catalan numbers). Having made this identification, one can connect the minimax recursion on our random tree to an analogous recursion in the model studied by Pemantle and

Ward [58], of a binary tree in which each node independently is a max or a a min with probability 1/2 each.

We end this section with two further open questions about the form of f in more general cases:

(4) Can f have an arbitrarily large number of fixed points in $[0, 1]$?

(5) Can f have infinitely many fixed points in $[0, 1]$ (without being equal to the identity)? Since f is analytic on $(0, 1)$, this would require the set of fixed points to accumulate at 0 and at 1.

2.3 Endogeny

Suppose we play the game on a tree where the depth $2n$ is large, so that the boundary values are far from the root. To be confident of making a good first move, do we need to see a large part of the structure of the tree, and the boundary values? – or can we play close to optimally by inspecting just the structure of the first few levels of the tree? This is a so-called *endogeny* question [4]. The answer to this question again depends on the offspring distribution and the distribution of the boundary values.

To formalise the question, first define an operator on distributions corresponding to the recursion given by (2.2). For a distribution μ on $[0, 1]$, let $T(\mu)$ be the distribution of the left-hand side of (2.2) when the random variables $W_{2n-n}^{(i,j)}$ on the right-hand side are i.i.d. with distribution μ. Equivalently, rewriting (2.3), $T(\mu)[0, y] = f(\mu[0, y])$ for all y.

We will be interested in fixed points of T. For example, for offspring distributions such that f is the identity, *every* μ is a fixed point of T. For more general offspring distributions, whenever x is a fixed point of f, the Bernoulli distribution which puts mass x at 0 and $1 - x$ at 1 is a fixed point of T; for a game with Bernoulli terminal values, there are only two possible values of the outcome and we can interpret 0 as a win and 1 as a loss (from the perspective of the first player).

Suppose indeed that μ is a fixed point of the operator T. Consider a tree of depth $2n$ (given by the Galton-Watson tree truncated at level $2n$) with the terminal values drawn independently from μ. Then the distribution of the value at the root is also μ. More generally, consider the structure of the first k levels of the tree; the distribution of these first k levels is the same for any n (such that $k \leq 2n$).

As a consequence of this *consistency* over different values of n, we may let $n \to \infty$ and, applying Kolmogorov's extension theorem, obtain a distribution of the entire infinite Galton-Watson tree along with values attached to each node that obey the minimax recursions (min at even levels, max at odd levels).

This gives a *stationary recursive tree process* in the language of [4]. The relevant stationarity property is the following: condition on the structure of the first two levels of the tree, and write $v_1, \dots, v_r$ for the level-2 nodes. Conditional on the structure of the first two levels, the structure of the subtrees rooted at $v_1, \dots, v_r$, along with the values associated to the nodes of those subtrees, are given by r i.i.d. copies of the original tree process. (More precisely, we might describe the tree process as '2-periodic' rather than stationary, since even and odd levels differ; we can recover a genuinely stationary process by considering only even levels.)

For a more formal and more general set-up, see for example [4] or [47].

We have defined a joint distribution of the structure of the tree and the values associated to each node of the tree. Now the recursive tree process is said to be *endogenous* if the value associated to the root is measurable with respect to the structure of the tree. Note that for the same offspring distribution, this endogeny property may hold for some fixed point distributions μ and not for others.

Being measurable with respect to the structure of the tree is equivalent to being approximable to any given degree of accuracy using the information only of some finite portion of the tree. That is, for any random variable X (in particular, the root value), X is measurable with respect to the structure of the tree if, for any x and any $\epsilon > 0$, there exists k such that with probability $1 - \epsilon$, the conditional probability of the event $\{X \leq x\}$, given the structure of the first k levels of the tree is in $[0, \epsilon] \cup [1 - \epsilon, 1]$, where X denotes the value at the root.

For a more concrete interpretation, we can concentrate only on the case of finite trees, truncated at some level $2n$. Then the property in the previous paragraph can be reformulated to say that the value at the root can be approximated arbitrarily closely using information from the structure of some appropriate number of levels at the top of the tree, *uniformly* in the value of n.

Note that endogeny does *not* indicate that the value at the root is insensitive to arbitrary changes in the boundary conditions. In our case, that would be trivially false. Rather, for a given distribution of boundary conditions, endogeny expresses the property that, if the boundary is far away, the root is typically not much affected by the difference between various realisations drawn from that distribution. In particular, endogeny may hold for some boundary distributions and not for others, as is indeed the case for our model.

Consider in particular the Bernoulli ('win/loss') boundary conditions described above.

Theorem 2.5. *Let $x \in (0,1)$ be a fixed point of f, and consider the stationary recursive tree process with Bernoulli($1-x$) marginals for the values at even levels. The process is endogenous if and only if $f'(x) \leq 1$.*

So, approximately speaking, the endogenous processes with Bernoulli marginals correspond to the *stable* fixed points of the function f, which are those fixed points that do *not* appear as atoms in the distribution of the limiting random variable W in Theorem 2.2. (An exception may occur when the derivative of f at a fixed point is precisely 1; further, in the cases $x = 0$ and $x = 1$ the values are constant and the process is trivially endogenous.)

To prove Theorem 2.5, we use a characterisation of endogeny in terms of uniqueness of bivariate distributions, introduced by Aldous and Bandyopadhyay in [4] and proved in somewhat more generality by Mach, Sturm and Swart [47]. See Section 2.5 for details.

For offspring distributions where f is the identity, any distribution μ gives rise to a recursive tree process. In particular, we can take μ to be the uniform distribution on $[0,1]$, as we did in previous sections. We have the following corollary of Theorem 2.5:

Corollary 2.6. *Suppose f is the identity. Then for any μ, the recursive tree process with marginals μ for the values at even levels is endogenous.*

2.4 Proofs: convergence and scaling limits

Proof of Theorem 2.2. From equation (2.3) and the monotonicity of f we see that

$$\lim_{n\to\infty} \mathbb{P}(W_{2n} \leq x) = \lim_{n\to\infty} f^n(x)$$

exists for all x, and therefore W_{2n} indeed converges in distribution as $n \to \infty$, to a limit W with a distribution function $F_W(x) = \lim_{n\to\infty} f^n(x)$.

Part (a) is immediate from (2.3). For part (b), note that since f is analytic in $(0,1)$ and f is not the identity function, the set of fixed points of f cannot have an accumulation point in $(0,1)$. Therefore, this set of fixed points of f defines a partition of the interval

$(0,1)$ into a disjoint union of open intervals plus the set of fixed points, each of which is an endpoint of exactly two intervals from the partition. Since f is monotone and continuous, $F_W(x) = \lim_{n\to\infty} f^n(x)$ is constant on those intervals; therefore W can have atoms only at fixed points of f.

Suppose $q \in (0,1)$ is such a fixed point. Then

$$\begin{aligned}\mathbb{P}(W = x) &= \lim_{\epsilon\to 0} \mathbb{P}(q - \epsilon < W \le x + \epsilon) \\ &= \lim_{\epsilon\to 0} \lim_{n\to\infty} f^n(q+\epsilon) - \lim_{\epsilon\to 0} \lim_{n\to\infty} f^n(q-\epsilon).\end{aligned} \tag{2.6}$$

Since f is monotone and continuous, the quantity above is equal to 0 precisely if and only if the fixed point q is stable. Hence indeed W has an atom at q precisely if q is unstable from at least one side. As commented immediately after Theorem 2.2, the right-hand side of (2.6) is equal to $q_+ - q_-$, as required.

The cases where $q = 0$ or $q = 1$ follow in a similar way. □

The rest of this section is devoted to the proof of Theorem 2.3.

2.4.1 Proof of Theorem 2.3(a): $1 < f'(q) < \infty$

Firstly we assume that q is the unique fixed point of f in $(0,1)$ and that it is unstable from both sides. In the second part of the proof we show how Lemma 2.7 below implies the general case.

Suppose $\xi = f'(q) > 1$. An example of this case is presented in Figure 2.2(a), where $p_1 = 0.45$ and $p_3 = 0.55$.

We will prove the following result:

Lemma 2.7. *Consider the recursion (2.3) and assume that q is the unique fixed point of f in $(0,1)$ and that it is unstable from both sides. Set $\xi = f'(q)$. If $\xi > 1$, then*

$$\xi^n (W_{2n} - q) \xrightarrow{d} V \text{ as } n \to \infty,$$

where the distribution function F_V of V is continuous and satisfies

$$F_V(x) = f(F_V(x/\xi)), \quad x \in \mathbb{R}.$$

Lemma 2.7 extends the result of Ali Khan, Devroye and Neininger [5] to the case of random trees. Note that Lemma 2.7 corresponds directly to the part (a) of Theorem 2.3 for f having a unique fixed point in $(0,1)$ that is unstable, as then $q_- = 0$ and $q_+ = 1$.

Proof of Lemma 2.7. We follow the lines of [5] but in our case the analysis is slightly more complicated because of the more general form that f can admit. We first prove that there exists a pointwise limit of distribution functions of $\xi^n(W_{2n} - q)$, which is not identical to 0 or 1, and then show that it is continuous, which completes the proof. Define

$$g_n(x) = \mathbb{P}\left(\xi^n(W_{2n} - q) \le x\right), \quad x \in \mathbb{R}$$

Therefore, for each x for sufficiently large n (such that $0 \le q + \frac{x}{\xi^n} \le 1$),

$$g_n(x) = \mathbb{P}\left(W_{2n} \le q + \frac{x}{\xi^n}\right) = f^n\left(q + \frac{x}{\xi^n}\right).$$

Note that $g_n(0) = q$ for all n. We need some local uniform bound for g_n around $x = 0$. This will be supplied by the following lemma:

Lemma 2.8. *Under the assumptions of Lemma 2.7, let k be the smallest number larger than 1 such that $f^{(k)}(q) \neq 0$. Denote $h_1(x) = q + x$ and $h_2(x) = q + x + cx^k$ for $x \in \mathbb{R}$. Then there exist c and an $\varepsilon > 0$ such that for all n and $|x| \le \varepsilon$, either $h_1(x) \le g_n(x) \le h_2(x)$ or $h_2(x) \le g_n(x) \le h_1(x)$.*

Note that such a number k exists since we assumed that f is not the identity function and f is analytic at q.

Proof of Lemma 2.8. Take any c such that c has the same sign as $f^{(k)}(q)$ and $|c| > \left|\frac{f^{(k)}(q)}{k!\xi(\xi^{k-1}-1)}\right|$. From analyticity of f, $f^{(k)}(x)$ does not change sign on some neighbourhood of q.

For simplicity assume that k is even and $f^{(k)}(q) > 0$. We would generally need to consider four cases depending on the parity of k and the sign of $f^{(k)}(q)$. For the other three cases the steps of the proof of the lemma are identical modulo the change of sign in the inequalities.

The proof is by induction on n and makes use of Taylor's formula up to order k. For $n = 0$ the assertion is true, as we have

$$h_1(x) = q + x = g_0(x) \le h_2(x).$$

Note that the above holds for all ε, thus we will choose ε later. Assume now that

$$h_1(x) \leq g_{n-1}(x) \leq h_2(x)$$

for some $n-1 \geq 0$, $|x| \leq \varepsilon$ and $\varepsilon > 0$. Since $\left|\frac{x}{\xi}\right| \leq \varepsilon$, as $|x| \leq \varepsilon$ and $\xi > 1$, and f is increasing, we have

$$g_n(x) = f^n\left(q + \frac{x}{\xi^n}\right) = f\left(f^{n-1}\left(q + \frac{x/\xi}{\xi^{n-1}}\right)\right) = f\left(g_{n-1}\left(x/\xi\right)\right) \geq f\left(h_1\left(x/\xi\right)\right),$$

and analogously

$$g_n(x) \leq f\left(h_2\left(x/\xi\right)\right).$$

The induction proof will be completed if we can show that for some $\varepsilon > 0$, for $|x| \leq \varepsilon$,

$$h_1(x) \leq f\left(h_1\left(x/\xi\right)\right), \quad f\left(h_2\left(x/\xi\right)\right) \leq h_2(x). \tag{2.7}$$

Taking the Taylor expansion of f around q at points $q + x/\xi$ and $q + x/\xi + c\left(x/\xi\right)^k$, we obtain

$$\begin{aligned} f(h_1(x/\xi)) &= q + x + \frac{1}{k!}\frac{1}{\xi^k}\left(f^{(k)}(q)\right)x^k + o(x^k), \\ f(h_2(x/\xi)) &= q + x + \frac{1}{k!}\frac{1}{\xi^k}\left(f^{(k)}(q) + \xi k! c\right)x^k + o(x^k). \end{aligned}$$

Since

$$0 < \frac{1}{k!}\frac{1}{\xi^k}\left(f^{(k)}(q) + \xi k! c\right) = \frac{f^{(k)}(q)}{k!\xi(\xi^{k-1}-1)}\frac{\xi^{k-1}-1}{\xi^{k-1}} + c\frac{1}{\xi^{k-1}} < c$$

and by assumption $f^{(k)}(q) > 0$, we are therefore able to pick $\varepsilon > 0$ such that (2.7) holds for $|x| \leq \varepsilon$. This completes the proof of Lemma 2.8. □

Note that if we show that for some g, $g_n(x) \to g(x)$ for all x, then the above lemma will imply that $g(x)$ is continuous and differentiable at $x = 0$ with $g'(0) = 1$. We now claim that for each x, $(g_n(x))$ is a monotone sequence for $n > n_x$. This is implied by the following lemma:

Lemma 2.9. *Under the assumptions of Lemma 2.7, for each M there exists n_M such that for $|x| \le M$, $(g_n(x))$ is a monotone sequence for $n \ge n_M$.*

Proof of Lemma 2.9. As in the proof of Lemma 2.8, we consider the case where k is even and $f^{(k)}(q) > 0$ – the other cases are identical. Using Taylor expansion up to order k, there exists $\varepsilon > 0$ such that for $|y| \le \varepsilon$,

$$f(q+y) \ge f(q) + f'(q)y \tag{2.8}$$

(recall that $f^{(i)}(q) = 0$ for $1 < i < k$). Now let $n_M = \lceil \log_\xi \left(\frac{M}{\varepsilon}\right) \rceil$ and note that for any $|x| \le M$ and $n \ge n_M$, $|x/\xi^n| < \varepsilon$ and therefore by (2.8),

$$f\left(q + \frac{x}{\xi^n}\right) \ge f(q) + f'(q)\frac{x}{\xi^n} = q + \frac{x}{\xi^{n-1}}.$$

Finally, since f^{n-1} is monotone increasing,

$$g_n(x) = f^n\left(q + \frac{x}{\xi^n}\right) = f^{n-1}\left(f\left(q + \frac{x}{\xi^n}\right)\right) \ge f^{n-1}\left(q + \frac{x}{\xi^{n-1}}\right) = g_{n-1}(x).$$

This proves the claim. □

Since $g_n(x) \in [0,1]$ for all n, by Lemma 2.9 $g_n(x)$ converges for all x – we denote the limit by $g(x)$. The continuity of f and the fact that $g_n(x) = f(g_{n-1}(x/\xi))$ imply that

$$g(x) = f(g(x/\xi)). \tag{2.9}$$

Therefore, from the continuity of f and the monotonicity of g, $\lim_{x\to-\infty} g(x)$ and $\lim_{x\to\infty} g(x)$ are fixed points of f. Using the fact that $\{0, q, 1\}$ are the only fixed points of f, g is non-decreasing, $g(0) = q$ and $g'(0) = 1$, we deduce that $\lim_{x\to-\infty} g(x) = 0$ and $\lim_{x\to\infty} g(x) = 1$. When we show that g is continuous at all x, it will then imply that $F_V = g$, i.e. that the limiting random variable V has a continuous distribution function.

We apply the following strategy to show that g is continuous: we showed that $g(x)$ is continuous at 0 and now we show separately that it is continuous on some $(-\varepsilon, 0)$ and on some $(0, \varepsilon)$. The identity (2.9), together with the continuity of f, then implies that g is continuous on all of $\mathbb{R}$ (since $\xi > 1$).

We still work under the assumption that $f^{(k)}(q) > 0$, where $k \ge 2$ is such that $f^{(r)}(q) = 0$

for $1 < r < k$ and $f^{(k)}(q) \neq 0$, and that k is even (if k is odd then reasoning in the two cases below should be swapped). Note that to prove that g is continuous on some interval I, it is sufficient to show that

$$\sup_{y\in I}\sup_{n\geq 0} g_n'(y) < \infty. \tag{2.10}$$

By the chain rule we obtain

$$g_n'(x) = \frac{1}{\xi^n}\prod_{i=0}^{n-1} f'\left(f^i\left(q+\frac{x}{\xi^n}\right)\right). \tag{2.11}$$

We consider first $g_n'(y)$ for $y < 0$. Since $f^{(k)}(q) > 0$, there exists an $\varepsilon > 0$ such that $f'(q+y) < \xi$ for $y \in (-\varepsilon, 0)$. Since $f(q) = q$, f is increasing and $\xi > 1$, this implies that for $y \in (-\varepsilon, 0)$ and $i < n$

$$f^i\left(q+\frac{y}{\xi^n}\right) = f^{i-1}\left(f\left(q+\frac{y}{\xi^n}\right)\right) > f^{i-1}\left(q+\xi\frac{y}{\xi^n}\right) > \ldots > q+\xi^i\frac{y}{\xi^n} > q-\varepsilon.$$

Therefore $f^i\left(q+\frac{y}{\xi^n}\right) \in (q-\varepsilon, q)$, hence $f'\left(f^i\left(q+\frac{y}{\xi^n}\right)\right) < \xi$. By (2.11) we conclude that $g_n'(y) < 1$ for all n and $y \in (-\varepsilon, 0)$. This implies that (2.10) holds with $I = (-\varepsilon, 0)$, hence g is continuous on $(-\varepsilon, 0)$.

Now we turn to the case of $y > 0$. The function f is non-decreasing, and $\xi > 1$; hence for all $0 < i < n$ and all $0 \leq y \leq x$,

$$q \leq f^i\left(q+\frac{y}{\xi^n}\right) \leq f^i\left(q+\frac{x}{\xi^n}\right). \tag{2.12}$$

Note also that

$$f^i\left(q+\frac{x}{\xi^n}\right) \leq f^i\left(q+\frac{x}{\xi^i}\right) = g_i(x). \tag{2.13}$$

Now by the assumption $f^{(k)} > 0$ there exists an $\varepsilon > 0$ such that $f'(q+x)$ is strictly increasing for $x \in (0, \varepsilon)$. By the continuity of g at 0, there exists $\gamma > 0$ such that $g(x) < q+\varepsilon$ for $x \in (0, \gamma)$. By Lemma 2.9, there exists n_γ such that for $0 < x \leq \gamma$, $(g_i(x))$ is a monotone sequence (an increasing one, since we assume $f^{(k)}(q) > 0$) for $i > n_\gamma$. Therefore, for $i \geq n_\gamma$, and for $0 < x < \gamma$,

$$g_i(x) \leq g(x) < q + \varepsilon. \tag{2.14}$$

On the other hand, for $i < n_\gamma \wedge n$, since $f(q+x) > q+x$ for $x \in (0, 1-q)$,

$$f^i\left(q + \frac{x}{\xi^n}\right) \leq f^{n_\gamma}\left(q + \frac{x}{\xi^n}\right) \leq f^{n_\gamma}\left(q+x\right). \tag{2.15}$$

f^{n_γ} is continuous and non-decreasing, hence we may pick $\tilde{\gamma} > 0$ such that for $0 < x < \tilde{\gamma}$,

$$f^{n_\gamma}(q+x) < q + \varepsilon. \tag{2.16}$$

Finally, combining (2.12) – (2.16), we obtain that for all $0 < i < n$ and for all $0 \leq y \leq x \leq \gamma \wedge \tilde{\gamma}$,

$$q \leq f^i\left(q + \frac{y}{\xi^n}\right) \leq f^i\left(q + \frac{x}{\xi^n}\right) \leq q + \varepsilon. \tag{2.17}$$

Recall that ε was chosen to be such that f' is strictly increasing on $(q, q+\varepsilon)$. Combining this with (2.11) and (2.17) we obtain that

$$g_n'(y) \leq g_n'(x). \tag{2.18}$$

We are now going to use (2.18) to show that (2.10) holds for $I = (0, \varepsilon)$ with $\varepsilon = \frac{1}{2}(\gamma \wedge \tilde{\gamma})$. For each $z \in (\varepsilon, 2\varepsilon)$ and all $n \geq 0$, by (2.18) we have

$$\sup_{y \in (0,\varepsilon)} g_n'(y) \leq g_n'(z).$$

Therefore,

$$\begin{aligned}
\sup_{n \geq 0} \sup_{y \in (0,\varepsilon)} g_n'(y) &\leq \sup_{n \geq 0} \frac{1}{\varepsilon} \int_\varepsilon^{2\varepsilon} g_n'(z) \mathrm{d}z \\
&\leq \sup_{n \geq 0} \frac{1}{\varepsilon} (g_n(2\varepsilon) - g_n(\varepsilon)) \\
&\leq \frac{1}{\varepsilon},
\end{aligned}$$

where the last inequality follows since each g_n is a distribution function. This proves that

$g(y)$ is continuous on $(0, \varepsilon)$. This completes the proof of Lemma 2.7.

□

Multiple atoms

In the previous section we found the correct order of fluctuations when f had a single fixed point in the interval $(0, 1)$. When f has more than one fixed point in the interval $(0, 1)$, we cannot simply consider the quantity $W_{2n} - q$, since the limiting distribution has multiple atoms, but it turns out that we can condition on W_{2n} being close enough to one of the atoms and straightforwardly apply Lemma 2.7. Note that the set of fixed points of f cannot have an accumulation point in the interval $(0, 1)$. To see this, recall that f is a composition of functions analytic in $(0, 1)$. Therefore, $f(x) - x$ is also analytic in $(0, 1)$ and we justify the claim using the fact that zeros of an analytic function not identical to 0 cannot have any accumulation points in the domain in which the function is analytic.

The case of multiple atoms of V is summarized in the following lemma:

Lemma 2.10. *Consider the recursion (2.3) and assume that q_-, q, q_+ are fixed points of f satisfying the following conditions:*

- $q \in (0, 1)$ *is unstable and* $f'(q) > 1$,
- $q_- < q < q_+$,
- q *is the only unstable from at least one side point of* f *in the interval* (q_-, q_+).

Then,

$$\mathcal{L}\left(\xi^n (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}(V),$$

where $\xi = f'(q)$ *and* V *is a random variable with a continuous distribution function.*

Note first that since $f'(q) > 1$, the definitions of q_- and q_+ coincide with those given in (2.4).

Proof of Lemma 2.10. Fix $x \in [0,1]$. Then

$$\begin{aligned}
\mathbb{P}\left(\frac{W_{2n}-q_-}{q_+-q_-} \le x \middle| W_{2n} \in [q_-, q_+]\right) &= \frac{\mathbb{P}(q_- \le W_{2n} \le x(q_+-q_-)+q_-)}{\mathbb{P}(q_- \le W_{2n} \le q_+)} \\
&= \frac{f(\mathbb{P}(W_{2n-2} \le x(q_+-q_-)+q_-)-q_-}{q_+-q_-} \\
&= \frac{f\left(\mathbb{P}\left(\frac{W_{2n-2}-q_-}{q_+-q_-} \le x\right)\right)-q_-}{q_+-q_-} \\
&= \tilde{f}\left(\mathbb{P}\left(\frac{W_{2n-2}-q_-}{q_+-q_-} \le x \middle| W_{2n-2} \in [q_-, q_+]\right)\right),
\end{aligned} \tag{2.19}$$

where

$$\tilde{f}(x) = \frac{f(x(q_+-q_-)+q_-)-q_-}{q_+-q_-}.$$

Furthermore, $\tilde{f}(x)$ is a continuous bijective mapping from $[0,1]$ to $[0,1]$ with a single fixed point $\tilde{q} = \frac{q-q_-}{q_+-q_-}$ in $(0,1)$ and

$$\begin{aligned}
\tilde{f}'(\tilde{q}) &= f'(q) = \xi, \\
\tilde{f}^{(k)}(\tilde{q}) &= f^{(k)}(q)(q_+-q_-)^{k-1}.
\end{aligned} \tag{2.20}$$

Consider a sequence of random variables $(\tilde{W}_{2n})_{n=0}^{\infty}$ such that

$$\tilde{W}_{2n} \overset{d}{=} \left(\frac{W_{2n}-q_-}{q_+-q_-} \;\middle|\; W_{2n} \in [q_-, q_+]\right).$$

We check that $\tilde{W}_0 \sim U(0,1)$, $\mathbb{P}(\tilde{W}_{2n} \le x) = \tilde{f}(\mathbb{P}(\tilde{W}_{2n-2} \le x)$. Combining this with (2.20), we may apply Lemma 2.7 to $(\tilde{W}_{2n})_{n=0}^{\infty}$ to conclude that

$$\mathcal{L}\left(\xi^n(W_{2n}-q) \mid W_{2n} \in [q_-, q_+]\right) \to \mathcal{L}((q_+-q_-)V),$$

where W is a random variable with a continuous distribution function $\tilde{g}$ that satisfies $\tilde{g}(x) = \tilde{f}(\tilde{g}(x))$. Finally, we note that the distribution function of $(q_+-q_-)V$ is also continuous, which completes the proof. □

Boundary fixed points

To finish the proof of part (a) of Theorem 2.3 we need to consider the case when one of q_-, q_+ is equal to q. This may happen either if q is at the boundary (i.e. $q \in \{0,1\}$) or when $q \in (0,1)$, but q is stable from one side. These cases can be treated simultaneously by repeating the reasoning from the proofs of Lemma 2.7 and Lemma 2.10. Note that the limiting distribution V is now concentrated on either the positive or negative half-line.

