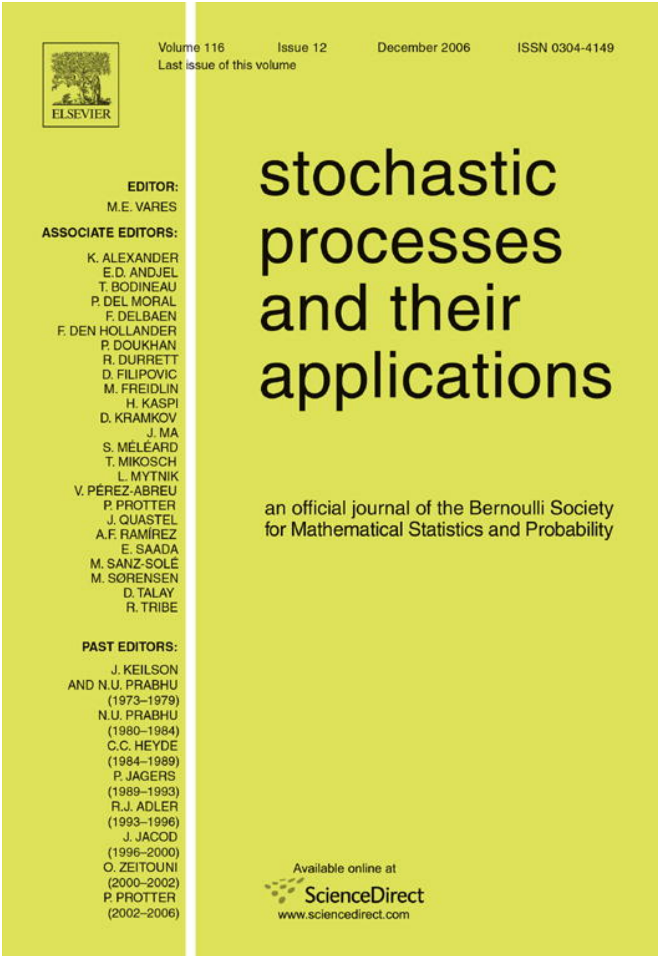
Volume 116
Issue 12
December 2006
ISSN 0304-4149
Last issue of this volume
ELSEVIER
stochastic processes and their applications
an official journal of the Bernoulli Society for Mathematical Statistics and Probability
EDITOR:
M.E. VARES
ASSOCIATE EDITORS:
K. ALEXANDER
E.D. ANDJEL
T. BODINEAU
P. DEL MORAL
F. DELBAEN
F. DEN HOLLANDER
P. DOUKHAN
R. DURRETT
D. FILIPOVIC
M. FREIDLIN
H. KASPI
D. KRAMKOV
J. MA
S. MÉLÉARD
T. MIKOSCH
L. MYTNIK
V. PÉREZ-ABREU
P. PROTTER
J. QUASTEL
A.F. RAMÍREZ
E. SAADA
M. SANZ-SOLÉ
M. SØRENSEN
D. TALAY
R. TRIBE
PAST EDITORS:
J. KEILSON AND N.U. PRABHU (1973–1979)
N.U. PRABHU (1980–1984)
C.C. HEYDE (1984–1989)
P. JAGERS (1989–1993)
R.J. ADLER (1993–1996)
J. JACOD (1996–2000)
O. ZEITOUNI (2000–2002)
P. PROTTER (2002–2006)
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 116 (2006) 1992–2013
stochastic processes and their applications
www.elsevier.com/locate/spa

A conceptual approach to a path result for branching Brownian motion

Robert Hardy, Simon C. Harris*

Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK

Received 30 September 2004; received in revised form 25 April 2006; accepted 23 May 2006
Available online 23 June 2006