2.4.2 Proof of Theorem 2.3(b): $f'(q) = 1$

We have already described the fluctuations of W_{2n} when we know that it converges to some fixed point q of f with $f'(q) \in (1, \infty)$. If the point was unstable from both sides, we obtained a two-sided continuous limiting distribution.

If q is a fixed point of f such that $f'(q) = 1$, it may be unstable, stable or unstable from one side and stable from the other. In this case it is more convenient to consider each side of q separately. For simplicity, we state and prove a lemma for the case where q is unstable from the right and then comment on the general case.

Note that the set of fixed points doesn't have an accumulation point in $(0,1)$, but it is not known whether this behaviour may be exhibited at the boundary, hence the additional assumption in the lemma.

Lemma 2.11. *Consider the recursion (2.3) and assume that q is a fixed point of f, that it is unstable from the right and let $q_+ = \inf\{x : x > q, x = f(x)$. Suppose that $f'(q) = 1$ and k is such that $f^{(r)}(q) = 0$ for $1 < r < k$ and $f^{(k)}(q) \neq 0$. If $q_+ \neq q$, then*

$$\mathcal{L}\left(n^{\frac{1}{k-1}}(W_{2n} - q) \mid W_{2n} \in [q, q_+]\right) \to \delta_a,$$

where $a = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}}$.

Figure 2.3(a) gives an example where Lemma 2.11 applies.

Proof of Lemma 2.11. We are going to show that the distribution function of $n^{\frac{1}{k-1}} \frac{W_{2n}-q}{q_+-q}$ conditioned on the event $W_{2n} \in [q, q_+]$ converges to some limit as n tends to infinity. Define

$$g_n(x) := \mathbb{P}\left(n^{\frac{1}{k-1}} \frac{W_{2n} - q}{q_+ - q} \leq x \;\middle|\; W_{2n} \in [q, q_+]\right). \tag{2.21}$$

By calculations similar to those (2.19) in the proof of Lemma 2.10, we obtain that for $x \in [0,1]$,

$$\mathbb{P}\left(\frac{W_{2n}-q}{q_+-q} \le x \;\middle|\; W_{2n} \in [q,q_+]\right) = \tilde{f}\left(\mathbb{P}\left(\frac{W_{2n}-q}{q_+-q} \le x \;\middle|\; W_{2n-2} \in [q,q_+]\right)\right),$$

where

$$\tilde{f}(x) = \frac{f(x(q_+-q)+q)-q}{q_+-q}.$$

Note that

$$\begin{aligned}\tilde{f}(0) &= 0,\\ \tilde{f}(1) &= 1,\\ \tilde{f}'(0) &= f'(q) = 1,\\ \tilde{f}^{(i)}(0) &= (q_+-q)^{i-1} f^{(i)}(q),\end{aligned}$$

and $\tilde{f}$ has no fixed points in $(0,1)$. Since for every $x > 0$, for sufficiently large n, $\frac{x}{n^{\frac{1}{k-1}}} \in [0,1]$, for such n we have

$$\begin{aligned} g_n(x) &= \mathbb{P}\left(\frac{W_{2n}-q}{q_+-q} \le xn^{-\frac{1}{k-1}} \;\middle|\; W_{2n} \in [q,q_+]\right)\\ &= \tilde{f}^n\left(\mathbb{P}\left(\frac{W_0-q}{q_+-q} \le xn^{-\frac{1}{k-1}} \;\middle|\; W_0 \in [q,q_+]\right)\right) \qquad (2.22)\\ &= \tilde{f}^n\left(xn^{-\frac{1}{k-1}}\right). \end{aligned}$$

The proof consists of two parts:

I We show that for each $x < a$, for sufficiently large n, $(g_n(x))$ forms a decreasing sequence, and for each $x > a$, for sufficiently large n, $(g_n(x))$ forms an increasing one,

II we show that for $x < a$, $g_n(x) \to 0$, and for $x > a$, $g_n(x) \to 1$.

Part I Fix $x \neq 0$. Using Taylor's expansion, we may expand $\tilde{f}(x)$ as follows:

$$\tilde{f}(x) = \tilde{f}(0) + x + \frac{\tilde{f}^{(k)}(0)}{k!}x^k + r_k(x)x^k,$$

where $\lim_{x\to 0} r_k(x) = 0$. Therefore,

$$\tilde{f}\left(\frac{x}{n^{\frac{1}{k-1}}}\right) < \frac{x}{(n-1)^{\frac{1}{k-1}}} \tag{2.23}$$

is equivalent to

$$\frac{x}{n^{\frac{1}{k-1}}} + \frac{\tilde{f}^{(k)}(0)}{k!}\left(\frac{x}{n^{\frac{1}{k-1}}}\right)^k + r_k\left(\frac{x}{n^{\frac{1}{k-1}}}\right)\left(\frac{x}{n^{\frac{1}{k-1}}}\right)^k < \frac{x}{(n-1)^{\frac{1}{k-1}}},$$

and to

$$\frac{\tilde{f}^{(k)}(0)}{k!}x^{k-1} + r_k\left(\frac{x}{n^{\frac{1}{k-1}}}\right)x^{k-1} < n\left(\left(\frac{n}{n-1}\right)^{\frac{1}{k-1}} - 1\right).$$

Letting $n \to \infty$, the right-hand side of the last formula converges to $\frac{1}{k-1}$, whereas the left-hand one converges to $\frac{\tilde{f}^{(k)}(0)}{k!}x^{k-1}$. Therefore, the last inequality is satisfied for large n if

$$x < \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}} \frac{1}{q_+ - q},$$

and similarly

$$\tilde{f}\left(\frac{x}{n^{\frac{1}{k-1}}}\right) > \frac{x}{(n-1)^{\frac{1}{k-1}}} \tag{2.24}$$

for large n if

$$x > \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}} \frac{1}{q_+ - q}.$$

This yields the claim, as $\tilde{f}^{n-1}$ is a strictly increasing function, hence recalling (2.22), the inequality (2.23) is equivalent to

$$g_n(x) = \tilde{f}^n\left(\frac{x}{n^{\frac{1}{k-1}}}\right) < \tilde{f}^{n-1}\left(\frac{x}{(n-1)^{\frac{1}{k-1}}}\right) = g_{n-1}(x),$$

and the inequality (2.24) is equivalent to

$$g_n(x) = \tilde{f}^n\left(\frac{x}{n^{\frac{1}{k-1}}}\right) > \tilde{f}^{n-1}\left(\frac{x}{(n-1)^{\frac{1}{k-1}}}\right) = g_{n-1}(x).$$

This completes the proof of the claim.

Part II Since each g_n is a distribution function, and by Part I above for each x, $(g_n(x))$ is a monotone sequence for large n (decreasing for $x < a$ and increasing for $x > a$), hence the limit $g(x) = \lim_{n\to\infty} g_n(x)$ exists for all x. To show that for $x < a$, $g(x) = 0$, assume that for some $0 < x < a$, $g(x) = \varepsilon > 0$. This implies that

$$\tilde{f}^n\left(\frac{x}{n^{\frac{1}{k-1}}}\right) \geq \varepsilon > 0 \tag{2.25}$$

for large n. Take $y \in \mathbb{R}, l \in \mathbb{N}$ such that $y = x\left(\frac{l}{l-1}\right)^{\frac{1}{k-1}} < a$. Note also that $g(y) \leq 1$, but since $(g_n(y))$ is a strictly decreasing sequence, the inequality is in fact sharp, thus

$$\lim_{n\to\infty} g_n(y) < 1. \tag{2.26}$$

On the other hand,

$$\begin{aligned} \lim_{n\to\infty} g_n(y) = \lim_{n\to\infty} g_{nl}(y) &= \lim_{n\to\infty} \tilde{f}^{nl}\left(\frac{y}{(nl)^{\frac{1}{k-1}}}\right) \\ &= \lim_{n\to\infty} \tilde{f}^n \circ \tilde{f}^{n(l-1)}\left(\frac{y\left(\frac{l-1}{l}\right)^{\frac{1}{k-1}}}{(n(l-1))^{\frac{1}{k-1}}}\right) \\ &= \lim_{n\to\infty} \tilde{f}^n \circ \tilde{f}^{n(l-1)}\left(\frac{x}{(n(l-1))^{\frac{1}{k-1}}}\right), \end{aligned} \tag{2.27}$$

and by (2.25),

$$\lim_{n\to\infty} \tilde{f}^n \circ \tilde{f}^{n(l-1)}\left(\frac{x}{(n(l-1))^{\frac{1}{k-1}}}\right) \geq \lim_{n\to\infty} \tilde{f}^n(\varepsilon) = 1,$$

as 1 is the only stable fixed point of $\tilde{f}$ in the interval $[0, 1]$. But this contradicts (2.26) and thus $g(x) = 0$.

Similarly, to show that for $x > a$, $g(x) = 1$, fix any such x and take $y \in \mathbb{R}, l \in \mathbb{N}$ such that $y = \left(\frac{l-1}{l}\right)^{\frac{1}{k-1}} x > a$. By calculations similar to (2.27),

$$\begin{aligned}\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} g_{nl}(x) &= \lim_{n\to\infty} \tilde{f}^{nl}\left(\frac{x}{(nl)^{\frac{1}{k-1}}}\right) \\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ \tilde{f}^{n(l-1)}\left(\frac{y}{(n(l-1))^{\frac{1}{k-1}}}\right) \\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ g_{n(l-1)}(y),\end{aligned}$$

but since $(g_n(y))$ is a strictly increasing sequence for large n, for these n, $g_n(y) \geq \delta$ for some $\delta > 0$, hence

$$\lim_{n\to\infty} \tilde{f}^{n} \circ g_{n(l-1)}(y) \geq \lim_{n\to\infty} \tilde{f}^{n}(\delta) = 1,$$

as again, 1 is the only stable fixed point of $\tilde{f}$ in $[0, 1]$.

Recall now the definition of $g_n(x)$ (2.21). Parts I and II prove that

$$\mathcal{L}\left(n^{\frac{1}{k-1}} \frac{W_{2n} - q}{q_+ - q} \;\middle|\; W_{2n} \in [q, q_+]\right) \to \delta_a,$$

and thus

$$\mathcal{L}\left(n^{\frac{1}{k-1}} (W_{2n} - q) \;\middle|\; W_{2n} \in [q, q_+]\right) \to \delta_{a(q_+ - q)},$$

where

$$a(q_+ - q) = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{\frac{1}{k-1}}.$$

This completes the proof of Lemma 2.11. □

To finish the proof of part (b) of Theorem 2.3 we apply Lemma 2.11 and its counterpart for points unstable from the left to $[q, q_+]$ and $[q_-, q]$ respectively. Checking that for each n

$$\begin{aligned}P(W_{2n} \in [q, q_+] \mid W_{2n} \in [q_-, q_+]) &= \frac{q_+ - q}{q_+ - q_-}, \\ P(W_{2n} \in [q_-, q] \mid W_{2n} \in [q_-, q_+]) &= \frac{q - q_-}{q_+ - q_-},\end{aligned}$$

shows that the masses in the formulation of the theorem are chosen appropriately, hence completes the proof.

2.4.3 Proof of Theorem 2.3(c): $f'(q) = \infty$

Note first that $f'(q) = \infty$ can happen only at $q \in \{0, 1\}$.

We start by describing behaviour of f near 0. The first step is supplied by the following technical lemma:

Lemma 2.12. *There exist functions $H(x)$ and $b(x)$ defined on $(-1, 1)$ such that $f(x) \sim H(x)$ as $x \to 0$ and*

$$H'(x) \sim \left(\sum_{n=1}^{\infty} np_n b(x)^{n-1} \right) Kp_K x^{K-1},$$

where $1 - b(x) \sim p_K x^K$ as $x \to 0$.

Proof of Lemma 2.12. By simple calculations,

$$\begin{aligned} f(x) = R(R(x)) &= \frac{R(R(x))}{1 - R(x)}(1 - R(x)) = \frac{1 - \sum_{n=1}^{\infty} p_n R(x)^n}{1 - R(x)} \sum_{n=1}^{\infty} p_n x^n \\ &= \frac{\sum_{n=1}^{\infty} p_n (1 - R(x)^n)}{1 - R(x)} \sum_{n=1}^{\infty} p_n x^n = \sum_{n=1}^{\infty} \left(p_n \sum_{i=0}^{n-1} R(x)^i \right) \sum_{n=1}^{\infty} p_n x^n \\ &= \sum_{i=0}^{\infty} \left(R(x)^i \sum_{n=i+1}^{\infty} p_n \right) \sum_{n=1}^{\infty} p_n x^n = \sum_{i=0}^{\infty} \left[R(x)^i \mathbb{P}(M > i) \right] \sum_{n=1}^{\infty} p_n x^n, \end{aligned} \tag{2.28}$$

where M is a random variable with law $\mathbb{P}(M = i) = p_i$. Furthermore, recalling that $K = \min\{i : p_i \neq 0\}$,

$$\begin{aligned} \frac{R(x)}{1 - x} &= \frac{\sum_{n=1}^{\infty} p_n (1 - x^n)}{1 - x} = \sum_{n=1}^{\infty} p_n \sum_{i=0}^{n-1} x^i = \sum_{i=0}^{\infty} x^i \sum_{n=i+1}^{\infty} p_n \\ &= \sum_{i=0}^{\infty} x^i \mathbb{P}(M > i) = 1 + x + \ldots + x^{K-1} + x^K \underbrace{\sum_{i=K}^{\infty} x^{i-K} \mathbb{P}(M > i)}_{\text{denote by } h(x)}. \end{aligned} \tag{2.29}$$

Therefore, substituting (2.29) into (2.28),

$$f(x) = \sum_{i=0}^{\infty} \left[(1 - x)(1 + x + \ldots + x^{K-1} + x^K h(x)) \right]^i \mathbb{P}(M > i) \sum_{n=1}^{\infty} p_n x^n, \tag{2.30}$$

where

$$h(x) = \sum_{i=K}^{\infty} x^{i-K} \mathbb{P}(M > i) \to \mathbb{P}(M > K) = 1 - p_K \quad \text{as } x \to 0. \tag{2.31}$$

Observe that $h'(x) \to \mathbb{P}(M > K+1)$ as $x \to 0$. Now for any $b < 1$,

$$\sum_{i=0}^{\infty} b^i \mathbb{P}(M > i) = \sum_{n=1}^{\infty} p_n \sum_{i=0}^{n-1} b^i = \frac{1}{1-b} \left(1 - \sum_{n=1}^{\infty} p_n b^n \right),$$

and thus, setting

$$b(x) = (1-x)(1 + x + \ldots + x^{K-1} + x^K h(x)) = 1 - x^K + x^K h(x) - x^{K+1} h(x),$$

from (2.30) we obtain

$$f(x) = \frac{1}{1-b(x)} \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n \right) \sum_{n=1}^{\infty} p_n x^n. \tag{2.32}$$

Observe that

$$1 - b(x) = x^K (1 - h(x)) + x^{K+1} h(x),$$

hence by (2.31),

$$1 - b(x) \sim p_K x^K \quad \text{as } x \to 0.$$

Moreover, from (2.32),

$$\begin{aligned} f(x) &= \frac{1}{x^K(1-h(x)) + x^{K+1} h(x)} \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n \right) \sum_{n=1}^{\infty} p_n x^n = \\ &= \frac{1}{1 - h(x) + x h(x)} \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n \right) \sum_{n=K}^{\infty} p_n x^{n-K}. \end{aligned}$$

Note that as $x \to 0$, the first fraction on the right-hand side converges to $\frac{1}{p_K}$, and the final sum converges to p_K. Thus

$$f(x) \sim \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n \right)$$

as $x \to 0$. Therefore, denoting $H(x) := (1 - \sum_{n=1}^{\infty} p_n b(x)^n)$, we have

$$f(x) \sim H(x) \tag{2.33}$$

and, recalling that $\lim_{x\to 0} h'(x) = \mathbb{P}(M > K+1)$,

$$H'(x) = \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right)' \sim \left(\sum_{n=1}^{\infty} n p_n b(x)^{n-1}\right) K p_K x^{K-1}$$

as $x \to 0^+$ which completes the proof of the lemma. □

Equipped with the relation from Lemma 2.12 we may now connect f with the underlying offspring distribution via *Karamata's Tauberian Theorem for Power Series* (a proof may be found e.g. in [14]). Recall first the theorem:

Theorem 2.13 (Karamata's Tauberian Theorem)**.** *If $a_n \geq 0$ and the power series $A(s) = \sum_{n=0}^{\infty} a_n s^n$ converges for $s \in [0,1)$, then for $c, \rho \geq 0$ the following are equivalent:*

$$\sum_{k=0}^{n} a_k \sim c n^{\rho} \text{ as } n \to \infty$$

and

$$A(s) \sim \frac{c\Gamma(1+\rho)}{(1-s)^{\rho}} \text{ as } s \uparrow 1.$$

Recall the assumption (2.5) of Theorem 2.3: for some $\rho \in (0,1)$,

$$\mathbb{E}(M \mathbb{I}_{M \leq n}) = \sum_{k=1}^{n} k p_k \sim c n^{\rho}.$$

By Theorem 2.13 applied to $a_k = k p_k$ and (2.5) we obtain that as we let $x \to 0$ (which implies $b(x) \to 1$),

$$\frac{1}{K p_K x^{K-1}} H'(x) \sim \sum_{n=1}^{\infty} n p_n b(x)^{n-1} \sim \frac{c\Gamma(1+\rho)}{(1-b(x))^{\rho}} \sim \frac{c\Gamma(1+\rho)}{p_K^{\rho}} \frac{1}{x^{K\rho}}.$$

Therefore,

$$H(t) = \int_0^t H'(x)\mathrm{d}x \sim \int_0^t \frac{c\Gamma(1+\rho)}{p_K^\rho} \frac{1}{x^{K\rho}} K p_K x^{K-1} \mathrm{d}x = \frac{c\Gamma(1+\rho) p_K^{1-\rho}}{1-\rho} t^{K-K\rho}, \tag{2.34}$$

as $t \to 0$, hence, by (2.33) and (2.34),

$$f(t) \sim \frac{c\Gamma(1+\rho) p_K^{1-\rho}}{1-\rho} t^{K(1-\rho)}$$

as $t \to 0$. This implies that for $f'(0) = \infty$ to hold it is necessary that $K < \frac{1}{1-\rho}$.

To provide the criterion for $q = 1$, we are interested in the behaviour of the quantity $1 - f(t)$ when $t \to 1$. Now

$$1 - f(t) = 1 - R(R(t)) = G(R(t)) \sim p_K R(t)^K. \tag{2.35}$$

By definition, $R(x) = 1 - \sum_{k=1}^{\infty} p_k x^k$, thus

$$R'(x) = -\sum_{k=1}^{\infty} k p_k x^{k-1},$$

and again by Theorem 2.13 applied to $a_k = k p_k$ and (2.5),

$$R'(x) \sim -\frac{c\Gamma(1+\rho)}{(1-x)^\rho}$$

as $x \to 1$, and thus

$$R(t) = R(1) - \int_t^1 R'(x)\mathrm{d}x \sim \frac{c\Gamma(1+\rho)}{1-\rho} (1-t)^{1-\rho}. \tag{2.36}$$

Substituting (2.36) into (2.35) we obtain that

$$1 - f(t) \sim p_K \left(\frac{c\Gamma(1+\rho)}{1-\rho} \right)^K (1-t)^{K(1-\rho)}$$

as $t \to 1$. Thus again, for $f'(1) = \infty$ to hold it is necessary that $K(1-\rho) < 1$.

We've shown that $f(t) \sim C_0 t^{K(1-\rho)}$ as $t \to 0$ and $1 - f(t) \sim C_1 (1-t)^{K(1-\rho)}$ as $t \to 1$ for some positive constants C_0, C_1 that we determined explicitly. Note that

$$C_1 = C_0^k p_K^{1-K(1-\rho)}$$

and since $K(1-\rho) < 1$, at least one of the constants C_0, C_1 is different from 1. The proof of Proposition 2.3(c) is completed by the following two lemmas applied as follows: Lemma 2.14 applied with $\alpha = K(1-\rho)$ proves existence of the distributional limit at either point q with $C_q \neq 1$, and Lemma 2.15 shows that the limit exists at $q = 0$ if and only if it exists at $q = 1$.

Lemma 2.14. *Consider the recursion (2.3);*

1. *Assume that* $f(t) \sim Ct^\alpha$ *with* $C \neq 1$ *and* $\alpha \in (0,1)$ *as* $t \to 0$*. Let* $q_+ = \inf\{x : x > 0,\ x = f(x)\}$*. Then*

$$\mathcal{L}\left(\alpha^n \log W_{2n} \mid W_{2n} \in [0, q_+]\right) \xrightarrow{d} V_0, \tag{2.37}$$

   where V_0 *is a random variable with* $\mathbb{P}(V_0 \in (-\infty, 0)) = 1$.

2. *Assume that* $1 - f(t) \sim C(1-t)^\alpha$ *with* $C \neq 1$ *and* $\alpha \in (0,1)$ *as* $t \to 1$*. Let* $q_- = \sup\{x : x < 1,\ x = f(x)\}$*. Then*

$$\mathcal{L}\left(\alpha^n \log(1 - W_{2n}) \mid W_{2n} \in [q_-, 1]\right) \xrightarrow{d} V_1, \tag{2.38}$$

   where V_1 *is a random variable with* $\mathbb{P}(V_1 \in (-\infty, 0)) = 1$.

Lemma 2.15. *Consider the recursion (2.2). For* $\alpha \in (0,1)$ *convergence (2.37) holds for some* V_0 *with* $\mathbb{P}(V_0 \in (-\infty, 0)) = 1$ *if and only if convergence (2.38) holds for some* V_1 *with* $\mathbb{P}(V_1 \in (-\infty, 0)) = 1$.

Note that in Lemma 2.15 we do not assume anything about f; in particular we do not assume that $C \neq 1$.

Proof of Lemma 2.14. Firstly we show how the second part can be obtained from the first one and then we prove the first part of Lemma 2.14, which corresponds to $q = 0$.

Assume that $1 - f(t) \sim C(1-t)^\alpha$ and set $\tilde{W}_{2n} = 1 - W_{2n}$ and $\tilde{f}(t) = 1 - f(1-t)$. Then,

$$\begin{aligned}\mathbb{P}(\tilde{W}_{2n} \le x) &= \mathbb{P}(1 - W_{2n} \le x) = 1 - \mathbb{P}(W_{2n} \le 1 - x) = 1 - f(\mathbb{P}(W_{2n-2} \le 1 - x)) \\ &= 1 - f(1 - \mathbb{P}(\tilde{W}_{2n-2} \le x)) = \tilde{f}(\mathbb{P}(\tilde{W}_{2n-2} \le x)) = \ldots = \tilde{f}^n(\mathbb{P}(\tilde{W}_0 \le x)) \\ &= \tilde{f}^n(x)\end{aligned}$$

and $\tilde{f}(t) = 1 - f(1-t) \sim Ct^\alpha$ as $t \to 0$. Hence it is enough to prove the result for the case $q = 0$.

Fix some $x < 0$. Note that for n large enough $\exp\left(\frac{x}{\alpha^n}\right) \le q_+$, hence for these n,

$$\begin{aligned}\mathbb{P}(\alpha^n \log W_{2n} \le x \mid W_{2n} \in [0, q_+]) = \mathbb{P}\left(W_{2n} \le \exp\left(\frac{x}{\alpha^n}\right) \;\middle|\; W_{2n} \in [0, q_+]\right) = \\ = \frac{1}{q_+} f^n\left(\exp\left(\frac{x}{\alpha^n}\right)\right).\end{aligned} \tag{2.39}$$

Define

$$f_l(y) = \log(f(\exp(y))),$$

and observe that

$$f_l^n(y) = \log(f^n(\exp(y))). \tag{2.40}$$

The idea behind $f_l(y)$ is to linearise $f(y)$: note that f_l is a monotone function, $f_l(\log q_+) = \log q_+$ and that

$$f_l(y) = \alpha y + O(1)$$

as $y \to -\infty$, hence there exist constants $\tilde{D}, \tilde{E}$ such that for $y \le \log q_+ < 0$,

$$\tilde{D} + \alpha y \le f_l(y) \le \tilde{E} + \alpha y. \tag{2.41}$$

In the first part of the proof we show (assuming that the limit (2.37) exists) that $\mathbb{P}(V \in$

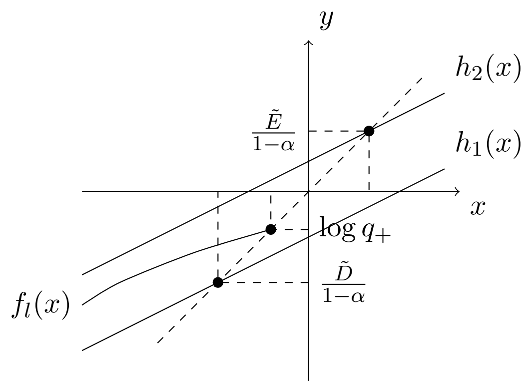

Figure 2.4: $h_1(x), f_l(x), h_2(x)$ together with their (stable) fixed points $\frac{\tilde{D}}{1-\alpha}, \log q_+, \frac{\tilde{E}}{1-\alpha}$ respectively. For $x \leq \log q_+$, $h_1(x) \leq f_l(x) \leq h_2(x)$. The dashed line represents the identity function.

$(-\infty, 0)) = 1$. Define

$$h_1(y) = \tilde{D} + \alpha y,$$
$$h_2(y) = \tilde{E} + \alpha y.$$

Now

$$\begin{aligned} \lim_{n\to\infty} h_1^n\left(\frac{y}{\alpha^n}\right) &= y + \frac{\tilde{D}}{1-\alpha}, \\ \lim_{n\to\infty} h_2^n\left(\frac{y}{\alpha^n}\right) &= y + \frac{\tilde{E}}{1-\alpha}, \end{aligned} \tag{2.42}$$

where $\frac{\tilde{D}}{1-\alpha}$, $\frac{\tilde{E}}{1-\alpha}$ are the (unique) fixed points of h_1 and h_2 respectively.

Equations (2.39), (2.40), (2.41) and (2.42) together imply that

$$\limsup_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leq x | W_{2n} \in [0, q_+]) \leq \frac{1}{q_+} \exp\left(x + \frac{\tilde{E}}{1-\alpha}\right),$$

and therefore

$$\lim_{x\to-\infty} \limsup_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leq x \mid W_{2n} \in [0, q_+]) = 0.$$

We shall now show that

$$\lim_{x\to 0^-} \liminf_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leq x \mid W_{2n} \in [0, q_+]) = 1. \tag{2.43}$$

To do so, note first that by (2.39) and (2.40)

$$\mathbb{P}(\alpha^n \log W_{2n} \leq x \mid W_{2n} \in [0, q_+]) = \frac{1}{q_+} \exp\left(f_l^n\left(\frac{x}{\alpha^n}\right)\right),$$

hence (2.43) is equivalent to

$$\lim_{x \to 0^-} \liminf_{n \to \infty} f_l^n\left(\frac{x}{\alpha^n}\right) = \log q_+. \tag{2.44}$$

Note that $\log q_+$ is a fixed point of $f_l(x)$. Equation (2.44) indicates that the scaling α^n is not strong enough to compensate the attraction of the fixed point $\log q_+$ of f_l.

Let $k_{x,n}$ be the smallest k such that $h_1^k\left(\frac{x}{\alpha^n}\right) \geq \frac{\tilde{D}}{1-\alpha} - 1$. Note that $k_{x,n}$ is properly defined as $\frac{\bar{D}}{1-\alpha}$ is the only fixed point of h_1 and is stable. Moreover, by (2.42) we have that for $x \in (-1, 0)$,

$$\lim_{n \to \infty} h_1^n\left(\frac{x}{\alpha^n}\right) = x + \frac{\bar{D}}{1-\alpha} \geq \frac{\tilde{D}}{1-\alpha} - 1,$$

hence for these x, $n - k_{x,n} \geq 0$ for large n. Define also

$$K_x = \liminf_{n \to \infty}(n - k_{x,n})$$

and note that since

$$\lim_{x \to 0^-} \lim_{n \to \infty} h_1^n\left(\frac{x}{\alpha^n}\right) = \frac{\tilde{D}}{1-\alpha},$$

and the right-hand side is a fixed point of h_1, we obtain that

$$\lim_{x \to 0^-} K_x = \infty.$$

Now define similarly $\tilde{k}_{x,n}$ to be the smallest k such that $f_l^k\left(\frac{x}{\alpha^n}\right) > \frac{\tilde{D}}{1-\alpha} - 1$. Since $f_l(y) \geq h_1(y)$ for $y \leq \log q_+$, we have $k_{x,n} \geq \tilde{k}_{x,n}$, and therefore

$$\lim_{x \to 0^-} \liminf_{n \to \infty}(n - \tilde{k}_{x,n}) = \infty.$$

This implies that

$$
\begin{aligned}
\lim_{x\to 0^-}\liminf_{n\to\infty} f_l^n\left(\frac{x}{\alpha^n}\right) &= \lim_{x\to 0^-}\liminf_{n\to\infty} f_l^{n-k_{x,n}}\left(f_l^{k_{x,n}}\left(\frac{x}{\alpha^n}\right)\right)\\
&\geq \lim_{x\to 0^-}\liminf_{n\to\infty} f_l^{n-k_{x,n}}\left(\frac{\tilde{D}}{1-\alpha}-1\right)\\
&= \log q_+.
\end{aligned}
$$

To finish the proof it is now enough to justify that the limit $\lim_{n\to\infty}\mathbb{P}(\alpha^n \log W_{2n} \leq x|W_{2n} \in [0,q_+])$ exists for all x; recalling (2.39) it is enough to check that the sequence $f^n\left(\exp\left(\frac{x}{\alpha^n}\right)\right)$ is monotone for large n. Since f is a strictly monotone function, the statements

$$
f^{n+1}\left(\exp\left(\frac{x}{\alpha^{n+1}}\right)\right) \geq f^n\left(\exp\left(\frac{x}{\alpha^n}\right)\right)
$$

and

$$
f\left(\exp\left(\frac{x}{\alpha^{n+1}}\right)\right) \geq \exp\left(\frac{x}{\alpha^n}\right). \tag{2.45}
$$

are equivalent. We set $y = \frac{x}{\alpha^{n+1}}$ and $z = \exp(y)$ (therefore $y \to -\infty$ corresponds to $z \to 0$) obtaining that (2.45) is equivalent to:

$$
f(z) \geq z^\alpha.
$$

Therefore, if $f(z) \sim Cz^\alpha$ for $C \neq 1$ we observe that for each $x < \log q_+$ the sequence $f^n\left(\exp\left(\frac{x}{(\alpha)^n}\right)\right)$ is monotone for n large enough which yields existence of the limit. This completes the proof of Lemma 2.14. □

Up to now we only defined W_m for even m. Before we prove Lemma 2.15 we extend it to odd m. Define the distribution of a random variable W_{2n-1} as follows:

$$
W_{2n-1} \overset{d}{=} \max_{1\leq i\leq M} W_{2n-2}^{(i)},
$$

where M is a random variable drawn from the tree's offspring distribution and $W_{2n-2}^{(i)}$ are independent copies of W_{2n-2} (independent of M). The quantity W_{2n-1} corresponds to the value at the root of a tree of height $2n-1$, with levels alternating between max and min, starting and ending with a max. One has similarly

$$W_{2n} \stackrel{d}{=} \min_{1\le i\le M} W_{2n-1}^{(i)}.$$

Lemma 2.16 below provides a useful identity which we are going to apply in the proof of Lemma 2.15.

Lemma 2.16. $W_{2n-1} \stackrel{d}{=} G^{-1}(1 - W_{2n-2})$.

Proof of Lemma 2.16. G is the probability generating function of the offspring distribution of the tree, so $G(t) = \mathbb{P}(\max_{1\le i\le M} U_i \le t)$ where U_i are independent uniform random variables and M follows the offspring distribution (independently of $(U_i, i \ge 1)$). Decomposing the minimax tree of height $2n-1$ with maximum at levels 1 and $2n-1$, we see that random variables at level $2n-2$ (i.e. one level above the leaves) are distributed as $\max_{1\le i\le M} U_i$. Therefore

$$W_{2n-1} \stackrel{d}{=} W_{2n-2}^{\max,G}, \tag{2.46}$$

where $W_{2n-2}^{\max,G}$ is a random variable corresponding to a max-min tree (i.e. with maximum at the even levels and minimum at the odd ones) where at the leaves instead of uniform random variables we put random variables with distribution function G. Noting that if U is a uniform random variable then $G^{-1}(U)$ has distribution function G, we see that

$$W_{2n-2}^{\max,G} \stackrel{d}{=} G^{-1}(W_{2n-2}^{\max}). \tag{2.47}$$

Now, since

$$\max_{1\le i\le M} U_i = 1 - \min_{1\le i\le M}(1 - U_i) \stackrel{d}{=} 1 - \min_{1\le i\le M} U_i$$

and

$$\min_{1\le i\le M} U_i = 1 - \max_{1\le i\le M}(1 - U_i) \stackrel{d}{=} 1 - \max_{1\le i\le M} U_i,$$

we obtain that

$$W_{2n-2}^{\max} \overset{d}{=} 1 - W_{2n-2}. \tag{2.48}$$

Finally, combining (2.46), (2.47) and (2.48) completes the proof. □

We are now ready to prove Lemma 2.15.

Proof of Lemma 2.15. The convergence (2.37) is equivalent to the convergence of

$$\lim_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \le x \mid W_{2n} \in [0, q_+]). \tag{2.49}$$

at all the continuity points of the corresponding limiting distribution function and similarly the convergence (2.38) is equivalent to the convergence of

$$\lim_{n\to\infty} \mathbb{P}(\alpha^n \log(1 - W_{2n}) \le x \mid W_{2n} \in [q_-, 1]). \tag{2.50}$$

at all the continuity points of the corresponding limiting distribution function. Fix $x < 0$. For large n,

$$\begin{aligned}
\mathbb{P}(\alpha^n \log W_{2n} \le x \mid W_{2n} \in [0, q_+]) &= \frac{1}{q_+}\mathbb{P}(\alpha^n \log W_{2n} \le x,\ W_{2n} \in [0, q_+]) \\
&= \frac{1}{q_+}\mathbb{P}(W_{2n} \le \exp(x/\alpha^n),\ W_{2n} \in [0, q_+]) \\
&= \frac{1}{q_+}\mathbb{P}(W_{2n} \le \exp(x/\alpha^n)).
\end{aligned}$$

Now by the branching structure of the tree,

$$\mathbb{P}(W_{2n} \le \exp(x/\alpha^n)) = 1 - G(\mathbb{P}(W_{2n-1} > \exp(x/\alpha^n))).$$

Since G is a continuous function, the convergence (2.49) is equivalent to the convergence

$$\lim_{n\to\infty} \mathbb{P}(W_{2n-1} > \exp(x/\alpha^n)).$$

By Lemma 2.16,

$$\begin{aligned}\mathbb{P}(W_{2n-1} > \exp(x/\alpha^n)) &= \mathbb{P}(G^{-1}(1 - W_{2n-2}) > \exp(x/\alpha^n)) \\ &= \mathbb{P}(1 - W_{2n-2} > G(\exp(x/\alpha^n))) \\ &= \mathbb{P}(\alpha^{n-2}\log(1 - W_{2n-2}) > \alpha^{n-2}\log(G(\exp(x/\alpha^n)))) \\ &= 1 - \mathbb{P}(\alpha^{n-2}\log(1 - W_{2n-2}) \le \alpha^{n-2}\log(G(\exp(x/\alpha^n)))).\end{aligned}$$

Since $G(t) \sim p_K t^K$ as $t \to 0$, we observe that

$$\alpha^{n-2}\log(G(\exp(x/\alpha^n))) = \alpha^{n-2}\log(p_K \exp((xK)/\alpha^n) + o(1) = \frac{xK}{\alpha^2} + o(1).$$

This implies that if the convergence (2.50) holds at some point $\frac{xK}{\alpha^2}$ that is a continuity point of the limiting distribution function, then the convergence (2.49) holds at x. Similarly, if the convergence (2.49) holds at some point x that is a continuity point of the limiting distribution function, then the convergence (2.50) holds at $\frac{xK}{\alpha^2}$. Since the set of discontinuity points of any distribution function is at most countable, this completes the proof. □

2.5 Proof of the endogeny result

To prove Theorem 2.5 we use the idea of *bivariate uniqueness* introduced by Aldous and Bandyopadhyay [4].

Informally, the idea is as follows: suppose we allow *two* values at each node. Each coordinate evolves separately, according to the minimax recursions (and using the same realisation of the tree structure). If we put bivariate values at the leaves of the tree, we then get a bivariate value at the root of the tree. Let us consider for the moment the case where the values are discrete (as for the Bernoulli case in Theorem 2.5). If the process is endogenous, and the tree is large, then with high probability the two components at the root agree with each other. On the other hand, if the process is not endogenous, then the probability that they disagree stays bounded away from zero as the size of the tree goes to infinity, and in fact we can obtain a bivariate process on the infinite tree that is two-periodic and non-degenerate (in the sense that the two components are not identically the same).

To formalise this we rewrite some of the ideas around (2.2) in new notation.

Let μ be a distribution on $[0, 1]$. We defined $T(\mu)$ to be the distribution of the LHS of

(2.2), given that the random variables $W_{2n-2}^{(i,j)}$ on the RHS of (2.2) are i.i.d. with distribution μ.

So T is a map from $\mathcal{P}$ to $\mathcal{P}$, where $\mathcal{P}$ is the space of distributions on $[0, 1]$. For Theorem 2.5 we assume that the Bernoulli$(1 - x)$ distribution is a fixed point of T.

Now consider the space $\mathcal{P}^{(2)}$ of distributions on $[0, 1]^2$. Define the map $T^{(2)}$ from $\mathcal{P}^{(2)}$ to itself as follows. As before let M and $M_1, M_2, \ldots$ be i.i.d. draws from the offspring distribution. Let $(X_1^{i,j}, X_2^{i,j})$, for each i, j, be i.i.d. with distribution $\mu^{(2)}$ (and independent of M and $\{M_i\}$). Then let $T^{(2)}(\mu^{(2)})$ be the distribution of (X_1, X_2), where

$$X_1 = \min_{1 \le i \le M} \max_{1 \le j \le M_i} X_1^{(i,j)},$$
$$X_2 = \min_{1 \le i \le M} \max_{1 \le j \le M_i} X_2^{(i,j)}.$$

Note particularly that the recursions for X_1 and X_2 use the *same* realisation of M and $\{M_i\}$.

If $\mu \in \mathcal{P}$ then we can define a *diagonal* distribution $\mu^{\nearrow}$ on $\mathcal{P}^{(2)}$ by $\mu^{\nearrow} = \mathrm{dist}(X, X)$ if $\mu = \mathrm{dist}(X)$, i.e. $\mu^{\nearrow}$ is a joint distribution of two identical copies of X.

If μ is a fixed point of T, then certainly $\mu^{\nearrow}$ is a fixed point of $T^{(2)}$. The question is whether there can be any fixed point of $T^{(2)}$, whose marginals are equal to μ, and which is *not* of the form of the diagonal distribution $\mu^{\nearrow}$. Mach, Sturm and Swart [47, Theorem 1], refining Aldous and Bandyopadhyay [4, Theorem 11], show that the recursive tree process is endogenous if and only if there are no such non-degenerate bivariate fixed points (i.e. if the 'bivariate uniqueness property' holds).

Proof of Theorem 2.5. We apply Theorem 1 of [47] (or indeed Theorem 11 of [4], since the additional technical condition relating to the continuity of the operator $T^{(2)}$ does in fact hold in this setting). To prove our result it is enough to show that the bivariate uniqueness property holds if and only if $f'(x) \le 1$.

Let us write μ for the Bernoulli$(1 - x)$ distribution on $\{0, 1\}$. We look for a distribution $\mu^{(2)}$ on $\{0, 1\}^2$ that is a fixed point of $T^{(2)}$, and whose marginals are both μ, but that is not the diagonal distribution $\mu^{\nearrow}$. Once these marginals are specified, we only need to specify one further parameter, say $b = \mu^{(2)}(1, 0)$, since then we can deduce $\mu^{(2)}(1, 1) = 1 - x - \mu^{(2)}(1, 0) = 1 - x - b$, and similarly $\mu^{(2)}(0, 1) = b$ and $\mu^{(2)}(0, 0) = x - b$. Note $b \in [0, \min(x, 1 - x)]$.

To show that $\mu^{(2)}$ is a fixed point of $T^{(2)}$, again it suffices to check just one entry of

$T^{(2)}(\mu^{(2)})$. To look at this we can consider a random tree with two levels, with bivariate marginals according to $\mu^{(2)}$ at level 2 of the tree; we wish to see the distribution $\mu^{(2)}$ again at the root. Then write also $\nu^{(2)}$ for the corresponding distribution of the marginals at level 1. Let us write o for the root and ι for a typical level-1 node.

So consider the probability of seeing values $(1,0)$ at the root. For this to happen, all children of the root must have 1 in the first coordinate, but at least one child of the root must have 0 in the second coordinate. That is, all children have values $(1,0)$ or $(1,1)$, but not all of them have values $(1,1)$. We obtain

$$\begin{aligned}\mathbb{P}(\text{values}(1,0) \text{ at } o) &= G(\nu^{(2)}(1,0)+\nu^{(2)}(1,1)) - G(\nu^{(2)}(1,1)) \\ &= R(\nu^{(2)}(1,1)) - R(\nu^{(2)}(1,0)+\nu^{(2)}(1,1)).\end{aligned} \tag{2.51}$$

We examine both terms on the RHS. First note that $\nu^{(2)}(1,0)+\nu^{(2)}(1,1)$ is the probability that ι has value 1 in the first coordinate. This is the probability that at least one child of ι has value 1 in the first coordinate, i.e. that not all the children of ι have value 0 in the first coordinate. Hence

$$\begin{aligned}\nu^{(2)}(1,0)+\nu^{(2)}(1,1) &= 1 - G(\mu(0,1)+\mu(0,0)) \\ &= 1 - G(x) \\ &= R(x).\end{aligned} \tag{2.52}$$

Similarly, for ι to have values $(1,1)$, we need to exclude the two events that all its children have value 0 in the first coordinate or that all its children have value 0 in the second coordinate. Both of these events have probability $G(x)$, while their intersection, i.e. that all children have values $(0,0)$, has probability $G(x-b)$. So applying inclusion-exclusion,

$$\begin{aligned}\nu^{(2)}(1,1) &= 1 - G(x) - G(x) + G(x-b) \\ &= 2R(x) - R(x-b).\end{aligned} \tag{2.53}$$

Combining (2.51), (2.52) and (2.53), we have that if the probability of values $(1,0)$ at level 2 is $b \in [0, \min(x, 1-x)]$, then the probability of values $(1,0)$ at the root is $h(b) \in$

$[0, \min(x, 1-x)]$, where

$$h(b) := R(2R(x) - R(x-b)) - R(R(x)). \tag{2.54}$$

For $\mu^{(2)}$ to be a fixed point of $T^{(2)}$, we therefore need $b = h(b)$. Also $\mu^{(2)}$ is diagonal iff $b = 0$. So non-endogeny is equivalent to the existence of a fixed point of h in the interval $(0, \min(x, 1-x)]$.

From (2.54) we have $h(0) = 0$, and differentiating with respect to b we get

$$h'(b) = R'(R(x) - [R(x-b) - R(x)])R'(x-b) \tag{2.55}$$

so that

$$\begin{aligned} h'(0) &= R'(R(x))R'(x) \\ &= \frac{d}{dx}R(R(x)) \\ &= f'(x). \end{aligned}$$

Differentiating once more we obtain

$$h''(b) = R''(2R(x) - R(x-b))R'(x-b)^2 - R'(2R(x) - R(x-b))R''(x-b). \tag{2.56}$$

Since R is positive, decreasing and strictly concave, it follows that (2.55) is positive and (2.56) is negative, hence that h is increasing and strictly concave.

So if $f'(x) \leq 1$, giving $h'(0) \leq 1$, then $h(u) < u$ for all $u > 0$. In that case the only non-negative fixed point of h is 0, and we must obtain $b = 0$. In that case the distribution $\mu^{(2)}$ must be a diagonal distribution, and we have bivariate uniqueness (and hence endogeny).

On the other hand, suppose that $f'(x) > 1$, so that $h'(0) > 1$. Then for sufficiently small $\epsilon > 0$, $h(\epsilon) > \epsilon$. Starting from some such ϵ and iterating h repeatedly gives an increasing sequence which is bounded above by $\min(x, 1-x)$. Its limit is a fixed point of h which lies in $(0, \min(x, 1-x)]$. Hence in this case there does exist a non-degenerate bivariate fixed point, and the process is non-endogenous, as required. □

Proof of Corollary 2.6. Since $f'(x) = 1$ everywhere, Theorem 2.5 tells us that all the processes with Bernoulli marginals are endogenous. This implies that for any μ, for the process with marginals μ, the event $\{Y \leq y\}$ is measurable with respect to the structure of the tree, for any y, where Y is the value at the root. But then in fact the random variable Y is measurable with respect to the structure of the tree, as required. $\square$

Chapter 3

Derivative martingale of the branching Brownian motion in dimension $d \geq 1$

Based on joint work with Julien Berestycki and Bastien Mallein [63]

Abstract

We consider a branching Brownian motion in $\mathbb{R}^d$. We prove that there exists a random subset Θ of $\mathbb{S}^{d-1}$ of full surface measure, such that the limit of the derivative martingale exists simultaneously for all directions $\theta \in \Theta$ almost surely. This allows us to define a random measure on $\mathbb{S}^{d-1}$ whose density is given by the derivative martingale.

The proof is based on first moment arguments: we approximate the martingale of interest by a series of processes, which do not take into account particles that travelled too far away. We show that these new processes are uniformly integrable martingales whose limits can be made to converge to the limit of the original martingale.

3.1 Introduction

Consider a branching Brownian motion in dimension $d \geq 1$. This is a particle system in which independent particles move in $\mathbb{R}^d$ as Brownian motions and branch independently at rate 1 into two particles. This system behaves as a growing cloud of diffusing particles. Let us fix the notation. We denote by $\mathbb{P}_x$ the law of the branching Brownian motion, or the Brownian motion (which will be clear from the context), starting from one particle at position $x \in \mathbb{R}^d$, (writing $\mathbb{P}$ for $\mathbb{P}_0$ for simplicity). To avoid ambiguity, we shall write $\mathbb{E}_{\mathbb{P}}$ for the expectation under the measure $\mathbb{P}$, and $\mathbb{E}_x$ for expectation under the measure $\mathbb{P}_x$. The Brownian motion at time t will be denoted by B_t. For all times $t \geq 0$, we denote by $\mathcal{N}_t$ the set of particles alive at time t, and for each particle $j \in \mathcal{N}_t$ and $s \leq t$, we write $X_s(j)$ for the position that j, or its ancestor at time s, occupied at time s. The natural filtration of the branching Brownian motion is denoted by $(\mathcal{G}_t, t \geq 0)$.

In [50], Mallein studied the maximal displacement of this model, i.e. the quantity

$$R_t = \max_{j \in \mathcal{N}_t} \|X_t(j)\|, \quad t \geq 0.$$

He showed that as $t \to \infty$

$$R_t = \sqrt{2}t + \frac{d-4}{2\sqrt{2}} \log t + O(1), \tag{3.1}$$

where $O(1)$ is a process Y_t such that $\lim_{K \to \infty} \mathbb{P}(\sup_t |Y_t| > K) = 0$, thus generalising a famous result of Bramson [18] for $d = 1$.

Imagine now that we want to know in which direction $D(t)$ is the particle at distance R_t at time t. Under $\mathbb{P}_0$, the process is completely spherically symmetric and it is thus evident that the distribution of the direction $D(t)$ of this extremal particle is uniform on the sphere $\mathbb{S}^{d-1}$. However, if we first observe the process up to time s and then try to guess the direction of the furthest particle at a later time t, the answer obviously depends on the configuration we observe at time s, even in the limit $t \to \infty$. Advantages gained or delays incurred early in a given direction are never forgotten.

It is believed that almost surely, for all measurable sets $A \subset \mathbb{S}^{d-1}$

$$\lim_{s \to \infty} \lim_{t \to \infty} \mathbb{P}(D(t) \in A \mid \mathcal{G}_s) = \mu(A),$$

where μ is a random probability measure which captures what happens early on in the life of the process. What should this measure be?