Abstract

We give a new, intuitive and relatively straightforward proof of a path large-deviations result for branching Brownian motion (BBM) that can be thought of as extending Schilder's theorem for a single Brownian motion. Our conceptual approach provides an elegant and striking new application of a change of measure technique that induces a 'spine' decomposition and builds on the new foundations for the use of spines in branching diffusions recently developed in Hardy and Harris [Robert Hardy, Simon C. Harris, A new formulation of the spine approach to branching diffusions, 2004, no. 0404, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/~massch/Research/Papers/spine-foundations.pdf; Robert Hardy, Simon C. Harris, Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales, 2004, no. 0405, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/~massch/Research/Papers/spine-Lp-cgce.pdf], itself inspired by related works of Kyprianou [Andreas Kyprianou, Travelling wave solutions to the K–P–P equation: alternatives to Simon Harris's probabilistic analysis, Ann. Inst. H. Poincaré Probab. Statist. 40 (1) (2004) 53–72] and Lyons et al. [Russell Lyons, A simple path to Biggins' martingale convergence for branching random walk, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221; Thomas Kurtz, Russell Lyons, Robin Pemantle, Yuval Peres, A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185; Russell Lyons, Robin Pemantle, Yuval Peres, Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes, Ann. Probab. 23 (3) (1995) 1125–1138]. Some of the techniques developed here will also apply in more general branching Markov processes, for example, see Hardy and Harris [Robert Hardy, Simon C. Harris, A spine proof of

* Corresponding author. Tel.: +44 1225386015; fax: +44 1225386492.
E-mail address: S.C.Harris@bath.ac.uk (S.C. Harris).

doi:10.1016/j.spa.2006.05.010

a lower bound for a typed branching diffusion, 2004, no. 0408, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/~massch/Research/Papers/spine-oubbm.pdf].

MSC: primary 60J80; secondary 60F10; 60J65

Keywords: Branching processes; Branching Brownian motion; Size-biased measures; Spine decomposition; Path large deviations; Schilder's theorem; Brownian motion

1. Overview

Suppose that under a measure $\tilde{P}$ the process $(\xi_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}$. The path large-deviations behaviour of ξ_t is controlled by Schilder's theorem (see Varadhan's proof in [20]), and in order to state this we first define a rescaling of the paths down to the time interval $s \in [0, 1]$:

Definition 1.1. If $(\xi_t)_{0\leq t\leq T}$ is the path in $\mathbb{R}$ followed over the time interval $t \in [0, T]$, then we define $(\xi^T(s))_{0\leq s\leq 1}$ to be a scaled-down version of this path:

$$\xi^T(s) := T^{-1}\xi_{sT},$$

and refer to this ξ^T as the time-T rescaled path.

Such a scaling of the path is also used in the related BBM large-deviations work of Git [5] and is equivalent to supposing that ξ^T is a Brownian motion on $[0, 1]$ with variance $\varepsilon(T) := 1/\sqrt{T}$, where we will soon consider $T \to \infty$ behaviour. Without losing generality, we can suppose that under $\tilde{P}$ the Brownian motion starts at the origin.

Definition 1.2. We use the label $C[0, 1]$ to refer to the set of all continuous functions on $[0, 1]$. We use $C_0[0, 1]$ to mean the set of paths $g \in C[0, 1]$ with $g(0) = 0$ whose derivative is square integrable.

Theorem 1.3 (*Schilder*). *There is a large-deviation principle for Brownian motion:*

- *Upper bound: If C is a closed subset of $C[0, 1]$ then*
$$\limsup_{T\to\infty} T^{-1} \log \tilde{P}(\xi^T \in C) \leq -\inf_{g\in C} I(g).$$
- *Lower bound: If V is an open subset of $C[0, 1]$ then*
$$\liminf_{T\to\infty} T^{-1} \log \tilde{P}(\xi^T \in V) \geq -\inf_{g\in V} I(g),$$

where
$$I(g) := \begin{cases} \int_0^1 \frac{1}{2} g'(s)^2\,\mathrm{d}s & \text{if } g \in C_0[0,1] \\ +\infty & \text{otherwise.} \end{cases}$$

Now consider a branching Brownian motion with constant branching rate r, which is the branching process whereby particles diffuse independently according to a Brownian motion on $\mathbb{R}$ and at any moment undergo fission at a rate r to produce two particles. We suppose that the probabilities of this are $\{P^x : x \in \mathbb{R}\}$ so that P^x is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of the process initiated from a single particle positioned at x.

Suppose that the configuration of this branching Brownian motion at time t is given by the point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t is the set of individuals alive at time t. Without loss of generality we suppose that the initial ancestor of the BBM starts out at the origin, and henceforth use P to mean P^0. We can likewise define a rescaling of the paths; we note that a particle u is born at the time $S_u - \sigma_u$ where S_u is the 'death' (fission) time and σ_u is the 'lifetime' of particle u, but for times earlier than this we interpret $X_u(t)$ as the spatial position of the unique ancestor of u that was alive at time t.

Definition 1.4. For each $T \geq 0$ and each $u \in N_T$ with path $X_u : [0, T] \to \mathbb{R}$, we define the function X_u^T on $[0, 1]$ to be the *time-T rescaled path*:

$$X_u^T : [0, 1] \to \mathbb{R}, \quad X_u^T(s) = T^{-1} X_u(sT).$$

In this article we will prove the following theorem concerning the probability that the path of *at least one* of the many particles in the branching diffusion stays near to a given continuous function.

Theorem 1.5. *There is a large-deviation result for the paths of a BBM:*

- *Upper bound: If C is a closed subset of $C[0, 1]$ then*

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \leq -\inf_{g\in C} S(g). \tag{1}$$

- *Lower bound: If V is an open subset of $C[0, 1]$ then*

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in V) \geq -\inf_{g\in V} S(g), \tag{2}$$

where

$$S(g) := \begin{cases} \sup_{w\in[0,1]} \left(\int_0^w \left\{ \frac{1}{2} g'(s)^2 - r \right\} \mathrm{d}s \right) & \text{if } g \in C_0[0, 1], \\ +\infty & \text{otherwise.} \end{cases}$$

This large-deviations result was previously proven by Tzong-Yow Lee [13] by relying heavily on Friedlin's work on rescalings of solutions of reaction–diffusion equations. In contrast, our approach is based on a conceptual spine change of measure technique and offers a very intuitive, neat and independent proof that can also be generalized to cover many different types of branching diffusions — in Hardy and Harris [9] we develop the ideas to deal with the typed branching diffusion originally studied in Harris and Williams [10].

The main application of Theorem 1.5 is for 'rare' or 'difficult' paths with $S(g) > 0$ where the probability any particle has stayed close to the scaled path decays roughly like $\exp\{-tS(g)\}$. However, we note that for some paths g we shall have $S(g) = 0$: for example if $g(s) = \lambda s$ with $\lambda^2 < 2r$. The large-deviations lower bound will then suggest that there is always a probability that a BBM path of this shape is present. In fact, in this regime, a stronger result has been proven by Git [5] which essentially states that *almost surely* we can be sure to have not just one of these paths with $S(g) = 0$ present in the BBM but an *exponentially growing number*. Indeed, intuitive spine techniques closely related to those contained within are in advanced development for this regime and we expect to cover such complementary results in a future paper.

Outline of proof

As far as the topological issues in our arguments are concerned, the main reference is Dembo and Zeitouni [4]. It is known that the δ-neighbourhoods make up a base for the topology of $C[0,1]$ induced by the metric $\|f\| := \sup_{w\in[0,1]} |f(w)|$.

Definition 1.6. For a given $g \in C[0,1]$ and $\delta > 0$ we define

$$\mathbb{B}_\delta(g) := \{f \in C[0,1] : \|f-g\| < \delta\},$$

as the δ-neighbourhood around the function g.

We aim to prove Theorem 1.5 by using spine techniques to initially prove the following local result:

Theorem 1.7. *For any fixed* $g \in C[0,1]$ *we have*

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g); \tag{3}$$

$$\lim_{\delta\to 0}\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g). \tag{4}$$

In Section 2, we shall give a brief review of some essential spine foundations from Hardy and Harris [7]. We will extend the usual BBM set-up by identifying a distinguished particle, the 'spine', and think of the enriched process as constructed by first laying down the path of the spine as a Brownian motion giving birth to independent BBMs as a Poisson process along its length. Importantly, we can change measure using martingales involving only the spine and its birth process so that the BBM under the new measure can be constructed by simply modifying the behaviour along the spine (including its birth rate). More significantly, projecting suitably chosen such martingales onto the original filtration (without information about the spine) gives rise to 'additive' martingales. In this way, our particular set-up and its various filtrations brings great flexibility and simplicity later on, but also see Chauvin and Rouault [2] and Kyprianou [12].

In Section 3, we prove the local upper bound (3) by first using the Many-to-One Theorem 2.8 to reduce the problem to one about only the spine's behaviour which enables us to directly appeal to Schilder's theorem.