To answer this question, it is instructive to look at the one-dimensional case. When $d = 1$, it is well-known that the asymptotic behaviour of the extremal particles (i.e. particles within distance $O(t^{1/2})$ from the maximal displacement at time t) is mainly driven by the limit of the so-called derivative martingale, defined by

$$Z_t^+ := \sum_{j \in \mathcal{N}_t} (\sqrt{2}t - X_t(j)) e^{\sqrt{2}(X_t(j) - \sqrt{2}t)}.$$

Although $(Z_t^+, t \geq 0)$ is known to be a non-uniformly integrable martingale, and clearly takes both positive and negative values, Lalley and Sellke [44] proved that it does have an almost sure limit $Z_\infty^+ := \lim_{t\to\infty} Z_t^+$ which is positive almost surely, and moreover

$$\max_{j \in \mathcal{N}_t} X_t(j) - m_t - \frac{\sqrt{2}}{2} \log Z_\infty^+$$

converges in law to a Gumbel random variable, where $m_t = \sqrt{2}t - \frac{3}{2\sqrt{2}} \log t$.

We introduce the maximal and minimal displacements, i.e. the largest displacement in the positive and negative direction:

$$M_t^+ := \max_{j \in \mathcal{N}_t} X_t(j) \quad \text{and} \quad M_t^- := \min_{j \in \mathcal{N}_t} X_t(j),$$

as well as the derivative martingale in the negative direction, which is the derivative martingale of the BBM $(-X_t(j), j \in \mathcal{N}_t)$. In other words, we set

$$Z_t^- := \sum_{j \in \mathcal{N}_t} (\sqrt{2}t + X_t(j)) e^{\sqrt{2}(X_t(j) + \sqrt{2}t)}$$

and $Z_\infty^- := \lim_{t\to\infty} Z_t^-$. As far as we are aware, the joint convergence in distribution of (M_t^+, M_t^-) had not been considered until now.

Theorem 3.1. *There exists a constant $c_\star$ such that for all $y, z \geq 0$ almost surely*

$$\lim_{s\to\infty} \lim_{t\to\infty} \mathbb{P}\left(M_t^+ - m_t \leq y, -M_t^- - m_t \leq z \,\middle|\, \mathcal{G}_s \right) = \exp\left(-c_\star Z_\infty^+ e^{-\sqrt{2}y} - c_\star Z_\infty^- e^{-\sqrt{2}z} \right).$$

In other words, $\left(M_t^+ - m_t - \frac{\sqrt{2}}{2} \log\left(c_\star Z_\infty^+\right), -M_t^- - m_t - \frac{\sqrt{2}}{2} \log\left(c_\star Z_\infty^-\right) \right)$ converges in distribution towards a pair of independent Gumbel random variables with scale parameter $\frac{\sqrt{2}}{2}$.

As a consequence, conditionally on (Z_∞^+, Z_∞^-) the probability that the direction of the furthest particle at a large time is in the positive direction is proportional to Z_∞^+.

Corollary 3.2. *We have*

$$\lim_{s\to\infty}\lim_{t\to\infty}\mathbb{P}\left(M_t^+ > -M_t^- \;\middle|\; \mathcal{G}_s\right) = \frac{Z_\infty^+}{Z_\infty^- + Z_\infty^+} \quad a.s.$$

It is straightforward from the definition of the branching Brownian motion, that for all $\theta \in \mathbb{S}^{d-1}$, its projection on the direction θ (the process $(X_t(j)\cdot\theta, j \in \mathcal{N}_t)$) is a branching Brownian motion in dimension one. Thus, for each $\theta \in \mathbb{S}^{d-1}$ we can define the derivative martingale of X in direction θ as

$$Z_t(\theta) := \sum_{j\in\mathcal{N}_t} (\sqrt{2}t - X_t(j)\cdot\theta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)} \tag{3.2}$$

and for each $\theta \in \mathbb{S}^{d-1}$, the limit $\lim_{t\to\infty} Z_t(\theta) = Z_\infty(\theta)$ exists a.s.

Coming back to the direction $D(t)$ of the extremal particle, it is natural to think that, as in dimension one, the random measure μ should give more mass to regions where $Z_\infty(\theta)$ is large. In fact, μ should have a density given by the normalized version of $\theta \mapsto Z_\infty(\theta)$. That is, for a measurable set $B \subset \mathbb{S}^{d-1}$, we would expect $\mu(B) = \int_B Z_\infty(\theta)\sigma(\mathrm{d}\theta)/\int_{\mathbb{S}^{d-1}} Z_\infty(\theta)\sigma(\mathrm{d}\theta)$, where $\sigma(\mathrm{d}\theta)$ stands for the surface measure of $\mathbb{S}^{d-1}$.

However, the problem is that we do not have a.s. existence of the limit $Z_\infty(\theta)$ for all $\theta \in \mathbb{S}^{d-1}$ simultaneously and so the above integrals are not *a priori* well defined. Observe for instance that by (3.1) one has

$$\inf_{\theta\in\mathbb{S}^{d-1}} Z_t(\theta) \le -C(\log t)t^{(d-4)/2} \text{ with high probability,}$$

hence the derivative martingale may be very small in exceptional directions, at least in dimension $d \ge 4$. This is due to the fact that in higher dimensions particles travel farther away from 0 than in dimension one (which follows from the Pythagoras' Theorem and an observation that coordinates of a multidimensional Brownian motion are i.i.d. one-dimensional Brownian motions), which has the effect of lowering the value of $Z_t(\theta)$ in the (random) direction at which these far away particles are located. As a result, one cannot hope for uniform convergence to hold for the process $(Z_t(\theta))$. It is nonetheless the main object of this chapter to show how one can make sense of the limit of the function $\theta \mapsto Z_t(\theta)$ in a weak sense.

We also prove that almost surely the limit of $Z_t(\theta)$ actually exists for all θ in a set of full measure. Hence a rigorous meaning can be given to the associated measure μ.

In this chapter we prove the weak convergence of $(Z_t(\theta)\sigma(\mathrm{d}\theta), \theta \in \mathbb{S}^{d-1})_{t\geq 0}$, seen as a random measure on the sphere. For two measurable functions $f, g : \mathbb{S}^{d-1} \mapsto \mathbb{R}$ we define

$$\langle f, g \rangle := \int_{\mathbb{S}^{d-1}} f(\theta)g(\theta)\sigma(\mathrm{d}\theta),$$

where σ is the Lebesgue measure on the sphere $\mathbb{S}^{d-1}$. We sometimes write $\langle f(\theta), g(\theta) \rangle$ to clarify how functions f and g depend on $\theta \in \mathbb{S}^{d-1}$.

The main result of this chapter is the following.

Theorem 3.3. *Almost surely there exists a measurable subset* Θ *of* $\mathbb{S}^{d-1}$ *of full measure (i.e.* $\sigma(\Theta) = \sigma(\mathbb{S}^{d-1})$*), such that* $Z_\infty(\theta) := \lim_{t\to\infty} Z_t(\theta)$ *exists for* $\theta \in \Theta$*, and for any bounded measurable function* f

$$\lim_{t\to\infty} \langle Z_t, f \rangle = \langle Z_\infty, f \rangle \ a.s., \tag{3.3}$$

writing $Z_\infty(\theta) = 0$ *for all* $\theta \notin \Theta$*. Additionally,* $0 < \lim_{t\to\infty}\langle Z_t, 1 \rangle < \infty$ *almost surely.*

Although we only consider the case of a *binary* branching mechanism (particles always split into two daughter particles), it would be straightforward to generalise our results to a situation in which an independent random number L of children is produced at each branching event, at least under the assumption $\mathbb{E}(L(\log L)^{2+\delta}) < \infty$ for some $\delta > 0$. Note that it was shown by Yang and Ren [66] that, in the case of the one-dimensional branching Brownian motion, the limit of the derivative martingale is non-degenerate if and only if $\mathbb{E}(L(\log L)^2) < \infty$. This result was then extended by Chen [25] to the case of branching random walks and recently Boutaud and Maillard simplified and streamlined the proofs of these limit theorems in [16]. We believe Theorem 3.3 would hold under similar optimal integrability conditions, but the proof would require additional control on the law of a Brownian motion conditioned to stay below a curve.

Let us now formulate a conjecture regarding the full extremal point process, from which the predicted behaviour of $D(t)$ follows. This conjecture is a multidimensional version of the description of the extremal point process of the one-dimensional branching Brownian motion obtained by Arguin Bovier and Kistler [9], and Aïdékon, Berestycki, Brunet and Shi

[1]. Recall from [50, Theorem 1.1] that

$$r_t := \sqrt{2}t + \frac{d-4}{2\sqrt{2}} \log t$$

is, up to an $O(1)$ error, the median of the maximal displacement of the d-dimensional branching Brownian motion, i.e. there exists $C_t = O(1)$ such that for all $t > 0$

$$\mathbb{P}\left(\max_{u\in\mathcal{N}_t} ||X_t(u)|| \le r_t + C_t\right) = \mathbb{P}\left(\max_{u\in\mathcal{N}_t} ||X_t(u)|| \ge r_t + C_t\right) = 1/2.$$

We also define the direction of a particle u at time t by $D_t(u) := X_t(u)/\|X_t(u)\|$ for $t \ge 0$, $u \in \mathcal{N}_t$.

Conjecture 3.4. *There exists $c_d^\star > 0$ such that*

$$\lim_{t\to\infty} \sum_{u\in\mathcal{N}_t} \delta_{D_t(u),\|X_t(u)\|-r_t} = \mathcal{L}(\mathrm{d}\theta, \mathrm{d}x) \text{ in law,}$$

where $\mathcal{L}$ is a decorated Poisson point process that can be constructed as follows. Let $(\theta_j, \xi_j)_{j\ge1}$ be the atoms of a Poisson point process with intensity $c_d^\star Z_\infty(\theta)\sigma(\mathrm{d}\theta)\mathrm{e}^{-\sqrt{2}x}\mathrm{d}x$ and $(D_j, j \ge 1)$ be i.i.d. point processes on $\mathbb{R}$ with common distribution $\mathcal{D}$. Then

$$\mathcal{L} = \sum_{j\ge1} \sum_{x\in D_j} \delta_{\theta_j,\xi_j+x}.$$

To be more explicit, the decoration point measure $\mathcal{D}$ above can be constructed as the weak limit of $\sum_{u\in\mathcal{N}_t} \delta_{\|X_t(u)\|-R_t}$ (the extremal process of moduli seen from the largest displacement) conditioned on $R_t \ge r_t + \frac{3}{2\sqrt{2}} \log t$ (c.f. [64] for a general result of convergence towards decorated Poisson point processes). In particular, $\mathcal{D}$ only charges $(-\infty, 0]$.

Let us discuss briefly some implications that would follow from Conjecture 3.4. Firstly, an easy Poisson point process computation would yield that

$$\lim_{t\to\infty} \mathbb{P}(R_t - r_t \le x) = \mathbb{E}_{\mathbb{P}}\left[\exp\left(-c_d^\star\langle Z_\infty, 1\rangle e^{-\sqrt{2}x}\right)\right].$$

This is the multidimensional version of [44] that gives the convergence in law of the maximum of the branching Brownian motion. Similarly, it would imply the following convergence for the law of the direction of the furthest particle at time t:

$$\lim_{s\to\infty}\lim_{t\to\infty}\mathbb{P}\left(D(t)\in B \mid \mathcal{G}_s\right)=\frac{1}{\langle Z_\infty,1\rangle}\int_B Z_\infty(\theta)\sigma(\mathrm{d}\theta)\qquad \text{a.s., } B\subseteq\mathbb{S}^{d-1}.$$

3.2 Proof strategy

Let us now review briefly how these results are proved in dimension $d=1$. The idea is to get rid of all particles that ever reach level $\sqrt{2}t+A$ at some time t (this is sometimes referred to as a *shaving argument*). However, as we push the barrier away by letting $A\to\infty$ the probability that any particle ever hits the barrier decreases to zero. More formally, one introduces the martingale

$$Z_t^A:=\sum_{j\in\mathcal{N}_t^A}(\sqrt{2}t+A-X_t(j))e^{\sqrt{2}(X_t(j)-\sqrt{2}t)},$$

where $\mathcal{N}_t^A=\{j\in\mathcal{N}_t : X_s(j)\le\sqrt{2}s+A, s\le t\}$. This martingale is non-negative (and uniformly integrable) and therefore converges to some Z_∞^A. Since in dimension one

$$\sup_{t\ge0}\sup_{j\in\mathcal{N}_t}|X_t(j)|-\sqrt{2}t<\infty\qquad\text{almost surely,}$$

hence taking A large enough ensures that no particle is killed with high probability. This proves that the derivative martingale converges and that almost surely $Z_\infty=\lim_{A\to\infty}Z_\infty^A$. Note that in higher dimensions similarly to the derivative martingale $Z_t(\theta)$ defined in equation (3.2) we can define $Z_t^A(\theta)$ as a function of the process $(X_t(j)\cdot\theta, j\in\mathcal{N}_t)$. Therefore, in every direction the limit $Z_\infty(\theta)=\lim_{A\to\infty}Z_\infty^A(\theta)$ exists almost surely.

In larger dimensions ($d\ge4$), however, one has

$$\sup_{t\ge0}\sup_{j\in\mathcal{N}_t}\|X_t(j)\|-\sqrt{2}t=\infty, \tag{3.4}$$

and this is the moment where the standard argument breaks. Even though for every direction θ we have that $Z_t(\theta)$ converges almost surely and $Z_\infty(\theta)=\lim_{A\to\infty}Z_\infty^A$, in larger dimensions ($d\ge4$) at each time t there are directions with displacements larger than $\sqrt{2}t$, and therefore in these directions the value of $Z_t(\theta)$ is much lower than the one of $Z_t^A(\theta)$.

To overcome this difficulty we need to introduce a different way of removing particles that fly too high in the one-dimensional branching Brownian motion. This is done by killing particles that reach a curved boundary $\sqrt{2}t + (\phi(t) \vee A)$ at some time t, with ϕ a well-chosen non-decreasing function. In particular, if ϕ grows fast enough, we can ensure that no particle will be removed with high probability by letting $A \to \infty$. The difficulty is then to find an analogue of the martingale Z^A for this curved boundary. We shall then apply this non-linear shaving to the projection of the BBM onto each direction separately.

The outline of the chapter is as follows: In Section 3.3 we study the standard one-dimensional Brownian motion killed when hitting the barrier $t \mapsto \phi(t) \vee A$. In Lemma 3.9 we prove in particular existence of some function $R^\phi(x,t)$ allowing us to describe the Brownian motion conditioned to stay below $\phi \vee A$ as a Doob h-transform. In Lemma 3.12 we show that this conditioned process behaves similarly to the Bessel process, i.e. its trajectory is asymptotically of the form $-t^{1/2+o(1)}$. In the next step we study properties of the function R^ϕ mentioned above. We shall use R^ϕ later to construct an approximation of the derivative martingale Z_t, therefore in Lemma 3.17 we prove its key property, i.e. that $\mathbb{R}^\phi(x,t)$ is asymptotically linear in $-x$, uniformly in t. This allows us to show that when $t \to \infty$, the limit of our approximation agrees with the limit of the derivative martingale with high probability.

In Section 3.4 we move to the setting of the branching Brownian motion. In Proposition 3.18, based on R^ϕ we define for the one-dimensional BBM an approximation Z_t^ϕ of the derivative martingale Z_t. We then define $Z_t^\phi(\theta)$ in each direction θ analogously to the definition of $Z_t(\theta)$ in equation (3.2). In Lemma 3.20 we find a sufficiently slowly increasing function $\tilde{r}(t)$ such that with high probability non of the projections on $\theta \in \mathbb{S}^{d-1}$ of any of the particles of the BBM ever reaches $\tilde{r}$. This implies that for a function ϕ growing faster than $\tilde{r}$, only for finitely many particles there exists a direction θ such that this particle hits the barrier $\sqrt{2}t + \phi(t)$, which means that our non-linear shaving overcomes the main issue of the linear shaving when applied in higher dimensions.

In Lemma 3.23 we use a spinal decomposition for the multidimensional branching Brownian motion and Bessel-type fluctuations for the one-dimensional Brownian motion conditioned not to hit ϕ, which we proved in Lemma 3.12, to show that the integrated martingale $\langle Z_t^\phi, f\rangle = \int_{\mathbb{S}^{d-1}} Z_t^\phi(\theta) f(\theta)\sigma(\mathrm{d}\theta)$ is uniformly integrable. In Proposition 3.26 we show that $Z_t^\phi(\theta)$ almost surely converges *simultaneously* for all θ from a random subset of $\mathbb{S}^{d-1}$ of full measure, and that $\lim_{t\to\infty}\langle Z_t^\phi, f\rangle = \langle Z_\infty^\phi, f\rangle$ almost surely. We use this result for a sequence of shaved martingales based on functions $(\phi(t) \vee A)_{A\in\mathbb{N}}$ to complete the proof of Theorem 3.3.

In Section 3.5 we treat the one-dimensional case and look at the joint law of the leftmost and rightmost particles in the branching Brownian motion.

3.3 One-dimensional Brownian motion conditioned to avoid $\phi(t)$

To prove Theorem 3.3, as explained in Section 3.2, we will need some estimates on the one-dimensional Brownian motion, which we shall denote by B_t, conditioned to stay below a curve. Unless otherwise stated, we assume that $B_0 = 0$. In this section we gather several results on this process, using Doob's h-transform theory.

Let ϕ be a $\mathcal{C}^1$-class increasing function $[0, \infty) \to \mathbb{R}$ such that $\phi(t) = o(t^{1/2-\epsilon})$ for some $\epsilon > 0$. We start by studying the Brownian motion conditioned not to hit the function ϕ from below until some finite time t. As the fluctuations of B_t, which are of order $t^{1/2}$, are much larger than $\phi(t)$, we expect that for $1 \ll s \ll t$ the Brownian motion on $[0, s]$ conditioned on not hitting ϕ until time t behaves roughly like a reflected Bessel process (i.e. a Brownian motion conditioned to stay below 0). More precisely, in Lemma 3.9 we introduce the relevant non-negative h-transform function R^ϕ defined as the renormalized probability of avoiding ϕ, and show that

$$(R^\phi(B_t, t)\mathbb{I}_{\{\forall s<t, B_s\leq\phi(s)\}})_{t\geq 0}$$

is a $\mathbb{P}$-martingale. In other words, R^ϕ is a harmonic function for the Markov process (B_t, t) confined to $\{(x, t) : x \leq \phi(t)\}$. The Doob h-transform obtained then describes a Brownian motion conditioned to stay below ϕ; we are going to denote the corresponding measure as $\mathbb{P}^\phi$. It will also be important to show that there exists $C > 0$ such that $R^\phi(x, t) \approx -Cx$ as $x \to -\infty$, as this will entail the 'Bessel-like' behaviour, i.e. that under $\mathbb{P}^\phi$ the trajectory of the Brownian motion is asymptotically of the form $-t^{1/2+o(1)}$.

Note that we want to condition the Brownian motion to stay below $\phi(t)$, where $\phi(t) = o(t^{1/2})$, and not below $\sqrt{2}t + \phi(t)$. This is because we shall apply these results later on to the spine of the BBM (i.e. some distinguished particle), which under some new measure in some direction θ_0 is going to behave like a Brownian motion with a drift $\sqrt{2}$ conditioned to stay below $\sqrt{2}t + \phi(t)$, i.e. effectively a Brownian motion conditioned to stay below $\phi(t)$.

The function R^ϕ will then be used to define approximations of the derivative martingale of the one-dimensional branching Brownian motion. Indeed, we wish to define a uniformly integrable martingale that approximates the derivative martingale

$$Z_t = \sum_{j\in\mathcal{N}_t} (\sqrt{2}t - X_t(j))e^{\sqrt{2}(X_t(j)-\sqrt{2}t)}, \tag{3.5}$$

that would be of the form

$$\sum_{j\in\mathcal{N}_t^\phi} H(X_t(j) - \sqrt{2}t, t)e^{\sqrt{2}(X_t(j)-\sqrt{2}t)} \tag{3.6}$$

where H is some function and $\mathcal{N}_t^\phi = \{j \in \mathcal{N}_t : X_s(j) \leq \sqrt{2}s + \phi(s),\ \forall s \leq t\}$ (so that the sum in (3.6) is taken only over the particles that did not hit the boundary $\sqrt{2}s + \phi(s)$). Recall the *many-to-one* lemma (a corollary of Lemma 1 in [37]):

Lemma 3.5. *For any $t \geq 0$ and for any measurable function ξ we have*

$$\mathbb{E}_\mathbb{P}\left[\sum_{j\in\mathcal{N}_t} \xi(X_u(j), 0 \leq u \leq t)\right] = e^t\mathbb{E}_\mathbb{P}[\xi(B_u, 0 \leq u \leq t)].$$

Applying the many-to-one lemma and using the Girsanov transform to add drift $\sqrt{2}$, one can see that assuming that (3.6) is a $\mathbb{P}$-martingale is equivalent to assuming that $(H(B_t, t)\mathbb{I}_{\{\forall s<t, B_s\leq\phi(s)\}})_{t\geq 0}$ is itself a $\mathbb{P}$-martingale. Hence, setting $H(x,t) = C^{-1}R^\phi(x,t)$ gives the desired approximation of (3.5).

The rest of the section is organised as follows. After proving Lemma 3.9, we define the measure $\mathbb{P}^\phi$ in equation (3.11). We characterise $\mathbb{P}^\phi$ in Lemma 3.10 as a limit of conditional distributions. In Lemma 3.11 we define a new measure $\mathbb{P}^V$, that corresponds to a Girsanov transform adding a drift $\sqrt{2}$ applied to a process with law $\mathbb{P}^\phi$. That is, we can interpret $\mathbb{P}^V$ as a measure of a Brownian motion with a drift $\sqrt{2}$ conditioned on never hitting $\sqrt{2}t + \phi(t)$. In Lemma 3.12 we formalize the 'Bessel-like' behaviour under $\mathbb{P}^\phi$. Finally, in Lemma 3.17 we study asymptotics of $R^\phi(x,t)$.

3.3.1 Brownian motion and non-linear barriers

For any continuous function $\phi : [0, \infty) \to \mathbb{R}$ set $\tau_\phi = \inf\{u > 0 \ : B_u \geq \phi(u)\}$, where B is a standard one-dimensional Brownian motion. The aim of this section is to give a precise asymptotic of the quantity $\mathbb{P}_x(\tau_\phi > t)$ as $t \to \infty$ for ϕ in a certain class. We are also interested in the dependence of this quantity on the shift of ϕ, i.e. we are going to consider functions $\phi_t(u) := \phi(t + u)$.

It is well-known that if ϕ grows slower than $t^{1/2}$ as $t \to \infty$, in a sense to be made precise soon, then $\tau_\phi < \infty$ a.s. and $\mathbb{P}(\tau_\phi > t)$ decays as $t^{-1/2}$. More precisely, Uchiyama proved the following upper bound.

Theorem 3.6 ([65], Proposition 3.1. (i))**.** *Let ϕ be a $\mathcal{C}^1$-class increasing function such that $\phi(0) = 0$ and $\lim_{t\to\infty} \phi(t)t^{-1/2} = 0$. If*

$$\phi(u) - \frac{u}{t}\phi(t) \geq 0 \quad \text{for } 0 < u < t,$$

then there exists a constant C such that for all $x \in \mathbb{R}$ and $t > 1$

$$\mathbb{P}_x(\tau_\phi > t) \leq \frac{1 + |x|}{t^{1/2}} \exp\left(\frac{\sqrt{2\pi}}{4} \int_1^t \frac{\phi(u)}{u^{3/2}} \mathrm{d}u + C \int_1^t \frac{\phi(u)^2}{u^2} \mathrm{d}u\right). \tag{3.7}$$

Novikov [56] obtained a precise asymptotic of $\mathbb{P}(\tau_\phi > t)$ as $t \to \infty$, expressed as a function of the law of $B_{\tau_\phi} = \phi(\tau_\phi)$.

Theorem 3.7 ([56], Theorem 2)**.** *If ϕ is a continuous non-decreasing function such that $\int_1^\infty \phi(t)t^{-3/2}dt < \infty$ and $\phi(0) > 0$, then*

$$\lim_{t\to\infty} \sqrt{t}\mathbb{P}(\tau_\phi > t) = \sqrt{\frac{2}{\pi}} \mathbb{E}_\mathbb{P} B_{\tau_\phi} < \infty.$$

We apply these two theorems to define and give the first property of the aforementioned function R^ϕ, which will be a key object of interest in the rest of the chapter. We will restrict ourselves to functions ϕ satisfying the following assumptions:

Assumption 3.8.