In Section 4, by projecting down from a martingale involving only the spine and the number of births along the spine, we introduce the additive martingale Z_{g_T} for the branching Brownian motion, where for each $T \ge 0$ and a differentiable function $g_T : [0,T] \to \mathbb{R}$,

$$Z_{g_T}(t) := \mathrm{e}^{-rt}\sum_{u\in N_t}\exp\left(\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s\right) \quad \text{for } t \in [0,T]. \tag{5}$$

We can define a new measure $\mathbb{Q}_T$ using this martingale via $\mathrm{d}\mathbb{Q}_T/\mathrm{d}P = Z_{g_T}(T)$ on $\mathcal{F}_T$. Under $\mathbb{Q}_T$, the process can be constructed on $[0,T]$ by running the spine as a BM with drift $g_T'(s)$ at time s, giving birth at an accelerated rate $2r$ to independent BBMs. In particular, we will take $g_T(s) := Tg(s/T)$ so that under $\mathbb{Q}_T$ the spine will naturally stay 'close' to the path of interest in Theorem 1.7. Of course, note that the simple special case of $g_T(s) = \lambda s$ gives the celebrated additive martingale from McKean [17] and Neveu [19] which is so fundamental in many other studies of BBM.

Now consider the following simple idea: for any $F_T \in \mathcal{F}_T$,

$$P(F_T) = \mathbb{Q}_T\left(\frac{1}{Z_{g_T}(T)}; F_T\right),$$

therefore, if F_T is the event that *at least one* particle stays 'close' to the path of interest over time interval $[0, T]$, the behaviour of the spine can guarantee that $\mathbb{Q}_T(F_T) \to 1$ as T goes to infinity, and a suitable lower bound on $Z_{g_T}(T)^{-1}$ under $\mathbb{Q}_T$ should then yield a lower bound for the probability $P(F_T)$. The path and time dependence inherent in the $Z_{g_T}(T)$ martingales can make them difficult to investigate, however, in Section 5, we introduce a spine technique to obtain a suitable upper bound on the growth of $Z_{g_T}(T)$ under $\mathbb{Q}_T$. In Section 6, we can then make this rough argument rigorous and prove the local lower bound at (4).

In Section 7, we will first see that an extension of a topological-type theorem from Dembo and Zeitouni means that the local results of Theorem 1.7 imply the existence of (at worst) a *weak* large-deviation result for the (sub-additive) probabilities of Theorem 1.5; that is, the lower bound holds in full for any open set $V \subset C[0, 1]$ but that the upper bound requires $C \subset C[0, 1]$ to be closed and *compact*. We will then use the Many-to-One Theorem 2.8 with the concept of *exponential tightness* for a single Brownian motion (the spine) to improve these local results to the full large-deviations statement of Theorem 1.5.

Note: (a) For the special case $r = 0$, we trivially note that Theorem 1.5 becomes Schilder's theorem for a single BM. (b) When $r > 0$, strictly speaking the above result is not a large-deviations *principle* in the precise sense since the probabilities $P(\exists u \in N_T : X_u^T \in \cdot)$ do not combine additively but are only *sub*-additive, and therefore cannot define a measure on $C[0, 1]$. For example, if for $T > 0$ we define a set function $\mu_T : C[0, 1] \to [0, 1]$ via

$$\mu_T(A) := P(\exists u \in N_T : X_u^T \in A),$$

then μ_T is sub-additive: $\forall A, A_1, A_2 \in C[0, 1]$ with $A \subset A_1 \cup A_2, \mu_T(A) \leq \mu_T(A_1) + \mu_T(A_2)$. However, consider the two sets $A_1 := \{g \in C[0, 1] : g(1) < 0\}$ and $A_2 := \{g \in C[0, 1] : g(1) \geq 0\}$, for which $A_1 \cup A_2 = C[0, 1]$. For T large enough, there will be so many Brownian particles in the BBM that we are highly likely to find at least one particle has a path in A_1 and at least one other in A_2, hence $\mu_T(A_1) + \mu_T(A_2) \simeq 1 + 1 \neq \mu_T(A_1 \cup A_2) = 1$. Thus, μ_T is *not* additive. Therefore, one must be careful when following 'standard arguments' from large-deviations theory. To reassure the reader, in Section 7 we are particularly careful to give full proofs of the topological properties that we use by closely mirroring some results from Dembo and Zeitouni [4] where, in fact, the issue of sub-additivity versus additivity only enters at two points corresponding to our Eqs. (28) and (31).

2. The spine approach foundations

In preparation for the proof of our main result, we must briefly review the formal constructions on which our spine analysis is based — full details can be consulted in the foundation articles Hardy and Harris [6] or [7]. The reader who is familiar with the work of Lyons et al. [15,11,16], or with Kyprianou's paper [12] will notice significant differences in our approach via our use of the filtrations on the single underlying space.

The set of Ulam–Harris labels is to be equated with the set Ω of finite sequences of strictly positive integers:

$$\Omega := \{\emptyset\} \cup \bigcup_{n \in \mathbb{N}} (\mathbb{N})^n,$$

where we take $\mathbb{N} = \{1, 2, \ldots\}$. For two words $u, v \in \Omega$, uv denotes the concatenated word ($u\emptyset = \emptyset u = u$), and therefore Ω contains elements like '213' (or '$\emptyset$213'), which we read as 'the individual being the 3rd child of the 1st child of the 2nd child of the initial ancestor $\emptyset$'. For two labels $v, u \in \Omega$ the notation $v < u$ means that v is an *ancestor* of u, and $|u|$ denotes the length of u. The set of all ancestors of u is equally given by

$$\{v : v < u\} = \{v : \exists w \in \Omega \text{ such that } vw = u\}.$$

Collections of labels, i.e. subsets of Ω, will therefore be groups of individuals. In particular, a subset $\tau \subset \Omega$ will be called a *Galton–Watson tree* if:

1. $\emptyset \in \tau$,
2. if $u, v \in \Omega$, then $uv \in \tau$ implies $u \in \tau$,
3. for all $u \in \tau$, $uj \in \tau$ if and only if $1 \le j \le 2$; we are supposing that each particle produces only two offspring.

The set of all Galton–Watson trees will be called $\mathbb{T}$. Typically we use the name τ for a particular tree, and whenever possible we will use the letters u or v or w to refer to the labels in τ, which we may also refer to as *nodes of* τ or *individuals in* τ or just as *particles*. Note, we present the notation in this section generalized to cover random numbers of offspring, and the ideas used in this paper will readily adapt to such situations. However, for the *binary* branching Brownian motion that we consider, there is, of course, only one $\tau \in \mathbb{T}$ of interest — the binary tree.

Each individual should have a *location* in $\mathbb{R}$ at each moment of its *lifetime*. Since a Galton–Watson tree $\tau \in \mathbb{T}$ in itself can express only the *family* structure of the individuals in our branching random walk, in order to give them these extra features we suppose that each individual $u \in \tau$ has a *mark* (X_u, σ_u) associated with it which we read as:

- $\sigma_u \in \mathbb{R}^+$ is the *lifetime* of u, which determines the *fission time* of particle u as $S_u := \sum_{v \le u} \sigma_v$ (with $S_\emptyset := \sigma_\emptyset$). The times S_u may also be referred to as the *death* times;
- $X_u : [S_u - \sigma_u, S_u) \to \mathbb{R}$ gives the *location* of u at time $t \in [S_u - \sigma_u, S_u)$.

To avoid ambiguity, we must decide whether a particle is in existence or not at its death time:

Remark 2.1. Our convention throughout will be that a particle u dies 'just before' its death time S_u (which explains why we have defined X_u on the interval $[S_u - \sigma_u, S_u)$). Thus at the time S_u the particle u has *disappeared*, replaced by its 2 children which are both alive and ready to go.

We denote a single marked tree by (τ, X, σ) or (τ, M) for shorthand, and the set of all marked Galton–Watson trees by $\mathcal{T}$:

- $\mathcal{T} := \{(\tau, X, \sigma) : \tau \in \mathbb{T}$ and for each $u \in \tau, \sigma_u \in \mathbb{R}^+, X_u : [S_u - \sigma_u, S_u) \to \mathbb{R}\}$.
- For each $(\tau, X, \sigma) \in \mathcal{T}$, the set of particles that are alive at time t is defined as $N_t := \{u \in \tau : S_u - \sigma_u \le t < S_u\}$.

For any given marked tree $(\tau, M) \in \mathcal{T}$ we can identify distinguished lines of descent from the initial ancestor: $\emptyset, u_1, u_2, u_3, \ldots \in \tau$, in which u_3 is a child of u_2, which itself is a child of u_1 which is a child of the original ancestor $\emptyset$. We'll call such a subset of τ a *spine*, and will refer to it as ξ:

- a spine ξ is a subset of nodes $\{\emptyset, u_1, u_2, u_3, \ldots\}$ in the tree τ that make up a unique line of descent. We use ξ_t to refer to the unique node in ξ that is alive at time t.