ϕ increasing, concave, $\mathcal{C}^1$-class with $\phi(0) > 0$,

$$\text{and there exists } \alpha \in (0, 1/2) \text{ such that } \lim_{t\to\infty} \frac{\phi(t)}{t^\alpha} = 0.$$

Recall that $\tau_{\phi_t} = \inf\{u > 0 \ : B_u \geq \phi(t+u)\}$.

Lemma 3.9. *Let ϕ be a function satisfying Assumption 3.8. Then the following limit exists for all $t \geq 0$ and $x \in \mathbb{R}$:*

$$R^\phi(x,t) := \sqrt{\frac{\pi}{2}} \lim_{s\to\infty} \sqrt{s}\mathbb{P}_x(\tau_{\phi_t} > s). \tag{3.8}$$

Moreover, there exists $C > 0$ such that for all $t \geq 0$ and $x \leq \phi(t)$,

$$R^\phi(x,t) \leq C(1 + (\phi(t) - x)).$$

Finally, $\left(R^\phi(B_t, t)\mathbb{I}_{\{\tau_\phi > t\}}\right)_{t\geq 0}$ is a $\mathbb{P}$-martingale.

$R^\phi(x,t)$ can be seen as a renormalized survival probability of a Brownian motion starting at time t from position x. The idea of using the renormalized survival probability to define an h-transform is classical. Here we draw inspiration from [11] (in which the law of a random walk conditioned to stay positive was constructed). There, as in the present work, we condition a random process not to hit some region (in our case the process of interest is $(B_t, t)_{t\geq 0}$ and the region to avoid is defined by $J = \{(x,t) : x \geq \phi(t)\}$).

In this setting the probability that the process $(B_t, t)_{t\geq 0}$ never hits J is equal to 0, irrespectively of its starting position (x_0, t_0). As a result, to define the h-transform in the sense of Doob, we need to renormalise the probability for the process (B_t, t) not to hit the region J for t units of time by $t^{1/2}$ so that the limit, that we denote by $R^\phi(x_0, t_0)$, is non-degenerate. It remains to check that the function R^ϕ which we defined is indeed a harmonic function for (B_t, t) on the domain J^c, i.e. that $\left(R^\phi(B_t, t)\mathbb{I}_{\{\tau_\phi > t\}}\right)_{t\geq 0}$ is a $\mathbb{P}$-martingale.

Proof. The assumptions on ϕ guarantee that Theorem 3.6 and Theorem 3.7 can be applied to the function ϕ_t for all $t \geq 0$. We note that for all $t \geq 0$, $s \geq 0$ and $x \geq \phi(t)$, we have

$$\mathbb{P}_x(\tau_{\phi_t} > s) = 0.$$

Note that $\tau_{(\phi_t - x)} = \inf\{u > 0\ : B_u + x \geq \phi(t+u)\}$. Applying Theorem 3.7 to the function ϕ_t, we deduce that for all x, t such that $x < \phi(t)$,

$$R^\phi(x,t) = \mathbb{E}_{\mathbb{P}}\left(B_{\tau_{(\phi_t - x)}}\right) \in (0, \infty),$$

which proves that R^ϕ is well-defined and finite. Additionally, using that ϕ is concave, and hence that $\phi(t+u) - \phi(t) \leq \phi(u) - \phi(0)$, we observe that for all $x \in \mathbb{R}$, $t \geq 0$ and $s > 0$, we have

$$\sqrt{s}\mathbb{P}_x(\tau_{\phi_t} > s) = \sqrt{s}\mathbb{P}_{x-\phi(t)}\left(\tau_{\phi_t - \phi(t)} > s\right) \leq \sqrt{s}\mathbb{P}_{x-\phi(t)}\left(\tau_{\phi - \phi(0)} > s\right).$$

Using Theorem 3.6, and observing that the exponential term in bound (3.7) is increasing in t, and hence may be bounded from above by its limit as $t \to \infty$, we obtain for $x \leq \phi(t)$ and $s \geq 0$,

$$\sqrt{s}\mathbb{P}_{x-\phi(t)}\left(\tau_{\phi - \phi(0)} > s\right) \leq C(1 + (\phi(t) - x)), \tag{3.9}$$

where $C > 0$ is a constant that does not depend on x, t, s.

Thanks to this bound, we can now prove that $(R^\phi(B_t, t)\mathbb{I}_{\{\tau_\phi > t\}}, t \geq 0)$ is a $\mathbb{P}$-martingale using the Dominated Convergence Theorem. Indeed, using the Markov property, note that it is enough to prove that for all $t, s \geq 0$ and $x \leq \phi(t)$,

$$R^\phi(x,t) = \mathbb{E}_x\left[R^\phi(B_s, t+s)\mathbb{I}_{\{\tau_{\phi_t} > s\}}\right].$$

Observe that by the Markov property of the Brownian motion, for all $r \geq 0$

$$\mathbb{P}_x(\tau_{\phi_t} > s + r) = \mathbb{E}_x\left[\mathbb{I}_{\{\tau_{\phi_t} > s\}}\mathbb{P}_{B_s}(\tau_{\phi_{t+s}} > r)\right]. \tag{3.10}$$

By definition, we have $\sqrt{r\pi/2}\mathbb{P}_x(\tau_{\phi_t} > s+r) \to R^\phi(x,t)$ as $r \to \infty$, and similarly, we have

$$\lim_{r\to\infty} \sqrt{r\pi/2}\mathbb{P}_{B_s}(\tau_{\phi_{t+s}} > r) = R^\phi(B_s, t+s) \text{ a.s.}$$

We now observe that by (3.9) we can bound $\sqrt{r}\mathbb{P}_{B_s}(\tau_{\phi_{t+s}} > r)$ uniformly in $r \geq 0$ by $C(1+|B_s|+|\phi(t+s)|)$. This quantity being integrable, letting $r \to \infty$, and applying Lebesgue's Dominated Convergence Theorem in (3.10) we get

$$R^\phi(x,t) = \mathbb{E}_x\left[R^\phi(B_s, t+s)\mathbb{I}_{\{\tau_{\phi_t}>s\}}\right],$$

which completes the proof. □

As mentioned above, the function R^ϕ can be used to construct the Brownian motion conditioned to stay below ϕ in the sense of Doob, as a process with law $\mathbb{P}^\phi$ defined by

$$\left.\frac{\mathrm{d}\mathbb{P}^\phi}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} := \frac{R^\phi(B_t,t)}{R^\phi(0,0)}\mathbb{I}_{\{\tau_\phi>t\}}, \tag{3.11}$$

using the fact that $R^\phi(B_t,t)\mathbb{I}_{\{\tau_\phi>t\}}$ is a non-negative $\mathbb{P}$-martingale with mean $R^\phi(0,0)$. The law $\mathbb{P}^\phi$ corresponds to the limit of the law of the Brownian motion on the time interval $[0,t]$ conditioned on $\tau_\phi > s$ when $s \to \infty$. More precisely, it can be characterized in the following way.

Proposition 3.10. *Assume that ϕ satisfies Assumption 3.8. For any $t > 0$ and $A \in \mathcal{F}_t$,*

$$\mathbb{P}^\phi(A) = \lim_{s\to\infty} \mathbb{P}(A \mid \tau_\phi > s).$$

The proof of Proposition 3.10 is inspired by ideas from the proof of Theorem 1 in [11].

Proof. Let $A \in \mathcal{F}_t$. We observe that for $s > t$,

$$\mathbb{P}(A \mid \tau_\phi > s) = \frac{\mathbb{P}(A, \tau_\phi > s)}{\mathbb{P}(\tau_\phi > s)} = \frac{\mathbb{E}_\mathbb{P}(\mathbb{I}_A\mathbb{I}_{\{\tau_\phi>t\}}\mathbb{P}_{B_t}(\tau_{\phi_t} > s-t))}{\mathbb{P}(\tau_\phi > s)}.$$

Then by (3.8), we have that $\lim_{s\to\infty}\sqrt{s\frac{\pi}{2}}\mathbb{P}(\tau_\phi > s) = R^\phi(0,0)$ and

$$\lim_{s\to\infty}\sqrt{s\frac{\pi}{2}}\mathbb{P}_{B_t}(\tau_{\phi_t} > s-t)) = R^\phi(B_t,t) \quad \text{a.s.}$$

Moreover, using (3.9), we can apply Lebesgue's Dominated Convergence Theorem to obtain

$$\lim_{s\to\infty}\sqrt{s\frac{\pi}{2}}\mathbb{E}_{\mathbb{P}}(\mathbb{I}_A\mathbb{I}_{\{\tau_\phi>t\}}\mathbb{P}_{B_t}(\tau_{\phi_t} > s-t)) = \mathbb{E}_{\mathbb{P}}(\mathbb{I}_A\mathbb{I}_{\{\tau_\phi>t\}}R^\phi(B_t,t)).$$

As a result, for any fixed t we have

$$\lim_{s\to\infty}\mathbb{P}(A \mid \tau_\phi > s) = \frac{1}{R^\phi(0,0)}\mathbb{E}_{\mathbb{P}}(\mathbb{I}_A\mathbb{I}_{\{\tau_\phi>t\}}R^\phi(B_t,t)) = \mathbb{P}^\phi(A),$$

by definition. □

To complete the section, note that one can make a Girsanov-type change of measure to give the Brownian motion we consider a linear drift. This additional change of measure will be used when working with a multidimensional BBM. In particular, in Lemma 3.22 we describe a decomposition of the size-biased law of the BBM with a spine particle that in some randomly chosen direction θ_0 behaves like a Brownian motion with drift $\sqrt{2}$ conditioned not to hit $\sqrt{2}t+\phi(t)$ for all $t \geq 0$.

More precisely, we introduce the hitting time

$$\tilde{\tau}_\phi := \inf\{u > 0 \ : B_u \geq \sqrt{2}u + \phi(u)\}$$

and the process

$$V_t := \frac{R^\phi(B_t - \sqrt{2}t, t)}{R^\phi(0,0)}\mathbb{I}_{\{\tilde{\tau}_\phi > t\}}e^{\sqrt{2}B_t - t}.$$

The following result then holds.

Lemma 3.11. *Assuming that ϕ satisfies Assumption 3.8, $(V_t, t \geq 0)$ is a mean one $\mathbb{P}$-martingale. Defining $\mathbb{P}^V$ by $\frac{\mathrm{d}\mathbb{P}^V}{\mathrm{d}\mathbb{P}}\big|_{\mathcal{F}_t} := V_t$, $\mathbb{P}^V$ is a probability measure corresponding to the law of a Brownian motion with drift $\sqrt{2}$ conditioned to stay below $\sqrt{2}t + \phi(t)$ at all times $t \geq 0$ (in the sense of Proposition 3.10).*

Proof. Set $Y_t := e^{\sqrt{2}B_t - t}$. It is then well-known that Y is a $\mathbb{P}$-martingale and that the law

$\tilde{\mathbb{P}} = Y \cdot \mathbb{P}$ corresponds to the law of a Brownian motion with drift $\sqrt{2}$, by Girsanov's Theorem. Observe that

$$\left.\frac{\mathrm{d}\mathbb{P}^V}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{R^\phi(B_t - \sqrt{2}t, t)\mathbb{I}_{\{\tau_\phi > t\}}}{R^\phi(0,0)} e^{\sqrt{2}B_t - t} = \frac{R^\phi(B_t - \sqrt{2}t, t)\mathbb{I}_{\{\tilde{\tau}_\phi > t\}}}{R^\phi(0,0)} \left.\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t}.$$

Using that under $\tilde{\mathbb{P}}$, $(B_t - \sqrt{2}t, t \geq 0)$ is a Brownian motion, we obtain immediately from Lemma 3.9 that $(R^\phi(B_t - \sqrt{2}t, t)\mathbb{I}_{\{\tilde{\tau}_\phi > t\}}, t \geq 0)$ is a non-negative $\tilde{\mathbb{P}}$-martingale, and therefore that V is a $\mathbb{P}$-martingale.

Additionally, we have that

$$\left.\frac{\mathrm{d}\mathbb{P}^V}{\mathrm{d}\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \frac{R^\phi(B_t - \sqrt{2}t, t)\mathbb{I}_{\{\tau_\phi > t\}}}{R^\phi(0,0)},$$

hence by Proposition 3.10 we have that under $\mathbb{P}^V$, $(B_t - \sqrt{2}t, t \geq 0)$ is a Brownian motion conditioned on not hitting the curve ϕ, which completes the proof. □

3.3.2 Behaviour of the conditioned process

We describe here the behaviour of the one-dimensional Brownian motion B_t under the law $\mathbb{P}^\phi$. We prove that for the Brownian motion conditioned to stay below ϕ, the process localizes at time t at position $-t^{1/2+o(1)}$. In other words, for any $\epsilon \in (0, 1/2)$, for all t large enough one has $t^{1/2-\epsilon} < -B_t < t^{1/2+\epsilon}$ $\mathbb{P}^\phi$-a.s. This result is similar to what happens with the Bessel process, i.e. as the Brownian motion typically has $\sqrt{t}$ fluctuation, conditioning it to stay below 0 or a smooth enough function of order $o(t^{1/2-\epsilon})$ does not make a difference, asymptotically.

Lemma 3.12. *Let ϕ be a function satisfying Assumption 3.8. We have*

$$\lim_{t\to\infty} \frac{\log(-B_t)}{\log t} = \frac{1}{2} \quad \mathbb{P}^\phi - a.s.,$$

i.e. $B_t = -t^{1/2+o(1)}$ as $t \to \infty$, $\mathbb{P}^\phi$-a.s.

We split this lemma into several pieces. We begin with an upper bound for the probability for B to be close to $\phi(t)$ at time t under the law $\mathbb{P}^\phi$.

Lemma 3.13. *Let ϕ be a function satisfying Assumption 3.8. There exists $C > 0$ such that* *for all $t, x \geq 0$ we have*

$$\mathbb{P}^{\phi}(B_t \geq \phi(t) - x) \leq C \left(\frac{1+x}{(1+t)^{1/2}} \right)^3 .$$

Proof. Let $x \geq 0$ and $t \geq 1$. Using the definition of $\mathbb{P}^{\phi}$ we have

$$\begin{aligned}
\mathbb{P}^{\phi}(B_t \geq \phi(t) - x) &= \mathbb{E}_{\mathbb{P}} \left(R(B_t, t) \mathbb{I}_{\{B_t \geq \phi(t) - x, \tau_\phi > t\}} \right) \\
&\leq \sup_{z \in [0,x]} R(\phi(t) - z, t) \mathbb{P} \left(B_t \geq \phi(t) - x, \tau_\phi > t \right) \\
&\leq C(1+x) \mathbb{P}(B_t \geq \phi(t) - x, \tau_\phi > t),
\end{aligned}$$

by (3.9). By the Markov property at time $t/2$, we have

$$\begin{aligned}
&\mathbb{P}(B_t \geq \phi(t) - x, \tau_\phi > t) \\
&\leq \mathbb{P}(\tau_\phi > t/2) \sup_{z \in \mathbb{R}} \mathbb{P}_z \left(B_{t/2} \geq \phi(t) - x, B_s \leq \phi(t/2 + s), \forall s \leq t/2 \right) \\
&\leq C t^{-1/2} \sup_{z \in \mathbb{R}} \mathbb{P}_z \left(B_{t/2} \geq \phi(t) - x, B_s \leq \phi(t), s \leq t/2 \right),
\end{aligned}$$

using Theorem 3.6 and Assumption 3.8.

We now use time-reversal of the Brownian motion, observing that under $\mathbb{P}_z$, $\hat{B}_s := B_{t/2} - B_{t/2-s}$ is a Brownian motion started from 0. We use it to estimate

$$\begin{aligned}
&\sup_{z \in \mathbb{R}} \mathbb{P}_z(B_{t/2} \geq \phi(t) - x, B_s \leq \phi(t), s \leq t/2) \\
= &\sup_{z \in \mathbb{R}} \mathbb{P}_z(\hat{B}_{t/2} + z \geq \phi(t) - x, \hat{B}_{t/2} + z - \hat{B}_s \leq \phi(t), s \leq t/2) \\
\leq &\sup_{z \in \mathbb{R}} \mathbb{P}_z(\hat{B}_{t/2} \geq \phi(t) - z - x, \hat{B}_{t/2} + z \leq \phi(t), \hat{B}_s \geq -x, s \leq t/2) \\
= &\sup_{z' \in \mathbb{R}} \mathbb{P}(\hat{B}_{t/2} \in [z', z' + x], \hat{B}_s \geq -x, s \leq t/2) \\
\leq &\mathbb{P}(\hat{B}_s \geq -x, s \leq t/4) \sup_{z \in \mathbb{R}} \mathbb{P}(\hat{B}_{t/4} \in [z, z + x]),
\end{aligned}$$

using the Markov property at time $t/4$. Then, using again Theorem 3.6, there exists $C > 0$ such that for all $x \geq 0$ and $t \geq 1$,

$$\mathbb{P}(\hat{B}_s \geq -x, s \leq t/4) \leq C(1+x)/t^{1/2}.$$

Additionally, we have $\mathbb{P}(\hat{B}_{t/4} \in [z, z+x]) \le \sqrt{\frac{2}{\pi t}} x$ for all $z \in \mathbb{R}$, noting that the density of $\hat{B}_{t/4}$ is bounded by $\sqrt{\frac{2}{\pi t}}$. Finally, we obtain the existence of $C > 0$ such that for all $t, x \ge 0$

$$\mathbb{P}^{\phi}(B_t \ge \phi(t) - x) \le C \frac{(1+x)^3}{(1+t)^{3/2}}. \qquad \square$$

We now use this result to bound from below the asymptotic behaviour of $\frac{\log(-B_t)}{\log t}$.

Lemma 3.14. *Given ϕ a function satisfying Assumption 3.8, we have*

$$\liminf_{t\to\infty} \frac{\log(-B_t)}{\log t} \ge \frac{1}{2} \quad \mathbb{P}^{\phi} - a.s.$$

Proof. To prove this result, we begin by using the Borel-Cantelli lemma to show that $\mathbb{P}^{\phi}$ almost surely, for all $\gamma < 1/2$,

$$\liminf_{n\to\infty} \frac{\log(-B_{t_n})}{\log t_n} \ge \gamma \tag{3.12}$$

along a well-chosen sequence t_n growing to ∞. We then use the observation that with high probability the Brownian motion between times t_n and t_{n+1} stays within a distance $O(t_{n+1} - t_n)^{1/2}$ from B_{t_n}. Therefore, as long as $(t_{n+1} - t_n)^{1/2}/t_n^{\gamma} \to 0$, we can extend (3.12) to any sequence growing to ∞, which completes the proof.

Let $\gamma < 1/2$. We assume without loss of generality that γ is close enough to $1/2$, such that $\phi(t) = o(t^{\gamma})$. Using Lemma 3.13 we have

$$\mathbb{P}^{\phi}(B_t \ge -t^{\gamma}) \le C t^{3\left(\gamma - \frac{1}{2}\right)}.$$

As a result, setting $t_n = n^{\frac{5}{6(1-2\gamma)}}$, we have

$$\mathbb{P}^{\phi}\left(\frac{\log(-B_{t_n})}{\log t_n} \le \gamma\right) \le C n^{-5/4}, \tag{3.13}$$

hence, by the Borel-Cantelli lemma,

$$\liminf_{n\to\infty} \frac{\log(-B_{t_n})}{\log t_n} \ge \gamma \quad \text{a.s.}$$

To complete the proof we now need to bound the maximal displacement of the Brownian motion in the time intervals $[t_n, t_{n+1}]$. Write $\alpha = \frac{5}{6(1-2\gamma)}$ so that $t_n = n^{\alpha}$ and compute for

$n \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{P}^{\phi}\left(\sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2, B_{t_n} \leq -t_n^{\gamma}\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(R^{\phi}(B_{t_{n+1}}, t_{n+1})\mathbb{I}_{\{\tau_\phi > t_{n+1}, B_{t_n} \leq -t_n^{\gamma}\}}\mathbb{I}_{\{\sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2\}}\right).
\end{aligned}$$

We can decompose this quantity depending on whether $B_{t_{n+1}}$ is smaller or larger than $-t_{n+1}^{2/3}$. Observe that for all $t \geq 1$ we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left(R^{\phi}(B_t, t)\mathbb{I}_{\{B_t < -t^{2/3}\}}\right) &\leq C\mathbb{E}_{\mathbb{P}}\left((1 + |B_t| + |\phi(t)|)\,\mathbb{I}_{\{B_t < -t^{2/3}\}}\right) \\
&\leq C\mathbb{E}_{\mathbb{P}}\left(|B_t|\mathbb{I}_{\{B_t < -t^{2/3}\}}\right) \leq Ce^{-\frac{t^{4/3}}{2t}},
\end{aligned}$$

using that $|\phi(t)| = o(t^{2/3})$ as $t \to \infty$ and integrating with respect to the Brownian density. Thus, there exists $C > 0$ such that for all $n \in \mathbb{N}$

$$\mathbb{E}_{\mathbb{P}}\left(R^{\phi}(B_{t_{n+1}}, t_{n+1})\mathbb{I}_{\{B_{t_{n+1}} < -t_{n+1}^{2/3}\}}\right) \leq C\exp\left(-t_{n+1}^{1/3}/2\right).$$

Hence, using that there exists $C > 0$ such that $R^{\phi}(x, t_{n+1}) \leq Ct_{n+1}^{2/3}$ for all $x \geq -t_{n+1}^{2/3}$,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left(R^{\phi}(B_{t_{n+1}}, t_{n+1})\mathbb{I}_{\{\tau_\phi > t_{n+1}, B_{t_n} \leq -t_n^{\gamma}\}}\mathbb{I}_{\{\sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2\}}\right) \\
&\qquad \leq Ct_{n+1}^{2/3}\mathbb{P}\left(\tau_\phi > t_{n+1}, \sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2, B_{t_n} \leq -t_n^{\gamma}\right) \\
&\qquad\qquad + C\exp\left(-t_{n+1}^{1/3}/2\right). \qquad (3.14)
\end{aligned}$$

We now bound $\mathbb{P}\left(\tau_\phi > t_{n+1}, \sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2, B_{t_n} \leq -t_n^{\gamma}\right)$. Using the Markov property at time t_n we have

$$\mathbb{P}\left(\tau_\phi > t_{n+1}, \sup_{s\in[t_n,t_{n+1}]} B_s \geq -t_n^{\gamma}/2, B_{t_n} \leq -t_n^{\gamma}\right) \leq \mathbb{E}_{\mathbb{P}}\left(G_n(B_{t_n})\mathbb{I}_{\{\tau_\phi > t_n\}}\mathbb{I}_{\{B_{t_n} \leq -t_n^{\gamma}\}}\right),$$

where $G_n(x) := \mathbb{P}_x\left(\mathbb{I}_{\{\sup_{s\leq t_{n+1}-t_n} B_s \geq -t_n^{\gamma}/2\}}\right)$. As $G_n(x)$ is non-decreasing in x, using the Brownian scaling, for all $x \leq -t_n^{\gamma}$ we have

$$G_n(x) \leq G_n(-t_n^{\gamma}) = \mathbb{P}_{-1}\left(\sup_{s\leq(t_{n+1}-t_n)/t_n^{2\gamma}} B_s \geq -1/2\right).$$

By definition of α and t_n we note that

$$\frac{t_{n+1}-t_n}{t_n^{2\gamma}} = \frac{(n+1)^\alpha - n^\alpha}{n^{2\alpha\gamma}} \sim \alpha n^{\alpha-1-2\gamma\alpha} = \alpha n^{-1/6} \text{ as } n \to \infty.$$

As the maximum of a Brownian motion on $[0,s]$ is distributed as the absolute value of a Gaussian random variable with parameter s, and using standard Gaussian estimates, we have

$$\begin{aligned}\mathbb{P}_{-1}\left(\sup_{s\le(t_{n+1}-t_n)/t_n^{2\gamma}} B_s \ge -1/2\right) &\le \mathbb{P}\left(\sup_{s\le C\alpha n^{-1/6}} B_s \ge 1/2\right)\\ &\le \frac{1}{\sqrt{\pi C\alpha n^{-1/6}}}\exp\left(-\frac{1}{8C\alpha n^{-1/6}}\right).\end{aligned}$$

Thus we deduce that for all $x \le -t_n^\gamma$ we have $G_n(x) \le Ce^{-cn^{1/6}}$. Since t_n has polynomial growth, we therefore obtain from (3.14) that there exists $C, \delta > 0$ such that

$$\mathbb{E}_\mathbb{P}\left(R^\phi(B_{t_{n+1}}, t_{n+1})\mathbb{I}_{\{\tau_\phi > t_{n+1}, B_{t_n}\le -t_n^\gamma\}}\mathbb{I}_{\{\sup_{s\in[t_n,t_{n+1}]} B_s \ge -t_n^\gamma/2\}}\right) \le Ce^{-n^\delta}.$$

We now conclude, using (3.13), that

$$\begin{aligned}&\sum_{n\in\mathbb{N}} \mathbb{P}^\phi(\sup_{s\in[t_n,t_{n+1}]} B_s \ge -t_n^\gamma/2)\\ \le& \sum_{n\in\mathbb{N}} \mathbb{P}^\phi(B_{t_n} \ge -t_n^\gamma) + \sum_{n\in\mathbb{N}} \mathbb{P}^\phi\left(B_{t_n} \le -t_n^\gamma, \sup_{s\in[t_n,t_{n+1}]} B_s \ge -t_n^\gamma/2\right)\\ \le& C\sum_{n\in\mathbb{N}} n^{-5/4} + \sum_{n\in\mathbb{N}} e^{-n^\delta} < \infty,\end{aligned}$$

which completes the proof, by the Borel-Cantelli lemma. □

A similar simpler proof also gives an upper bound for $\log(-B_t)/\log t$ under the law $\mathbb{P}^\phi$.