In a more formal definition, which can for example be found in the paper by Liu and Rouault [14], a spine is thought of as a point on $\partial\tau$ the boundary of the tree — in fact the boundary is *defined* as the set of all infinite lines of descent. This explains the notation $\xi \in \partial\tau$ in the following definition: we augment the space $\mathcal{T}$ of marked trees to become

- $\tilde{\mathcal{T}} := \{(\tau, M, \xi) : (\tau, M) \in \mathcal{T}$ and $\xi \in \partial\tau\}$ is the set of *marked trees with distinguished spines*.

It is natural to speak of the *position of the spine at time t* which we think of just as the position of the unique node that is in the spine and alive at time t:

- we define the time-t position of the spine as $\xi_t := X_u(t)$, where $u \in \xi \cap N_t$.

By using the notation ξ_t to refer to *both* the node in the tree and that node's spatial position we are introducing potential ambiguity, but in practice the context will make clear which we intend. However, in case of needing emphasis, we shall give the node a longer name, writing:

- $\text{node}_t(\xi) = u$, if $u \in \xi$ is the node in the spine alive at time t.

As the spine ξ_t diffuses, at the fission times S_u for $u \in \xi$ it gives birth to some offspring, one of which continues the spine whilst the others go off to create sub-trees like copies of the BBM. These times on the spine are especially important for the later spine decomposition of the martingale Z_λ, and we therefore give them a name:

- the sequence of random times $\{S_u : u \in \xi\}$ are known as the *fission times on the spine*;

Finally, it will later be important to know how many fission times there have been in the spine, or what is the same, to know which generation of the family tree the node ξ_t is in (where the original ancestor Ø is considered to be the 0th generation).

Definition 2.2. We define the counting function

$$n_t = |\text{node}_t(\xi)|,$$

or equivalently,

$$n_t := |\{u : u \in \xi \text{ and } S_u \le t\}|,$$

which tells us which generation the spine node is in, or equivalently how many fission times there have been on the spine.

For example, if the node that corresponds to the spine at time t is $\emptyset 21 \in \xi$, representing the first child of the second child of the initial ancestor, we write $\text{node}_t(\xi) = \emptyset 21$ (or sometimes $\xi_t = \emptyset 21$), and then $n_t = |\text{node}_t(\xi)| = 2$ since there have been exactly two fissions along the spine's path by time t.

The collection of all marked trees with a distinguished spine $(\hat{\tau}, \xi)$ is given the label $\tilde{\mathcal{T}}$. On this space we define four filtrations of key importance that encapsulate different knowledge, but see Hardy and Harris [7] for more precise details:

- $\mathcal{F}_t$ knows everything that has happened to all the branching particles up to the time t, *but does not know which one is the spine*;
- $\tilde{\mathcal{F}}_t$ knows everything that $\mathcal{F}_t$ knows and also knows which line of descent is the spine (it is in fact the finest filtration);
- $\mathcal{G}_t$ knows only about the spine's motion on $\mathbb{R}$ up to time t, but does not actually know which line of descent in the family tree makes up the spine;

- $\tilde{\mathcal{G}}_t$ knows about the spine's motion and also knows which nodes it is composed of. Furthermore it knows about the fission times of these nodes and how many children were born at each time.

Having now defined the underlying space for our probabilities, we remind ourselves of the probability measures:

Definition 2.3. For each $x \in \mathbb{R}$, let P^x be the measure on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ such that the filtered probability space $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, P^{\cdot})$ makes the $\mathbb{R}$-valued point process $\mathbb{X}_t = \{X_u(t) : u \in N_t\}$ the canonical model for BBM.

For details of how the measures P^x are formally constructed on the underlying space of trees, we refer the reader to the work of Neveu [18] and Chauvin [3,1].

Our spine approach relies on building a measure $\tilde{P}^x$ under which the spine is a single genealogical line of descent chosen from the underlying tree. Starting the spine at the initial ancestor in a tree, at next birth event along the spine we will choose completely at random between the offspring for the one which continues to be the spine, and so on. Thus, if we are given a sample tree (τ, M) for the branching process it can be verified that the corresponding 'harmonic' choice of which line of descent makes up the spine ξ implies that if $u \in \tau$ then

$$\mathrm{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{2}. \tag{6}$$

This observation is the key to our method for extending the measures, and for this we make use of the following representation also used in Lyons [15].

Theorem 2.4. *If f is a $\tilde{\mathcal{F}}_t$-measurable function then we can write:*

$$f = \sum_{u \in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{7}$$

where f_u is $\mathcal{F}_t$-measurable.

We use this representation to extend the measures P^x.

Definition 2.5. Given the measure P^x on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ we extend it to the probability measure $\tilde{P}^x$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by defining

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P}^x := \int_{\mathcal{T}} \sum_{u \in N_t} f_u \prod_{v<u} \frac{1}{2} \, \mathrm{d}P^x, \tag{8}$$

for each $\tilde{\mathcal{F}}_t$-measurable f with representation like (7).

The previous approach to spines, exemplified in Lyons [15], used the idea of *fibres* to get a (non-probability) measure instead of our $\tilde{P}$ that could measure the spine. However, a weakness in this approach was that the corresponding measure had exponentially increasing total mass and could not simply be normalized to become a 'natural' probability measure. Our approach of using the down-weighting term of (6) in the definition of $\tilde{P}$ is crucial in ensuring that we have a probability measure with a very natural interpretation, which also leads to the very useful situation in which *all* measure changes in our formulation are carried out by *martingales*.

Theorem 2.6. *This measure $\tilde{P}^x$ really is an extension of P^x in that $P = \tilde{P}|_{\mathcal{F}_\infty}$.*

Proof. If f is $\mathcal{F}_t$-measurable then the representation (7) is trivial and therefore by definition

$$\int_{\tilde{\mathcal{T}}} f \,\mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \times \left(\sum_{u \in N_t} \prod_{v<u} \frac{1}{2} \right) \mathrm{d}P.$$

However, it can be shown that $\sum_{u \in N_t} \prod_{v<u} \frac{1}{2} = 1$ by retracing the sum back through the lines of ancestors to the original ancestor $\emptyset$, factoring out the product terms as each generation is passed. Thus

$$\int_{\tilde{\mathcal{T}}} f \,\mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \,\mathrm{d}P. \quad \square$$

The spine diffusion ξ_t is $\tilde{\mathcal{F}}_t$-measurable, and it is immediate that

Theorem 2.7. *Under $\tilde{P}^x$ the spine diffusion ξ_t is a Brownian motion that starts at x.*

'Many-to-One' results have previously been extremely useful in reducing expectation calculations of sums over the whole collection of branching particles to expectation calculations depending on just a single particle (the spine in our model), for example, see Harris and Williams [10]. We shall need to use the following 'Many-to-One' theorem in the next section; a proof of a more general result is given in Hardy and Harris [7].

Theorem 2.8 (*Many-to-One*). *If $g(t)$ is $\mathcal{G}_t$-measurable with the representation*

$$g(t) = \sum_{u \in N_t} g_u(t) \mathbf{1}_{(\xi_t = u)},$$

where $g_u(t)$ is $\mathcal{F}_t$-measurable, then

$$\mathrm{e}^{rt} \tilde{P}\,(g(t)) = P \left(\sum_{u \in N_t} g_u(t) \right).$$

Recall that r is the rate of binary fission in the branching Brownian motion so that on average there will be e^{rt} particles at time t. Further, each individual particle moves with the same law as the spine process $\{\xi_t\}_{t \ge 0}$ under $\tilde{P}$, so the *spine* can be thought of as representing the motion of a *'typical'* particle. Considering the special case when $g(t) = h(\xi_t) = \sum_{u \in N_t} h(X_u(t))\mathbf{1}_{(\xi_t = u)}$, the Many-to-One theorem gives $P\left(\sum_{u \in N_t} h(X_u(t))\right) = \tilde{P}\left(\mathrm{e}^{rt} h(\xi_t)\right)$. Intuitively, the expectation of a sum of contributions from *many* particles is simplified to the expectation of just *one* 'typical' contribution (from the spine) multiplied by the average number of such 'typical' particles. Note, the general 'Many-to-One' also covers spatially dependent branching rates, general offspring distributions, and so forth.

3. A local upper bound

Theorem 3.1. *Let $g \in C[0, 1]$. Then,*

$$\lim_{\delta \to 0} \limsup_{T \to \infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \le -S(g). \tag{9}$$

Proof. We first note that a monotonicity holds:

$$0 \geq \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \downarrow \quad \text{as } \delta \downarrow 0$$

and therefore the $\delta \to 0$ limit (9) exists (though it could potentially be $-\infty$).