Lemma 3.15. *Given ϕ a function satisfying Assumption 3.8, we have*

$$\limsup_{t\to\infty} \frac{\log(-B_t)}{\log t} \le \frac{1}{2} \quad \mathbb{P}^\phi - a.s.$$

Proof. Let $\alpha > 1/2$. We observe that for all $n \in \mathbb{N}$ we have

$$\mathbb{P}^{\phi}(\inf_{s\in[n,n+1]} B_s \leq -n^{\alpha}) \leq \frac{1}{R^{\phi}(0,0)} \mathbb{E}_{\mathbb{P}}(R^{\phi}(B_{n+1}, n+1)\mathbb{I}_{\{\inf_{s\in[n,n+1]} B_s \leq -n^{\alpha}\}})$$
$$\leq Ce^{-cn^{2\alpha-1}},$$

using that $R^{\phi}(x, n+1)$ grows at most linearly in $-x$, and the Gaussian concentration of $\inf_{s\in[n,n+1]} B_s$. As a result, by the Borel-Cantelli lemma we conclude that

$$\limsup_{t\to\infty} \frac{\log(-B_t)}{\log t} \leq \alpha \quad \text{a.s.}$$

We complete the proof by letting $\alpha \to 1/2$. □

The proof of Lemma 3.12 is then a combination of Lemmas 3.14 and 3.15.

3.3.3 Linear growth

In this section we prove the key property of R^{ϕ}: the function grows linearly in $-x$ uniformly in t. We begin with the following lower bound on R^{ϕ}, which is a straightforward consequence of the definition in Theorem 3.7.

Lemma 3.16. *Let ϕ be a function satisfying Assumption 3.8, then for all $t \geq 0$ and $x \leq \phi(t)$,*

$$R^{\phi}(x,t) \geq \phi(t) - x.$$

Proof. Recall that for all $s \geq 0$ we have $\phi_t(s) = \phi(t+s) \geq \phi(t)$, as ϕ is increasing, and also that $\tau_{(\phi_t - x)} = \inf\{u > 0 \ : B_u \geq \phi(t+u) - x\}$. Therefore, by Theorem 3.7 we obtain that for $x \leq \phi(t)$ we have $\tau_{(\phi_t - x)} < \infty$ a.s. and

$$R^{\phi}(x,t) = \mathbb{E}_{\mathbb{P}} B_{\tau_{(\phi_t - x)}} \geq \phi(t) - x,$$

completing the proof. □

To obtain a uniform upper bound on R^{ϕ}, we need to add an assumption on the growth rate of the derivative of ϕ.

Lemma 3.17. *Let ϕ be a function satisfying Assumption 3.8, and assume additionally that $\phi'(t) = o(t^{-1/2-\epsilon})$ for some $\epsilon > 0$. Then for all $\delta > 0$ and $D > 0$ there exists $t_0 > 0$ such* *that*

$$\forall t \geq t_0,\ \forall x \in [\phi(t) - Dt, \phi(t) - Dt_0],\ R^{\phi}(x,t) \leq (\phi(t) - x)(1+\delta). \tag{3.15}$$

Proof. Observe that by the assumption on the function ϕ, there exists $\gamma < 1/2$ and $\alpha > 0$ such that for all $t \geq 0$ we have $0 \leq \phi'(t) \leq \alpha\gamma t^{\gamma-1}$. By integration we immediately obtain that for all $s, t \geq 0$

$$\phi(t+s) - \phi(t) \leq \psi(t+s) - \psi(t),$$

where we have set $\psi(t) = \alpha t^{\gamma}$. It is then straightforward to note that for all $s, t \geq 0$ and $x \leq \phi(t)$

$$\mathbb{P}_x\left(B_u \leq \phi(t+u), u \leq s\right) \leq \mathbb{P}_x\left(B_u \leq \psi(t+u) - \psi(t) + \phi(t), u \leq s\right).$$

As a result, by Theorem 3.7 and using that $R^{\phi}(x,t) = 0$ for $x \geq \phi(t)$, we obtain that

$$R^{\phi}(x,t) \leq R^{\psi}(x + \psi(t) - \phi(t), t) \tag{3.16}$$

for all $x \in \mathbb{R}$ and $t \geq 0$. Therefore, we shall work with R^{ψ} which will simplify some arguments, and use (3.16) to prove (3.15).

For $t, x \geq 0$ set

$$S^{\psi}(x,t) := R^{\psi}(\psi(t) - x, t) - x = \mathbb{E}_{-x}(B_{\tau_{\psi_t - \psi(t)}}).$$

Observe that as ψ is concave, for all $s \geq 0$ we have that $\psi_t(s) - \psi(t)$ is decreasing with t. Therefore $t \mapsto S^{\psi}(x,t)$ is decreasing, hence for all $D > 0$ one has $S^{\psi}(x,t) \leq S^{\psi}(x, x/D)$ as long as $x \leq Dt$. We shall show that for any $D > 0$ we have $S^{\psi}(Dt,t)/t \to 0$ as $t \to \infty$.

Fix $D > 0$. For all $\lambda, t > 0$ define

$$\psi^{\lambda}(t) := \frac{1}{\lambda}(\psi(\lambda + \lambda^2 t) - \psi(\lambda)),$$

and observe that by the scaling property of the Brownian motion we have

$$\frac{S^{\psi}(\lambda D, \lambda)}{\lambda} = \frac{1}{\lambda}\mathbb{E}_{-\lambda D}(B_{\tau_{\psi_\lambda - \psi(\lambda)}}) = \mathbb{E}_{-D}(B_{\tau_{\psi^\lambda}}). \tag{3.17}$$

Observe that $(\psi^\lambda, \lambda > 1)$ decreases to 0 as $\lambda \to \infty$. We can also note that the convergence is monotone outside of a compact set. Indeed, for all $u > 0$,

$$\begin{aligned}
\frac{1}{\alpha}\frac{\mathrm{d}\psi^\lambda(u)}{\mathrm{d}\lambda} &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{1}{\lambda}\left((\lambda+u\lambda^2)^\gamma - \lambda^\gamma\right)\right) \\
&= \frac{1}{\lambda^2}\Bigg(\gamma\lambda\left((1+2u\lambda)(\lambda+u\lambda^2)^{\gamma-1} - \lambda^{\gamma-1}\right) \\
&\qquad\qquad - \left((\lambda+u\lambda^2)^\gamma - \lambda^\gamma\right)\Bigg) \\
&= \frac{1}{\lambda^2}\left((1-\gamma)\lambda^\gamma - (\lambda+u\lambda^2)^{\gamma-1}\left(\lambda(1-\gamma)+u\lambda^2(1-2\gamma)\right)\right) \\
&= \frac{1}{\lambda^2}\left((1-\gamma)\lambda^\gamma - (1-2\gamma)(\lambda+u\lambda^2)^\gamma - \gamma\lambda(\lambda+u\lambda^2)^{\gamma-1}\right).
\end{aligned}$$

In particular, it appears there exists $\lambda_0 > 0$ such that for all $u > 1$ and $\lambda > \lambda_0$ we have that $\frac{\mathrm{d}\psi^\lambda(u)}{\mathrm{d}\lambda} < 0$. Therefore, setting $\bar{\psi}^\lambda(u) := \psi^\lambda(u \vee 1)$, we have

$$0 \leq \mathbb{E}_{-D}(B_{\tau_{\psi^\lambda}}) \leq \mathbb{E}_{-D}(B_{\tau_{\bar{\psi}^\lambda}}) \to 0 \quad \text{ as } \lambda \to \infty,$$

by the monotone convergence theorem, using that $\bar{\psi}^\lambda$ decreases to 0 when $\lambda \to \infty$. Therefore, (3.17) yields

$$\lim_{t\to\infty} S^\psi(Dt,t)/t = 0.$$

Choose $\delta > 0$. There exists $t_0 > 0$ such that for all $t > t_0$ we have $S^\psi(Dt,t) \leq \delta Dt$. Then, recalling (3.16), for all $Dt_0 \leq y \leq Dt$ we have

$$R^\phi(\phi(t)-y,t) \leq R^\psi(\psi(t)-y,t) = y + S^\psi(y,t) \leq y + S^\psi(y, y/D) \leq (1+\delta)y,$$

which, setting $x := \phi(t) - y$, completes the proof. □

3.4 Multidimensional Branching Brownian Motion and uniformly integrable approximations of the martingale

In this section we prove Theorem 3.3, showing that the derivative martingale almost surely converges in almost every direction simultaneously. As we mentioned in the introduction, the techniques are based on a shaving argument: removing all particles that travel too far away from the origin, and therefore carry most of the fluctuations of Z. It turns the derivative martingale into a uniformly integrable martingale. We use here the results obtained in the previous section to construct a shaving argument with a function satisfying Assumption 3.8.

Before moving to the multidimensional setting, we are going to define the martingale Z^ϕ in dimension 1, that will serve as a uniformly integrable approximation of the derivative martingale Z. To be precise, set

$$\mathcal{N}_t^\phi := \{j \in \mathcal{N}_t : X_s(j) \leq \sqrt{2}s + \phi(s), s \leq t\}.$$

The martingale Z^ϕ is then defined in the following way.

Proposition 3.18. *Let ϕ be a function satisfying Assumption 3.8. We set R^ϕ as in* (3.8). *Then the process defined for all $t \geq 0$ by*

$$Z_t^\phi := \sum_{j \in \mathcal{N}_t^\phi} R^\phi(X_t(j) - \sqrt{2}t, t) e^{\sqrt{2}(X_t(j) - \sqrt{2}t)}$$

is a non-negative $\mathbb{P}$-martingale with mean $R^\phi(0,0)$.

Proof. We first note that by definition, $\mathbb{E}_{\mathbb{P}} Z_0^\phi = R^\phi(0,0)$, and that for all t, x, we have $R^\phi(x,t) \geq 0$. We thus only need to check that Z_t^ϕ is a $\mathbb{P}$-martingale. By the branching property, for all $s, t \geq 0$ we have

$$\mathbb{E}_{\mathbb{P}}(Z_{t+s}^\phi \mid \mathcal{F}_t) = \sum_{j \in \mathcal{N}_t^\phi} G_s(X_t(j)),$$

where we have set

$$\begin{aligned}
&G_s(x) \\
&:= \mathbb{E}_{\mathbb{P}}\left(\sum_{j \in \mathcal{N}_s^{\sqrt{2}t + \phi_t - x}} R^\phi(X_s(j) + x - \sqrt{2}(t+s), t+s) e^{\sqrt{2}(X_s(j) + x - \sqrt{2}(t+s))} \right) \\
&= e^{\sqrt{2}(x - \sqrt{2}t)} \mathbb{E}_{\mathbb{P}}\Bigg(\sum_{j \in \mathcal{N}_s^{\sqrt{2}t + \phi_t - x}} R^\phi(X_s(j) + x - \sqrt{2}(t+s), t+s) \\
&\qquad\qquad \cdot e^{\sqrt{2}(X_s(j) - \sqrt{2}s)} \Bigg) \\
&= e^{\sqrt{2}(x - \sqrt{2}t)} e^s \mathbb{E}_{\mathbb{P}}\Big(R^\phi(B_s + x - \sqrt{2}(t+s), t+s) \\
&\qquad\qquad \cdot e^{\sqrt{2}B_s - 2s} \mathbb{I}_{\{\forall u \leq s, B_u + x - \sqrt{2}t \leq \sqrt{2}u + \phi_t(u)\}} \Big),
\end{aligned}$$

by Lemma 3.5 (many-to-one). Thus by Lemma 3.11 we obtain

$$G_s(x) = e^{\sqrt{2}(x-\sqrt{2}t)} R^{\phi}(x - \sqrt{2}t, t),$$

from which we deduce that $\mathbb{E}_{\mathbb{P}}(Z^{\phi}_{t+s} \mid \mathcal{F}_t) = Z^{\phi}_t$ a.s., completing the proof. □

3.4.1 Construction of $(Z^{\phi}_t(\theta), t \geq 0)$: radial shaving

We may now turn to our main object of interest : the d-dimensional branching Brownian motion $X_t = (X_t(i), i \in \mathcal{N}_t)$. Recall that this is a d-dimensional branching particle system in which particles move according to i.i.d. Brownian motions and split into two at rate one. For a direction $\theta \in \mathbb{S}^{d-1}$ recall that we denoted

$$Z_t(\theta) = \sum_{j \in \mathcal{N}_t} (\sqrt{2}t - X_t(j) \cdot \theta) e^{\sqrt{2}(X_t(j)\cdot\theta - \sqrt{2}t)}.$$

We now introduce the shaved martingale Z^{ϕ}, where the shaving is done along a curve ϕ satisfying Assumption 3.8. Set $\mathcal{N}^{\phi,\theta}_t = \{j \in \mathcal{N}_t : X_s(j) \cdot \theta \leq \sqrt{2}s + \phi(s), s \leq t\}$ for $t \geq 0$ and $\phi \in \mathbb{S}^{d-1}$.

We now set

$$Z^{\phi}_t(\theta) := \sum_{j \in \mathcal{N}^{\phi,\theta}_t} R^{\phi}(X_t(j) \cdot \theta - \sqrt{2}t, t) e^{\sqrt{2}(X_t(j)\cdot\theta - \sqrt{2}t)}. \tag{3.18}$$

The function ϕ will be chosen to grow fast enough to guarantee that

$$\lim_{A\to\infty} \mathbb{P}\left(\forall t \geq 0, \bigcap_{\theta \in \mathbb{S}^{d-1}} \mathcal{N}^{\phi \vee A,\theta}_t = \mathcal{N}_t\right) = 1.$$

We will show in Section 3.4.2 that choosing ϕ growing faster than $\frac{d-1}{2\sqrt{2}} \log t$ as $t \to \infty$ is enough.

In Section 3.4.3 we prove (using classical spinal decomposition techniques along the lines of [43] and [59]) that for all measurable bounded functions f the process $(\langle Z^{\phi}_t, f\rangle, t \geq 0)$ is a uniformly integrable $\mathbb{P}$-martingale. We then use convergence of these martingales in Section 3.4.4 to show that $\lim_{t\to\infty} Z^{\phi}_t(\theta)$ exists for almost all $\theta \in \mathbb{S}^{d-1}$ almost surely. Finally, we complete the proof of Theorem 3.3 by using that, with high probability, Z and Z^{ϕ} coincide asymptotically as $t \to \infty$, for which we shall apply Lemma 3.17.

3.4.2 Bounds on the maximal displacement of the BBM

We prove here that with high probability all particles in the multidimensional BBM are at all times t within a ball of radius $\sqrt{2}t + \frac{d-1}{2\sqrt{2}}\log(t+1) + A$. First, recall the following lemma due to Mallein:

Lemma 3.19 ([50], Lemma 3.1)**.** *Let*

$$r_s^{t,y} := \sqrt{2}s + \frac{d-1}{2\sqrt{2}}\log(s+y) - \frac{3}{2\sqrt{2}}\log\frac{t+1}{t-s+1} + y.$$

Then there exists $C > 0$ such that for any $t \geq 1$ and $y \in [1, \sqrt{t}]$

$$\mathbb{P}\left(\exists j \in \mathcal{N}_t, \exists s \leq t : ||X_s(j)|| \geq r_s^{t,y}\right) \leq Cye^{-\sqrt{2}y}.$$

We use Lemma 3.19 to prove the following result.

Lemma 3.20. *Let $\tilde{r}(s) := \sqrt{2}s + \frac{d-1}{2\sqrt{2}}\log(1+s)$. For any $\epsilon > 0$ there exists C_ϵ such that*

$$\mathbb{P}\left(\exists t \geq 0, \exists j \in \mathcal{N}_t : ||X_t(j)|| \geq \tilde{r}(t) + C_\epsilon\right) \leq \epsilon.$$

Proof. Observe first that by Lemma 3.19, for any $y > 0$ and $t \geq 0$, we have

$$\mathbb{P}\left(\exists s \leq t, \exists j \in \mathcal{N}_s : ||X_s(j)|| \geq \sqrt{2}s + \frac{d-1}{2\sqrt{2}}\log(s+y) + y\right)$$
$$\leq \mathbb{P}\left(\exists s \leq t, \exists j \in \mathcal{N}_s : ||X_s(j)|| \geq r_s^{t,y}\right) \leq Cye^{-\sqrt{2}y}.$$

Hence, choosing y large enough such that $Cye^{-\sqrt{2}y} < \epsilon$ and letting $t \to \infty$, we deduce that

$$\mathbb{P}(\exists s \geq 0, j \in \mathcal{N}_s : ||X_s(j)|| \geq \sqrt{2}s + \frac{d-1}{2\sqrt{2}}\log(s+y) + y) \leq \epsilon.$$

To complete the proof, it is therefore enough to choose C_ϵ as

$$\sup_{t\geq 0} \frac{d-1}{2\sqrt{2}}\log(t+y) + y - \frac{d-1}{2\sqrt{2}}\log(t+1) < \infty. \qquad \square$$

3.4.3 Uniform integrability of $(Z_t^\phi, t \geq 0)$

Let f be a non-negative function such that $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$. By Fubini's theorem, it is a straightforward calculation to verify that the process defined by

$$\langle Z_t^\phi, f\rangle = \int_{\mathbb{S}^{d-1}} Z_t^\phi(\theta) f(\theta)\sigma(\mathrm{d}\theta)$$

is a non-negative $\mathbb{P}$-martingale. To prove its uniform integrability we use a spinal decomposition method. This technique, appearing earlier e.g. in work of Evans [31], but popularised by Lyons, Pemantle and Peres [46] for studying Galton-Watson processes, and adapted by Lyons [45] to spatial branching settings, consists in an alternative description of the law of the branching Brownian motion biased by the martingale $\langle Z_t^\phi, f\rangle$. More precisely, we define

$$\left.\frac{\mathrm{d}\mathbb{P}^f}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} := R^\phi(0,0)^{-1}\langle Z_t^\phi, f\rangle.$$

The spinal decomposition consists in a construction of the BBM under the law $\mathbb{P}^f$, where a distinguished particle, called the spine, moves and reproduces differently to typical BBM particles. The offspring of that spine particle then start independent copies of the original BBM with law $\mathbb{P}$, from their birth time and position.

Thanks to the spinal decomposition we will be able to show that under the measure $\mathbb{P}^f$, $\limsup_{t\to\infty}\langle Z_t^\phi, f\rangle$ is almost surely finite, which will allow us to deduce that the $\mathbb{P}$-martingale $\langle Z_t^\phi, f\rangle$ is uniformly integrable.

Before presenting the spinal decomposition for the branching Brownian motion, we introduce the law of the multi-dimensional Brownian motion biased by a martingale similar to the one introduced in Lemma 3.11. This will allow us to describe the trajectory of the spine under the biased law $\mathbb{P}^f$.

Let B be a Brownian motion in $\mathbb{R}^d$. For all $\theta \in \mathbb{S}^{d-1}$ we define a non-negative $\mathbb{P}$-martingale $(V_t(\theta), t \geq 0)$ as

$$V_t(\theta) := \frac{R^\phi(B_t\cdot\theta - \sqrt{2}t, t)}{R^\phi(0,0)}\mathbb{I}_{\{\tau_\phi(\theta)>t\}}e^{\sqrt{2}B_t\cdot\theta - t},$$

where $\tau_\phi(\theta) := \inf\{u > 0 \ : B_u\cdot\theta - \sqrt{2}u \geq \phi(u)\}$. Writing $B^{(1)} = B\cdot\theta$ and $B^{(2)}$ for the projection of B on $\theta^\perp$, we note that these are two independent Brownian motions. Applying

Lemma 3.11 to $B^{(1)}$, we deduce that under the law defined as

$$\left.\frac{\mathrm{d}\mathbb{P}^{V(\theta)}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} := V_t(\theta)$$

the process B is a d-dimensional Brownian motion with drift $\sqrt{2}\theta$, conditioned on $B_t \cdot \theta \leq \sqrt{2}t + \phi(t)$ for all $t \geq 0$ (in the sense of Doob).

The key point of Theorem 3.3 is to consider several directions at the same time. To do so, we will consider integrated versions of the martingale $V(\theta)$. Given f a non-negative function satisfying $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$, we set

$$U_t := \langle V_t, f \rangle$$

and we define the measure $\mathbb{P}^U$ by

$$\left.\frac{\mathrm{d}\mathbb{P}^{U}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} := U_t.$$

Lemma 3.21. *Let f be a non-negative function with $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$, then the process U is a non-negative $\mathbb{P}$-martingale. Moreover, setting θ_0 a random variable in $\mathbb{S}^{d-1}$ with law $f(\theta)\sigma(\mathrm{d}\theta)$ and writing (B_t) for a process with law $\mathbb{P}^{V(\theta_0)}$ conditionally on θ_0, the process $(B_t, t \geq 0)$ has law $\mathbb{P}^U$.*

Proof. The process U is a $\mathbb{P}$-martingale using Fubini's theorem. Additionally, for all $t \geq 0$ and $G \in \mathcal{G}_t$ we have

$$\mathbb{P}^U(G) = \int_G \langle V_t, f \rangle \mathrm{d}\mathbb{P} = \left\langle \int_G V_t \mathrm{d}\mathbb{P}, f \right\rangle = \int_{\mathbb{S}^{d-1}} \mathbb{P}^{V(\theta)}(G) f(\theta)\sigma(\mathrm{d}\theta),$$

which justifies the description of B under the law $\mathbb{P}^U$. □

Observe that one can decompose

$$\langle Z_t^\phi, f \rangle = R^\phi(0,0) \sum_{j \in \mathcal{N}_t} U_t(j) e^{-t},$$

where

$$U_t(j) := \left\langle \frac{R^\phi(X_t(j) \cdot \theta - \sqrt{2}t, t)}{R^\phi(0,0)} \mathbb{I}_{\{\forall u < t, X_u(j) \cdot \theta - \sqrt{2}t < \phi(u)\}} e^{\sqrt{2}X_t(j) \cdot \theta - t}, f \right\rangle.$$

Thanks to this decomposition we can describe the BBM under the law $\mathbb{P}^f$ in terms of a spinal decomposition, which follows e.g. from [34, Lemma 6.7].

Lemma 3.22. *Let f be a non-negative function with $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$. The law of the BBM under $\mathbb{P}^f$ can be constructed as follows*

1. *we pick a direction θ_0 according to a random variable on $\mathbb{S}^{d-1}$ with density $f(\theta)\sigma(\mathrm{d}\theta)$;*
2. *conditionally on this direction we sample a trajectory (Ξ_t) with law $\mathbb{P}^{V(\theta_0)}$ that will be followed by the spine particle;*
3. *the spine particle creates offspring at rate 2;*
4. *every child of the spine then starts an independent standard BBM with law $\mathbb{P}$.*

An analogous decomposition in dimension one was given in [24] or in [43]. We are now ready to present the key lemma that states the uniform integrability of Z_t^{ϕ}.

Lemma 3.23. *Let ϕ be a function satisfying Assumption 3.8. For any bounded measurable function f the $\mathbb{P}$-martingale $\left(\langle Z_t^{\phi}, f\rangle\right)_{t\geq 0}$ is uniformly integrable.*

Before we present the proof of Lemma 3.23, note that applying it in dimension one with the binary function f (i.e. $f(-1) = 0$ and $f(1) = 1$) we obtain the following corollary.

Corollary 3.24. *For any $\theta \in \mathbb{S}^{d-1}$, $\left(Z_t^{\phi}(\theta)\right)_{t\geq 0}$ is a uniformly integrable $\mathbb{P}$-martingale.*

Proof of Lemma 3.23. Note first that without loss of generality we may assume that $f \geq 0$ and that $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$, as otherwise we may write f as a linear combination of functions satisfying these assumptions and consider each of these functions separately.

Set $\mathcal{Z} := \limsup_{t\to\infty}\langle Z_t^{\phi}, f\rangle$ (which is also equal to $\lim_{t\to\infty}\langle Z_t^{\phi}, f\rangle$ $\mathbb{P}$-a.s. because $\langle Z_t^{\phi}, f\rangle$ is a non-negative $\mathbb{P}$-martingale). Recall the following measure theoretic dichotomy (see e.g. Theorem 5.3.3. in [30]):

Theorem 3.25. *Let $(\mathcal{F}_n)$ be a filtration, and let $\mathcal{F}_\infty$ be the smallest σ-field containing all $\mathcal{F}_n$. Let $\mathbb{P}, \mathbb{Q}$ be two probability measures on $(\Omega, \mathcal{F}_\infty)$. Assume that for any n, $\mathbb{Q}_{|\mathcal{F}_n} \ll \mathbb{P}_{|\mathcal{F}_n}$ and let $X_n := \frac{\mathrm{d}\mathbb{Q}_{|\mathcal{F}_n}}{\mathrm{d}\mathbb{P}_{|\mathcal{F}_n}}$ and $X := \limsup_{n\to\infty} X_n$ which is $\mathbb{P}$-a.s. finite. Then*

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(X\mathbb{I}_{\{A\}}) + \mathbb{Q}(A \cap \{X = \infty\}), \quad \forall A \in \mathcal{F}_\infty.$$

From Theorem 3.25 we obtain that

$$\mathbb{P}^f\left(\frac{\mathcal{Z}}{R^\phi(0,0)} < \infty\right) = 1 \qquad \Longleftrightarrow \qquad \int \frac{\mathcal{Z}}{R^\phi(0,0)} \mathrm{d}\mathbb{P} = 1,$$

thus instead of proving that $\mathbb{E}_{\mathbb{P}}\mathcal{Z} = 1$, we shall prove that under $\mathbb{P}^f$, $\mathcal{Z}$ is almost surely finite. To show that, we are going to use the spinal decomposition from Lemma 3.22.

Let $\mathcal{F}_\infty$ be the filtration generated by the movement and the branching of the spine Ξ, and $\mathcal{B}_t$ be the set of branching times of the spine until time t. For the ease of notation we shall write $\mathbb{P}^f$ for the expectation under the measure $\mathbb{P}^f$. From the decomposition mentioned above and the martingale property from Proposition 3.18 we see that

$$\begin{aligned}\mathbb{P}^f[\langle Z_t^\phi, f\rangle|\mathcal{F}_\infty] = &\left\langle \sum_{s\in\mathcal{B}_t} R^\phi(\Xi_s\cdot\theta - \sqrt{2}s, s)e^{\sqrt{2}(\Xi_s\cdot\theta-\sqrt{2}s)}, f\right\rangle \\ &+ \left\langle R^\phi(\Xi_t\cdot\theta - \sqrt{2}t, t)e^{\sqrt{2}(\Xi_t\cdot\theta-\sqrt{2}t)}, f\right\rangle.\end{aligned}$$

To complete the proof it is enough to show that

$$\limsup_{t\to\infty} \mathbb{P}^f[\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] < \infty, \tag{3.19}$$

as by Fatou's lemma we have

$$\begin{aligned}&\mathbb{P}^f[\liminf_{t\to\infty}\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] \\ \leq\ &\liminf_{t\to\infty} \mathbb{P}^f[\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] \\ \leq\ &\limsup_{t\to\infty} \mathbb{P}^f[\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] < \infty,\end{aligned}$$

which implies that

$$\liminf_{s\to\infty}\langle Z_t^\phi, f\rangle < \infty, \quad \mathbb{P}^f - \text{a.s.} \tag{3.20}$$

Recalling the definition of $\mathbb{P}^f$, $(\langle Z_t^\phi, f\rangle)^{-1}$ is a non-negative $\mathbb{P}^f$-supermartingale, hence it converges to a finite limit $\mathbb{P}^f$ almost surely. Combined with (3.20) this implies that $\mathbb{P}^f$ almost surely

$$\liminf_{s\to\infty}\langle Z_t^\phi, f\rangle = \limsup_{s\to\infty}\langle Z_t^\phi, f\rangle,$$

from which we would deduce that $\lim_{t\to\infty}\langle Z_t^\phi, f\rangle < \infty$ $\mathbb{P}^f$-a.s.

It remains to show (3.19). We first upper bound $||\Xi_t|| = \sup_\theta \Xi_t \cdot \theta$. Fix the direction θ_0 in which the movement of the spine is altered. Observe that we can decompose the spine as $\Xi_t = \xi_t\theta_0 + Y_t$ where ξ_t and Y_t are independent processes such that ξ_t is a Brownian motion with drift $\sqrt{2}$ conditioned on never hitting $\sqrt{2}t+\phi(t)$ and Y_t is a $(d-1)$-dimensional Brownian motion living in the space $\theta_0^\perp$. Thus

$$||\Xi_t|| = \sqrt{|\xi_t|^2 + ||Y_t||^2}.$$

By Lemma 3.12 almost surely for any $\delta > 0$ there exist C_1, t_0 such that for all $t \geq t_0$, $|\xi_t| \leq \sqrt{2}t - C_1 t^{1/2-\delta}$. Similarly, by e.g. the law of the iterated logarithm, for any $\delta' > 0$ there exists C_2 such that up to enlarging t_0, for all $t \geq t_0$, $||Y_t|| \leq C_2 t^{1/2+\delta'}$. Choose δ, δ' such that $\delta + 2\delta' < 1/2$, then for t large enough,

$$\begin{aligned}
||\Xi_t|| &\leq \sqrt{2t^2 + C_1^2 t^{1-2\delta} - 2\sqrt{2}C_1 t^{3/2-\delta} + C_2^2 t^{1+2\delta'}} \\
&\leq \sqrt{2t^2 + (C_1/2)^2 t^{1-2\delta} - 2\sqrt{2}(C_1/2)t^{3/2-\delta}} \\
&= \sqrt{2}t - C_1/2 t^{1/2-\delta}. \qquad (3.21)
\end{aligned}$$

Let $C_f = \sup_{\mathbb{S}^{d-1}} f(\theta)$. By Lemma 3.9 we know that for some $C \geq 0$, $R^\phi(x,t) \leq C(1+|x|+\phi(t))$ for all $x \in \mathbb{R}, t \geq 0$, thus since the spine particle has zero contribution in the limit,

$$\begin{aligned}
&\limsup_{t\to\infty} \mathbb{P}^f[\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] \\
&\leq C\left\langle \sum_{s\in\mathcal{B}_\infty} (1 + |\sqrt{2}s - \Xi_s\cdot\theta| + \phi(s))e^{\sqrt{2}(\Xi_s\cdot\theta - \sqrt{2}s)}, C_f \right\rangle \\
&\leq P_{d-1}CC_f \sum_{s\in\mathcal{B}_\infty} (1 + \sqrt{2}s + ||\Xi_s|| + \phi(s))e^{\sqrt{2}(||\Xi_s|| - \sqrt{2}s)},
\end{aligned}$$

where P_{d-1} is the surface area of a d dimensional sphere. Combining it with (3.21) we obtain that almost surely there exists a constant C_1 such that

$$\limsup_{t\to\infty} \mathbb{P}^f[\langle Z_t^\phi, f\rangle \mid \mathcal{F}_\infty] \leq P_{d-1}CC_f \sum_{s\in\mathcal{B}_\infty} (1 + 2\sqrt{2}s + \phi(s))e^{-s^{1/2-\delta}C_1/2},$$

which is almost surely finite, as $\mathcal{B}$ is a Poisson point process with intensity 2 and hence $s_i \in \mathcal{B}$ are i.i.d. spaced and $(s_i/i)_{i\geq 1}$ satisfies the Strong Law of Large Numbers. The proof is now complete. ∎

3.4.4 Simultaneous limits on the sphere

The main aim of this section is the proof of the following proposition, which shows that $(Z_t(\theta))$ converges a.s. both on a random set of full Lebesgue measure, and as a random measure.

Proposition 3.26. *Let ϕ be a function satisfying Assumption 3.8. Then almost surely there exists $\Theta \subset \mathbb{S}^{d-1}$ of full Lebesgue measure (i.e. $\sigma(\Theta) = \sigma(\mathbb{S}^{d-1})$) such that for all $\theta \in \Theta$, $Z^\phi_\infty(\theta) := \lim_{t\to\infty} Z^\phi_t(\theta)$ exists a.s., and for any bounded measurable function f,*

$$\lim_{t\to\infty} \langle Z^\phi_t, f\rangle = \langle Z^\phi_\infty, f\rangle \ a.s.$$

Moreover, the limit is almost surely finite.

Proof. Without loss of generality we may and will assume that $f \geq 0$ and $\int_{\mathbb{S}^{d-1}} f(\theta)\sigma(\mathrm{d}\theta) = 1$. The integrated $\mathbb{P}$-martingale $\langle Z^\phi_t, f\rangle$ is non-negative, hence it converges a.s. to some limit, and we set $\mathcal{Z} := \lim_{t\to\infty}\langle Z^\phi_t, f\rangle$. Furthermore, by Lemma 3.23 this martingale is uniformly integrable, thus

$$\mathbb{E}_{\mathbb{P}}\mathcal{Z} = \mathbb{E}_{\mathbb{P}}\langle Z^\phi_0, f\rangle = R^\phi(0,0).$$

We want to show that

$$\mathcal{Z} = \langle \lim_{t\to\infty} Z^\phi_t, f\rangle$$

but *a priori* we don't even know that the right hand side is well defined.

As $Z^\phi_t(\theta) \geq 0$ a.s., we observe that by Fatou's lemma

$$\mathcal{Z} = \liminf_{t\to\infty}\langle Z^\phi_t, f\rangle \geq \langle \liminf_{t\to\infty} Z^\phi_t, f\rangle. \tag{3.22}$$

Note that $\liminf_{t\to\infty} Z^\phi_t(\theta)$ exists simultaneously for all $\theta \in \mathbb{S}^{d-1}$.

On the other hand, by the uniform integrability of $\langle Z^\phi_t, f\rangle$ (Lemma 3.23) and Fubini's theorem,

$$\mathbb{E}_{\mathbb{P}}\mathcal{Z} = \lim_{t\to\infty} \mathbb{E}_{\mathbb{P}}\langle Z_t^\phi, f\rangle = \lim_{t\to\infty} \langle \mathbb{E}_{\mathbb{P}} Z_t^\phi, f\rangle.$$

Since the distribution of $Z_t^\phi(\theta)$ does not depend on θ (used in the first equality), and again using the uniform integrability, but of $Z_t^\phi(\theta)$, and also Fubini's theorem, we obtain that

$$\begin{aligned}\lim_{t\to\infty} \langle \mathbb{E}_{\mathbb{P}} Z_t^\phi, f\rangle = \langle \lim_{t\to\infty} \mathbb{E}_{\mathbb{P}} Z_t^\phi, f\rangle = \langle \mathbb{E}_{\mathbb{P}} \lim_{t\to\infty} Z_t^\phi, f\rangle \\ = \langle \mathbb{E}_{\mathbb{P}} \liminf_{t\to\infty} Z_t^\phi, f\rangle = \mathbb{E}_{\mathbb{P}}\langle \liminf_{t\to\infty} Z_t^\phi, f\rangle.\end{aligned}$$

Thus we have shown that

$$\mathbb{E}_{\mathbb{P}}\mathcal{Z} = \mathbb{E}_{\mathbb{P}}\langle \liminf_{t\to\infty} Z_t^\phi, f\rangle,$$

and recalling (3.22) this means that almost surely

$$\mathcal{Z} = \langle \liminf_{t\to\infty} Z_t^\phi, f\rangle.$$

By Fubini's theorem

$$\mathbb{E}_{\mathbb{P}}\langle \limsup_{t\to\infty} Z_t^\phi, f\rangle = \langle \mathbb{E}_{\mathbb{P}} \lim_{t\to\infty} Z_t^\phi, f\rangle = \mathbb{E}_{\mathbb{P}}\langle \liminf_{t\to\infty} Z_t^\phi, f\rangle,$$

hence almost surely for almost all θ

$$\limsup_{t\to\infty} Z_t^\phi(\theta) = \liminf_{t\to\infty} Z_t^\phi(\theta).$$

Therefore, almost surely $\lim_{t\to\infty} Z_t^\phi(\theta)$ exists simultaneously for all θ besides a random set of Lebesgue surface measure 0, and

$$\lim_{t\to\infty} \langle Z_t^\phi, f\rangle = \langle \lim_{t\to\infty} Z_t^\phi, f\rangle.$$

□

3.4.5 Proof of Theorem 3.3

We start with the following technical lemma:

Lemma 3.27. *Let ϕ be a function satisfying Assumption 3.8. If the function ϕ additionally satisfies $\lim_{t\to\infty}\phi(t) - \frac{d-1}{2\sqrt{2}}\log(1+t) = \infty$ and $\phi'(t) = o(t^{-1/2-\epsilon})$, then for any bounded measurable function f,*

$$\lim_{t\to\infty}\left\langle \sum_{j\in\mathcal{N}_t^{\phi,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta + \phi(t))e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle$$
$$= \left\langle \lim_{t\to\infty}\sum_{j\in\mathcal{N}_t^{\phi,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta + \phi(t))e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle \quad (3.23)$$

almost surely and the limit is finite with probability one.

Note that there are two differences between Lemma 3.27 and Theorem 3.3: firstly, we don't take a sum over all particles, and secondly we have an additional term ϕ appearing. We solve both of these issues in the remainder of this section.

Proof of Lemma 3.27. Since

$$\lim_{t\to\infty}\phi(t) - \frac{d-1}{2\sqrt{2}}\log(1+t) = \infty,$$

from Lemma 3.20 we obtain that

$$\lim_{t\to\infty}\inf_{j\in\mathcal{N}_t}\left(\sqrt{2}t - ||X_t(j)|| + \phi(t)\right) = +\infty$$

and (since $(-X_t(j), j\in\mathcal{N}_t)$ has the same distribution as $(X_t(j), j\in\mathcal{N}_t)$)

$$\limsup_{t\to\infty}\sup_{\theta\in\mathbb{S}^{d-1}, j\in\mathcal{N}_t}\frac{1}{t}(\sqrt{2}t - X_t(j)\cdot\theta) = 2\sqrt{2}$$

almost surely. Recall the definition (3.18). We are now going to make use of the asymptotic behaviour of $R^\phi(x,t)$: we apply Lemma 3.17 with $D > 2\sqrt{2}$ and an arbitrarily small δ to obtain that almost surely

$$\langle Z_\infty^\phi, f\rangle = \langle \lim_{t\to\infty}\sum_{j\in\mathcal{N}_t^{\phi,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta + \phi(t))e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\rangle. \quad (3.24)$$

From Proposition 3.26 we know that

$$\langle Z^\phi_\infty, f\rangle = \lim_{t\to\infty} \langle Z^\phi_t, f\rangle.$$

and again, applying Lemma 3.17 with $D > 2\sqrt{2}$ and an arbitrarily small δ, we obtain that

$$\langle Z^\phi_\infty, f\rangle = \lim_{t\to\infty} \left\langle \sum_{j\in\mathcal{N}_t^{\phi,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta + \phi(t))e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f \right\rangle. \tag{3.25}$$

Combining (3.24) and (3.25) completes the proof. □

We now get rid of the term involving ϕ in (3.23):

Lemma 3.28. *Let ϕ be such that $\phi(t) = o(t^{1/2-\epsilon})$ for some $\epsilon > 0$. Then*

$$\lim_{t\to\infty} \phi(t) \left\langle \sum_{j\in\mathcal{N}_t} e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1 \right\rangle = 0$$

almost surely.

Proof. Without loss of generality assume that $\phi(t)$ is an increasing, concave, $\mathcal{C}^1$-class function such that $\lim_{t\to\infty}\phi(t) - \frac{d-1}{2\sqrt{2}}\log(1+t) = \infty$ and $\phi'(t) = o(t^{-1/2-\epsilon})$. Set $\psi(t) := t^{1/2-\epsilon/2}$ and observe that by Lemma 3.20, for any $\delta > 0$ we can choose A_δ such that with probability $1-\delta$ none of the particles ever hit the sphere of an increasing radius $\sqrt{2}t + \frac{d-1}{2\sqrt{2}}\log(1+t) + A_\delta$, thus conditioning on this event

$$\begin{aligned}
&\limsup_{t\to\infty} \left\langle \sum_{j\in\mathcal{N}_t^{\psi+A_\delta,\theta}} (X_t(j)\cdot\theta - \sqrt{2}t)\mathbb{I}_{\{X_t(j)\cdot\theta\geq\sqrt{2}t\}} e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1 \right\rangle \\
\leq\ &\limsup_{t\to\infty} \left\langle \sum_{j\in\mathcal{N}_t^{\psi+A_\delta,\theta}} (\phi(t)+A_\delta)\mathbb{I}_{\{X_t(j)\cdot\theta\geq\sqrt{2}t\}} e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1 \right\rangle \\
\leq\ &\limsup_{t\to\infty} \frac{\phi(t)+A_\delta}{\psi(t)+A_\delta} \left\langle \sum_{j\in\mathcal{N}_t^{\psi+A_\delta,\theta}} (\psi(t)+A_\delta) e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1 \right\rangle.
\end{aligned} \tag{3.26}$$

Consider the following decomposition:

$$\begin{aligned}
\sqrt{2}t - X_t(j)\cdot\theta + \psi(t) + A_\delta = &(\sqrt{2}t - X_t(j)\cdot\theta)\mathbb{I}_{\{X_t(j)\cdot\theta\geq\sqrt{2}t\}} \\
&+ (\sqrt{2}t - X_t(j)\cdot\theta)\mathbb{I}_{\{X_t(j)\cdot\theta\leq\sqrt{2}t\}} \\
&+ \psi(t) + A_\delta.
\end{aligned} \tag{3.27}$$

Note that only the first term on the right-hand side of (3.27) is negative. Since by Lemma 3.27

$$\lim_{t\to\infty}\left\langle \sum_{j\in\mathcal{N}_t^{\psi+A_\delta,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta + \psi(t) + A_\delta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1\right\rangle \tag{3.28}$$

exists almost surely, from (3.27), (3.26) and $\lim_{t\to\infty}\frac{\phi(t)+A_\delta}{\psi(t)+A_\delta} = 0$ we deduce that the limit

$$\limsup_{t\to\infty}\left\langle \sum_{j\in\mathcal{N}_t^{\psi+A_\delta,\theta}} (\psi(t) + A_\delta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1\right\rangle$$

is finite with probability $1-\delta$: if it wasn't finite with probability larger than δ, then by (3.26) and (3.27), with positive probability (3.28) would diverge to infinity, as its negative part is negligible in comparison to the positive one.

Since $\lim_{t\to\infty}\frac{\phi(t)}{\psi(t)} = 0$, this implies further that with probability $1-\delta$

$$\lim_{t\to\infty}(\phi(t)+A_\delta)\left\langle \sum_{j\in\mathcal{N}_t^{\phi+A_\delta,\theta}} e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, 1\right\rangle = 0.$$

Taking A_δ arbitrarily large completes the proof. □

We are now ready to present the last step of the proof of Theorem 3.3. Recalling Lemma 3.27 we show that in fact we can sum over all the particles and we can still swap integration with taking the limit. As was mentioned before, this is the step where we consider a sequence of functions $\phi\vee A$ for $A\in\mathbb{N}$.

Proof of Theorem 3.3. Set $\phi = t^{1/2-\epsilon}$ for some $\epsilon\in(0,1)$. By combining Lemma 3.27 with Lemma 3.28 we obtain that almost surely for all $A\in\mathbb{N}$

$$\begin{aligned}\lim_{t\to\infty}&\left\langle \sum_{j\in\mathcal{N}_t^{\phi\vee A,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle \\ &= \left\langle \lim_{t\to\infty}\sum_{j\in\mathcal{N}_t^{\phi\vee A,\theta}} (\sqrt{2}t - X_t(j)\cdot\theta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle. \end{aligned}\tag{3.29}$$

By Lemma 3.20 for any δ we can choose A_δ such that the event defined by $B_\delta := \{\forall s > 0, u\in\mathcal{N}_s : ||X_s(j)|| \le \sqrt{2}s + \phi(s)\vee A_\delta\}$ happens with probability $\mathbb{P}(B_\delta)\ge 1-\delta$. Therefore, conditioning on B_δ and taking $A \geq A_\delta$ in (3.29), we obtain that

$$\lim_{t\to\infty}\left\langle \sum_{j\in\mathcal{N}_t}(\sqrt{2}t - X_t(j)\cdot\theta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle = \left\langle \lim_{t\to\infty}\sum_{j\in\mathcal{N}_t}(\sqrt{2}t - X_t(j)\cdot\theta)e^{\sqrt{2}(X_t(j)\cdot\theta-\sqrt{2}t)}, f\right\rangle. \quad (3.30)$$

holds almost surely on B_δ. Taking δ arbitrarily small we conclude that (3.30) holds with probability one, which proves (3.3).

Finally, to show that $\langle Z_\infty(\theta), 1\rangle > 0$ we observe that by Fubini's theorem

$$0 = \int_{\mathbb{S}^{d-1}} \mathbb{P}(Z_\infty(\theta) = 0)\sigma(\mathrm{d}\theta) = \mathbb{E}_{\mathbb{P}}\left[\int_{\mathbb{S}^{d-1}} \mathbb{I}_{\{Z_\infty(\theta)=0\}}\sigma(\mathrm{d}\theta)\right],$$

which completes the proof. □

3.5 Direction of the largest displacement in dimension one

In this section we prove Theorem 3.1 but we start by showing how Corollary 3.2 follows from Theorem 3.1. Set

$$G_t^+ := \sqrt{2}(M_t^+ - m_t - \tfrac{\sqrt{2}}{2}\log Z_\infty)$$
$$\text{and}\quad G_t^- := \sqrt{2}(-M_t^- - m_t - \tfrac{\sqrt{2}}{2}\log Z_\infty^-).$$

Then we can rewrite

$$\mathbb{P}\left(M_t^+ > -M_t^- \;\middle|\; \mathcal{F}_s\right) = \mathbb{P}\left(G_t^+ + \log Z_\infty > G_t^- + \log Z_\infty^- \;\middle|\; \mathcal{G}_s\right).$$

Theorem 3.1 tells us that (G_t^+, G_t^-) conditioned on $\mathcal{G}_s$ converges in the double limit, first letting $t\to\infty$ and then $s\to\infty$, to a pair of independent standard Gumbel random variables. Thus the proof of Corollary 3.2 is a consequence of the following lemma.

Lemma 3.29. *Let $G_1, \ldots, G_n$ be independent standard Gumbel-distributed random variables. Then for any $a_1, \ldots, a_n$,*

$$\mathbb{P}\Big(G_1 + a_1 \geq \max(a_2 + G_2, \ldots, a_n + G_n)\Big) = \frac{e^{a_1}}{\sum_{i=1}^n e^{a_i}}.$$

Proof. Recall that the pdf of the standard Gumbel distribution is given by $e^{-(x+e^{-x})}$, and the cdf is given by $e^{-e^{-x}}$. Then by simple computations, setting $K := \log(1+\sum_{i=2}^n e^{a_i - a_1})$, we have

$$\begin{aligned}
&\mathbb{P}\Big(a_1 + G_1 \geq \max(a_2+G_2, \dots, a_n+G_n)\Big) \\
&= \int_{\mathbb{R}} e^{-(g_1+e^{-g_1})} \prod_{i=2}^n e^{e^{-(a_1+g_1-a_i)}} \mathrm{d}g_1 \\
&= \int_{\mathbb{R}} e^{-g_1 - e^{-g_1}(1+\sum_{i=2}^n e^{a_i-a_1})} \mathrm{d}g_1 \\
&= e^{-K} \int_{\mathbb{R}} e^{-(g_1-K)-e^{-(g_1-K)}} \mathrm{d}g_1 \\
&= \frac{1}{1+\sum_{i=2}^n e^{a_i-a_1}} = \frac{e^{a_1}}{\sum_{i=1}^n e^{a_i}}.
\end{aligned}$$

□

We now prove the main theorem of this section.

Proof of Theorem 3.1. Let $y, z \geq 0$. Note that for all $0 \leq s \leq t$ we have

$$\mathbb{P}\left(M_t^+ - m_t \leq y, -M_t^- - m_t \leq z \;\Big|\; \mathcal{G}_s\right) = \prod_{j \in \mathcal{N}_s} \nu_{s,t}(X_s(j), y, z) \tag{3.31}$$

by the branching property, where we have set

$$\nu_{s,t}(x,y,z) := \mathbb{P}\left(M_{t-s}^+ - m_t \leq y - x, -M_{t-s}^- - m_t \leq z + x\right).$$

We now bound $\nu_{s,t}$ from above and from below to obtain an asymptotically tight estimate for the joint cdf of (M_t^+, M_t^-) given $\mathcal{G}_s$. We begin by computing a lower bound. Observe first that since $m_t - m_{t-s} = \sqrt{2}s + \frac{3}{2\sqrt{2}} \log \frac{t-s}{t}$, from the inequality $x/(x+1) \leq \log(1+x) \leq x$ we obtain that

$$\sqrt{2}s - \frac{3}{2\sqrt{2}} \frac{s}{t-s} \leq m_t - m_{t-s} \leq \sqrt{2}s - \frac{3}{2\sqrt{2}} \frac{s}{t}. \tag{3.32}$$

Therefore, noting that for any events F, G one has $\mathbb{P}(F \cap G) \geq 1 - \mathbb{P}(F^c) - \mathbb{P}(G^c)$, we obtain that

$$\begin{aligned}
\nu_{s,t}(x,y,z) \geq 1 &- \mathbb{P}\left(M_{t-s}^+ - m_{t-s} \geq y - (x - \sqrt{2}s) - \frac{3}{2\sqrt{2}} \frac{s}{t-s}\right) \\
&- \mathbb{P}\left(-M_{t-s}^- - m_{t-s} \geq z + (x + \sqrt{2}s) - \frac{3}{2\sqrt{2}} \frac{s}{t-s}\right).
\end{aligned} \tag{3.33}$$

From [19, Theorem 1] we know that $\mathbb{P}(M_t^+ - m_t \geq x)$ converges uniformly as $t \to \infty$ to $\omega(x)$, which further satisfies

$$1 - \omega(x) \sim c_\star x e^{-\sqrt{2}x} \qquad \text{as } x \to \infty.$$

Hence, (3.33) yields

$$\begin{aligned}
&\liminf_{t\to\infty} \prod_{j\in\mathcal{N}_s} \nu_{s,t}(X_s(j), y, z) \\
\geq &\prod_{j\in\mathcal{N}_s} \Big[1 - \omega(\sqrt{2}s - X_s(j) + y) - \omega(\sqrt{2}s + X_s(j) + z))\Big] \\
\geq &\prod_{j\in\mathcal{N}_s} \Big[1 - \Big\{c_\star(\sqrt{2}s - X_s(j) + y)e^{\sqrt{2}(X_s(j)-\sqrt{2}s-y)} \\
&+ c_\star(\sqrt{2}s + X_s(j) + z))e^{\sqrt{2}(-X_s(j)-\sqrt{2}s-z))}\Big\}(1 + \epsilon(s))\Big],
\end{aligned}$$

where $s \mapsto \epsilon(s)$ is a random process such that $\lim_{s\to\infty} \epsilon(s) = 0$ a.s., where we used that $\liminf_{s\to\infty} \min_{j\in\mathcal{N}_s} \Big\{\sqrt{2}s - |X_s(j)|\Big\} = \infty$ a.s. This result follows plainly from the fact that the additive martingale converges to 0 a.s. which can be found in [44].

Therefore, since for any numbers $a_i \in (0,1)^n$

$$\prod_{i=1}^n (1 - a_i) \geq e^{-\sum_i^n \frac{a_i}{1-a_i}} \geq e^{-\frac{1}{1-\max a_i}\sum_i^n a_i},$$

and recalling also that

$$\lim_{s\to\infty} \max_{j\in\mathcal{N}_s} \Big\{|\sqrt{2}s - X_s(j)|e^{\sqrt{2}(X_s(j)-\sqrt{2}s)}\Big\} = 0$$

almost surely, we obtain from (3.31) that

$$\begin{aligned}
&\liminf_{s\to\infty} \liminf_{t\to\infty} \mathbb{P}\left(M_t^+ - m_t \leq y, -M_t^- - m_t \leq z \;\middle|\; \mathcal{F}_s\right) \\
\geq &\liminf_{s\to\infty} \exp\Big(- c_\star \sum_{j\in\mathcal{N}_s} (y - (X_s(j) - \sqrt{2}s))e^{-\sqrt{2}(y-(X_s(j)-\sqrt{2}s))} \\
&\qquad - c_\star \sum_{j\in\mathcal{N}_s} (z + (X_s(j) + \sqrt{2}s))e^{-\sqrt{2}(z+(X_s(j)+\sqrt{2}s))}\Big).
\end{aligned}$$

Using again that the additive martingale $\sum_{j\in\mathcal{N}_s} e^{\sqrt{2}(X_s(j)-\sqrt{2}s)}$ converges to 0 a.s. we eventually obtain that

$$\begin{aligned}&\liminf_{s\to\infty}\liminf_{t\to\infty}\mathbb{P}\left(M_t^+ - m_t \le y, -M_t^- - m_t \le z \;\middle|\; \mathcal{F}_s\right)\\ &\ge \exp\left(-c_\star Z_\infty e^{-\sqrt{2}y} - c_\star Z_\infty^- e^{-\sqrt{2}z}\right).\end{aligned} \tag{3.34}$$

To obtain a similar upper bound, we use that for any pair of events F, G, $\mathbb{P}(F\cap G) = 1 - \mathbb{P}(F^c) - \mathbb{P}(G^c) + P(F^c \cap G^c)$, hence recalling (3.32),

$$\begin{aligned}\nu_{s,t}(x,y,z) \le 1 &- \mathbb{P}\left(M_{t-s}^+ - m_{t-s} \ge y - (x-\sqrt{2}s) - \frac{3}{2\sqrt{2}}\frac{s}{t}\right)\\ &- \mathbb{P}\left(-M_{t-s}^- - m_{t-s} \ge z + (x+\sqrt{2}s) - \frac{3}{2\sqrt{2}}\frac{s}{t}\right)\\ &+ \zeta_{s,t}(x,y,z),\end{aligned}$$

where

$$\begin{aligned}\zeta_{s,t}(x,y,z) := \mathbb{P}\Big(&M_{t-s}^+ - m_{t-s} \ge y - (x-\sqrt{2}s) - \frac{3}{2\sqrt{2}}\frac{s}{t},\\ &- M_{t-s}^- - m_{t-s} \ge z + (x+\sqrt{2}s) - \frac{3}{2\sqrt{2}}\frac{s}{t}\Big).\end{aligned}$$

Note that

$$\zeta_{s,t}(x,y,z) \le \mathbb{P}\Big(M_{t-s}^+ - m_{t-s} \ge y\wedge z + \sqrt{2}s - \frac{3}{2\sqrt{2}}\frac{s}{t}\Big),$$

hence

$$\limsup_{t\to\infty}\zeta_{s,t}(x,y,z) \le \omega(y\wedge z + \sqrt{2}s).$$

As a result, with similar computations as in the proof of the lower bound, there exists a process $\epsilon(s)$ converging a.s. to 0 as $s\to\infty$ such that,

$$\begin{aligned}&\limsup_{t\to\infty}\prod_{j\in\mathcal{N}_s}\nu_{s,t}(X_s(j),y,z)\\ &\le \prod_{j\in\mathcal{N}_s}\Big[1 - \Big\{c_\star(\sqrt{2}s - X_s(j) + y)e^{\sqrt{2}(X_s(j)-\sqrt{2}s-y)}\\ &\quad + c_\star(\sqrt{2}s + X_s(j) + z))e^{\sqrt{2}(-X_s(j)-\sqrt{2}s-z)}\\ &\quad - c_\star(\sqrt{2}s + y\wedge z)e^{\sqrt{2}(-\sqrt{2}s-y\wedge z)}\Big\}(1+\epsilon(s))\Big].\end{aligned}$$

Using that for any numbers $a_i < 1$, $\prod_{i=1}^n (1-a_i) \leq e^{-\sum_i^n a_i}$, and noting that for any $C \in \mathbb{R}$

$$\lim_{s\to\infty} \sum_{j\in\mathcal{N}_s} (\sqrt{2}s + C)e^{\sqrt{2}(-\sqrt{2}s-C)} = 0$$

almost surely, we obtain that

$$\begin{aligned}&\limsup_{s\to\infty}\limsup_{t\to\infty} \mathbb{P}(M_t^+ - m_t \leq y, -M_t^- - m_t \leq z \mid \mathcal{F}_s) \\ &\leq \exp(-c_\star Z_\infty e^{-\sqrt{2}y} - c_\star Z_\infty^- e^{-\sqrt{2}z}),\end{aligned}$$

which, together with (3.34), completes the proof. □

Remark 3.30. With similar computations to the ones made in the proof of Theorem 3.1 we would be able to prove joint convergence in distribution of

$$\left(\sum_{j\in\mathcal{N}_t} \delta_{X_t(j)-m_t}, \sum_{j\in\mathcal{N}_t} \delta_{-X_t(j)-m_t}\right)$$

towards a pair of decorated Poisson point processes with random intensities $c_\star Z_\infty e^{-\sqrt{2}x}\mathrm{d}x$ and $c_\star Z_\infty^- e^{-\sqrt{2}x}\mathrm{d}x$ respectively, and such that these processes are independent conditionally on (Z_∞, Z_∞^-). This result can be thought of as a unidimensional version of Conjecture 3.4.

Chapter 4

Conclusions

Minimax functions on Galton-Watson trees In Chapter 2 we studied minimax functions on Galton-Watson trees. Our main object of interest was the distribution of the value assigned to the root in such trees. In Theorem 2.2 we described possible limits of such distributions as the height of the tree tends to infinity. We proved in particular that if f (recall the definition in equation (2.1)) is not the identity function, then the limiting distribution is discrete with atoms located exactly at fixed points of f unstable from at least one side. In Theorem 2.3 we fixed one such atom and conditioned the random variable at the root of the tree to be close to this atom. For this setting we determined the rate of convergence in different regimes (i.e. for different offspring distributions of the tree). Finally, in Theorem 2.5 we provided a condition for the stationary minimax tree to be endogenous.

Multidimensional branching Brownian motion In Chapter 3 we defined the derivative martingale as a process on the sphere. In Theorem 3.3 we proved that it converges weakly, which was the main result of this chapter and a stepping stone for future studies.

The long-term goal would be to prove a multi-dimensional analogue of the convergence of the extremal process of the branching Brownian motion. Such a result in the one-dimensional case was proved in [9] and [1]. Set

$$r_t := \sqrt{2}t + \frac{d-4}{2\sqrt{2}} \log t.$$

From [50] we know that, up to an $O(1)$ error, r_t is equal to the median position of the maximal displacement of the d-dimensional branching Brownian motion at time t. Define the direction of a particle u at time t by $D_t(u) := X_t(u)/\|X_t(u)\|$ for $t \geq 0, u \in \mathcal{N}_t$. We conjectured the following:

Conjecture 3.4. *There exists $c_d^\star > 0$ such that*

$$\lim_{t\to\infty} \sum_{u\in\mathcal{N}_t} \delta_{D_t(u),\|X_t(u)\|-r_t} = \mathcal{L}(\mathrm{d}\theta, \mathrm{d}x) \text{ in law,}$$

where $\mathcal{L}$ is a decorated Poisson point process that can be constructed as follows. Let $(\theta_j, \xi_j)_{j\geq 1}$ be the atoms of a Poisson point process with intensity $c_d^\star Z_\infty(\theta)\sigma(\mathrm{d}\theta)e^{-\sqrt{2}x}\mathrm{d}x$ and $(D_j, j \geq 1)$ be i.i.d. point processes on $\mathbb{R}$ with common distribution $\mathcal{D}$. Then

$$\mathcal{L} = \sum_{j\geq 1}\sum_{x\in D_j} \delta_{\theta_j,\xi_j+x}.$$

This conjecture can be interpreted as follows. Recall that the extremal process of the one-dimensional BBM, i.e. the limit of

$$\sum_{u\in\mathcal{N}_t} \delta_{X_t(u)-m_t}$$

can be constructed as a decorated Poisson point process with (random) intensity

$$c_\star Z_\infty e^{-\sqrt{2}x}\mathrm{d}x,$$

where $c_\star$ is a positive constant. The decoration is constructed as the limiting process of a BBM conditioned on the event that the maximal displacement exceeds $\sqrt{2}t$.

In the multidimensional case we claim that we should again recover a Poisson point process, since for any two angles θ_1, θ_2, particles with large projections on θ_1 and θ_2 at time t form with high probability two disjoint families. Therefore, the behaviour in two distinct directions can be seen as almost independent (conditionally on the initial behaviour), which should lead to a Poisson point process. Secondly, we claim that the intensity at direction θ of the limiting process should be proportional to $Z_\infty(\theta)$, since the projection of the multidimensional BBM on θ is a one-dimensional BBM, in which case the limiting intensity is proportional to Z_∞. Moreover, we believe that the factor $e^{-\sqrt{2}x}$ comes from the fact that for any direction θ and for any constant A, if we look at particles with projections on θ at time t exceeding $r_t - A$, then with high probability projections of these particles onto the space orthogonal to θ will be of order $O(\sqrt{t})$. Therefore, the radial size of a cone in which all such particles are located tends to zero as t goes to infinity. On the other hand, in direction θ we see exactly the one-dimensional BBM, we thus expect to see the same decoration emerging, which results in the intensity factor $e^{-\sqrt{2}x}$. To sum up, we expect to see an intensity proportional to $Z_\infty(\theta)\sigma(\mathrm{d}\theta)e^{-\sqrt{2}x}\mathrm{d}x$, and we claim that the proportionality constant will depend on the dimension (recall that in the multidimensional case we centre at r_t, and not at m_t).

We believe that Conjecture 3.4 could be proved using the concept of 'clan leaders' introduced in the proofs of the one-dimensional case. More precisely, for all $s \in [0,t]$ we say that a particle is an s-clan leader if it belongs to the set

$$\mathcal{C}_t^s := \{u \in \mathcal{N}_t : \forall v \in \mathcal{N}_t, ||X_t(v)|| < ||X_t(u)|| \text{ or } \mathrm{MRCA}(u,v) > s\}, \tag{4.1}$$

where $\mathrm{MRCA}(u,v)$ is the age of the most recent common ancestor of u and v. In other words, $u \in \mathcal{C}_t^s$ if and only if $||X_t(u)||$ is maximal among the individuals that share a common ancestor with u in the last s units of time. The point process of the s-clan leaders is defined as

$$L_t^s := \sum_{u \in \mathcal{C}_t^s} \delta_{X_t(u)/||X_t(u)||, ||X_t(u)|| - r_t}. \tag{4.2}$$

We thus conjecture the following intermediate step.

Conjecture 4.1. *If s satisfies $\lim_{t\to\infty} s(t) = \lim_{t\to\infty} t - s(t) = \infty$, then we have*

$$\lim_{t\to\infty} L_t^{s(t)} = L_\infty \text{ in law,}$$

where L_∞ is a Poisson point process on $\mathbb{S}^{d-1} \times \mathbb{R}$ with (random) intensity $c_d^\star Z_\infty(\mathrm{d}\theta)e^{-\sqrt{2}x}\mathrm{d}x$.

This conjecture can be seen as a multidimensional version of Theorem 1.1 of [7]. Moving one step backwards, to prove Conjecture 4.1 one needs to show that

$$\left(\max_{u\in\mathcal{N}_t} ||X_t(u)|| - r_t, t \geq 0\right) \tag{4.3}$$

is not only tight (which was proved by Mallein [50]) but that it converges in distribution. This is a subject of our ongoing work.

Bibliography

[1] E. Aïdékon, J. Berestycki, E. Brunet, and Z. Shi. Branching Brownian motion seen from its tip. *Probability Theory and Related Fields*, 157(1-2):405–451, 2013.

[2] E. Aidekon and Z. Shi. The Seneta–Heyde scaling for the branching random walk. *Annals of Probability*, 42(3):959–993, 2014.

[3] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110, pages 1–72. Springer Berlin Heidelberg, 2004.

[4] D. J. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Annals of Applied Probability*, 15(2):1047–1110, 2005.

[5] T. Ali Khan, L. Devroye, and R. Neininger. A limit law for the root value of minimax trees. *Electronic Communications in Probability*, 10:273–281, 2005.

[6] L.-P. Arguin, A. Bovier, and N. Kistler. Genealogy of extremal particles of branching Brownian motion. *Communications on Pure and Applied Mathematics*, 64(12):1647–1676, 2011.

[7] L.-P. Arguin, A. Bovier, and N. Kistler. Poissonian statistics in the extremal process of branching Brownian motion. *Annals of Applied Probability*, 22(4):1693–1711, 2012.

[8] L.-P. Arguin, A. Bovier, and N. Kistler. An ergodic theorem for the frontier of branching Brownian motion. *Electronic Journal of Probability*, 18, 2013.

[9] L.-P. Arguin, A. Bovier, and N. Kistler. The extremal process of branching Brownian motion. *Probability Theory and Related Fields*, 157(3-4):535–574, 2013.

[10] R. Basu, A. E. Holroyd, J. B. Martin, and J. Wästlund. Trapping games on random boards. *Annals of Applied Probability*, 26(6):3727–3753, 2016.

[11] J. Bertoin and R. A. Doney. On conditioning a random walk to stay nonnegative. *Annals of Probability*, 22(4):2152–2167, 1994.

[12] J. D. Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, 14(1):25–37, 1977.

[13] J. D. Biggins. Uniform convergence of martingales in the branching random walk. *Annals of Probability*, 20(1):137–151, 1992.

[14] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*. Cambridge University Press, 1987.

[15] C. Bordenave, M. Lelarge, and J. Salez. Matchings on infinite graphs. *Probability Theory and Related Fields*, 157(1-2):183–208, 2013.

[16] P. Boutaud and P. Maillard. A revisited proof of the Seneta-Heyde norming for branching random walks under optimal assumptions. *Electronic Journal of Probability*, 24, 2019.

[17] A. Bovier and L. Hartung. Extended convergence of the extremal process of branching Brownian motion. *Annals of Applied Probability*, 27(3):1756–1777, 2017.

[18] M. D. Bramson. Maximal displacement of branching Brownian motion. *Communications on Pure and Applied Mathematics*, 31(5):531–581, 1978.

[19] M. D. Bramson. *Convergence of Solutions of the Kolmogorov Equation to Travelling Waves*. American Mathematical Society: Memoirs of the American Mathematical Society. American Mathematical Society, 1983.

[20] N. Broutin, L. Devroye, and N. Fraiman. Recursive functions on conditional Galton-Watson trees. *Random Structures & Algorithms*, 57(2):304–316, 2020.

[21] N. Broutin and C. Mailler. And/Or trees: a local limit point of view. *Random Structures & Algorithms*, 53(1):15–58, 2018.

[22] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton. A survey of Monte Carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, 4(1):1–43, 2012.

[23] B. Chauvin. Product martingales and stopping lines for Branching brownian motion. *Annals of Probability*, 19(3):1195–1205, 1991.

[24] B. Chauvin and A. Rouault. KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probability Theory and Related Fields*, 80(2):299–314, 1988.

[25] X. Chen. A necessary and sufficient condition for the nontrivial limit of the derivative martingale in a branching random walk. *Advances in Applied Probability*, 47(3):741–760, 2015.

[26] X. Chen, V. Dagard, B. Derrida, Y. Hu, M. Lifshits, and Z. Shi. The Derrida–Retaux conjecture on recursive models, 2019, 1907.01601.

[27] X. Chen, B. Derrida, Y. Hu, M. Lifshits, and Z. Shi. A max-type recursive model: some properties and open questions. In *Sojourns in Probability Theory and Statistical Physics - III*, pages 166–186. Springer Singapore, 2019.

[28] P. Collet, J. P. Eckmann, V. Glaser, and A. Martin. Study of the iterations of a mapping associated to a spin glass model. *Communications in Mathematical Physics*, 94(3):353–370, 1984.

[29] R. A. Doney. A limit theorem for a class of supercritical branching processes. *Journal of Applied Probability*, 9(4):707–724, 1972.

[30] R. Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[31] S. N. Evans. Two representations of a conditioned superprocess. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 123(5):959–971, 1993.

[32] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, 1937.

[33] S. Gelly, L. Kocsis, M. Schoenauer, M. Sebag, D. Silver, C. Szepesvári, and O. Teytaud. The grand challenge of computer Go: Monte Carlo tree search and extensions. *Communications of the ACM*, 55(3):106–113, 2012.

[34] R. Hardy and S. C. Harris. A new formulation of the spine approach to branching diffusions. *arXiv preprint math/0611054*, 2006.

[35] R. Hardy and S. C. Harris. A spine approach to branching diffusions with applications to L^p-convergence of martingales. In *Séminaire de Probabilités XLII*, pages 281–330. Springer Berlin Heidelberg, 2009.

[36] S. C. Harris. Travelling-waves for the FKPP equation via probabilistic arguments. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 129(3):503–517, 1999.

[37] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):226–242, 2017.

[38] A. E. Holroyd and J. B. Martin. Galton-Watson games. *arXiv preprint arXiv:1904.04150*, 2019.

[39] Y. Hu. The almost sure limits of the minimal position and the additive martingale in a branching random walk. *Journal of Theoretical Probability*, 28(2):467–487, 2015.

[40] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Annals of Probability*, 37(2):742–789, 2009.

[41] J. F. C. Kingman. The first birth problem for an age-dependent branching process. *The Annals of Probability*, 3(5):790–801, 1975.

[42] A. N. 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. *Bull. Univ. Moskow, Ser. Internat., Sec. A*, 1:1–25, 1937.

[43] A. E. Kyprianou. Travelling wave solutions to the KPP equation: alternatives to Simon Harris' probabilistic analysis. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 40(1):53–72, 2004.

[44] S. P. Lalley and T. Sellke. A conditional limit theorem for the frontier of a branching Brownian motion. *Annals of Probability*, 15(3):1052–1061, 1987.

[45] R. Lyons. A simple path to Biggins' martingale convergence for branching random walk. In *Classical and Modern Branching Processes*, pages 217–221. Springer New York, 1997.

[46] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Annals of Probability*, 23(3):1125–1138, 1995.

[47] T. Mach, A. Sturm, and J. M. Swart. A new characterization of endogeny. *Mathematical Physics, Analysis and Geometry*, 21:30, 2018.

[48] T. Madaule. First order transition for the branching random walk at the critical parameter. *Stochastic Processes and their Applications*, 126(2):470 – 502, 2016.

[49] P. Maillard and M. Pain. 1-stable fluctuations in branching Brownian motion at critical temperature I: The derivative martingale. *Annals of Probability*, 47(5):2953–3002, 2019.

[50] B. Mallein. Maximal displacement of d-dimensional branching Brownian motion. *Electronic Communications in Probability*, 20(0), 2015.

[51] J. B. Martin and R. Stasiński. Minimax functions on Galton–Watson trees. *Combinatorics, Probability and Computing*, 29(3):455–484, 2020.

[52] H. P. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on Pure and Applied Mathematics*, 28(3):323–331, 1975.

[53] A. Montanari. Statistical mechanics and algorithms on sparse and random graphs. https://web.stanford.edu/~montanar/OTHER/STATMECH/stflour.pdf, 2013.

[54] J. Neveu. Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987*, pages 223–242. Birkhäuser Boston, 1988.

[55] J. Nolen, J.-M. Roquejoffre, and L. Ryzhik. Refined long-time asymptotics for Fisher–KPP fronts. *Communications in Contemporary Mathematics*, 21(07):1850072, 2019.

[56] A. A. Novikov. Martingales, Tauberian theorem, and strategies of gambling. *Theory of Probability & Its Applications*, 41(4):716–729, 1997.

[57] J. Pearl. Asymptotic properties of minimax trees and game-searching procedures. *Artificial Intelligence. An International Journal*, 14(2):113–138, 1980.

[58] R. Pemantle and M. D. Ward. Exploring the average values of Boolean functions via asymptotics and experimentation. In *Proceedings of the Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 253–262. SIAM, 2006.

[59] M. I. Roberts. Spine changes of measure and branching diffusions (Ph.D. thesis), 2010.

[60] M. I. Roberts. A simple path to asymptotics for the frontier of a branching Brownian motion. *Annals of Probability*, 41(5):3518–3541, 2013.

[61] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[62] A. V. Skorokhod. Branching diffusion processes. *Theory of Probability & Its Applications*, 9(3):445–449, 1964.

[63] R. Stasiński, J. Berestycki, and B. Mallein. Derivative martingale of the branching Brownian motion in dimension $d \geq 1$. *arXiv preprint arXiv:2004.00162*, 2020.

[64] E. Subag and O. Zeitouni. Freezing and decorated Poisson point processes. *Communications in Mathematical Physics*, 337(1):55–92, 2015.

[65] K. Uchiyama. Brownian first exit from and sojourn over one sided moving boundary and application. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 54(1):75–116, 1980.

[66] T. Yang and Y.-X. Ren. Limit theorem for derivative martingale at criticality w.r.t. branching Brownian motion. *Statistics & Probability Letters*, 81(2):195–200, 2011.