The probability that a single particle has a path near g is smaller than the *expected number* of such particles:

$$P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \leq P\left(\sum_{u\in N_t} \mathbf{1}\left\{X_u^T \in \mathbb{B}_\delta(g)\right\}\right),$$

and an application of the Many-to-One Theorem 2.8 gives:

$$\begin{aligned} P\left(\sum_{u\in N_T} \mathbf{1}\left\{X_u^T \in \mathbb{B}_\delta(g)\right\}\right) &= \tilde{P}\left(\mathrm{e}^{rT}\mathbf{1}\left\{\xi^T \in \mathbb{B}_\delta(g)\right\}\right) \\ &= \mathrm{e}^{rT}\tilde{P}\left(\xi^T \in \mathbb{B}_\delta(g)\right). \end{aligned} \tag{10}$$

If it is the case that $g \notin C_0[0,1]$, so that its derivative is not square integrable, then from the simple fact that the open set $\mathbb{B}_\delta(g)$ is a subset of the closed δ-neighbourhood $\overline{\mathbb{B}_\delta(g)}$, we can use the above reasoning to deduce that

$$\begin{aligned} \limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) &\leq \limsup_{T\to\infty} T^{-1}\log P\left(\exists u \in N_T : X_u^T \in \overline{\mathbb{B}_\delta(g)}\right) \\ &\leq \limsup_{T\to\infty} T^{-1}\log \mathrm{e}^{rT}\tilde{P}\left(\xi^T \in \overline{\mathbb{B}_\delta(g)}\right), \end{aligned}$$

and an application of the upper bound in Schilder's theorem to the right-hand probability will give us the correct result:

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \leq r - \lim_{\delta\to 0}\inf_{g\in\overline{\mathbb{B}_\delta(g)}} I(g) = -\infty = -S(g).$$

Therefore we can assume that throughout the following proof we have the more interesting case of $g \in C_0[0,1]$.

The reasoning that gave (10) can immediately be strengthened by the simple observation that if the rescaled path is near g throughout the whole interval $[0,1]$, then it *must* be near g throughout all shorter intervals $[0,w]$, and a similar argument to the above would imply that for all $w \in [0,1]$,

$$\begin{aligned} P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) &\leq P\left(\exists u \in N_T : |X_u^T(s) - g(s)| < \delta, \forall s \in [0,w]\right) \\ &\leq P\left(\sum_{u\in N_T}\mathbf{1}\left\{\left|X_u^T(s) - g(s)\right| < \delta, \forall s\in[0,w]\right\}\right) \\ &= \mathrm{e}^{rwT}\tilde{P}\left(\left|\xi^T(s) - g(s)\right| < \delta, \forall s \in [0,w]\right). \end{aligned} \tag{11}$$

For $g \in C_0[0,1]$ it is clear that the supremum in the definition of the rate functional $S(g)$ will be reached at some point $\hat{w} \in [0,1]$, whence

$$S(g) = -r\hat{w} + \int_0^{\hat{w}} \frac{1}{2} g'(s)^2\,\mathrm{d}s.$$

Choosing $w = \hat{w}$ in (11) gives

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g))$$
$$\le r\hat{w} + \limsup_{T\to\infty} T^{-1} \log \tilde{P}\left(\left|\xi^T(s) - g(s)\right| < \delta, \forall s \in [0, \hat{w}]\right). \tag{12}$$

Since the spine diffusion ξ_t is just a Brownian motion, Schilder's theorem says that over the time interval $[0, \hat{w}]$, its rescaled path $\xi^T(s)$ will satisfy a large-deviations principle with rate functional $I^{\hat{w}}(g) := \frac{1}{2}\int_0^{\hat{w}} g'(s)^2\,\mathrm{d}s$, and therefore,

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1} \log \tilde{P}\left(\left|\xi^T(s) - g(s)\right| < \delta, \forall s \in [0, \hat{w}]\right) = -\frac{1}{2}\int_0^{\hat{w}} g'(s)^2\,\mathrm{d}s. \tag{13}$$

Our local upper bound for the BBM now follows directly from (12) and (13):

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \le r\hat{w} - \frac{1}{2}\int_0^{\hat{w}} g'(s)^2\,\mathrm{d}s = -S(g). \quad \square$$

4. The martingale Z_{g_T} for BBM

Let $g \in C_0[0, 1]$ be a fixed path; note that from here until Section 6 we must insist that the derivative g' is square integrable, since otherwise the change of measure martingales cannot be defined.

Given that the spine diffusion ξ_t is itself a $\tilde{P}$-Brownian motion, it follows that on the sub-filtration $(\mathcal{G}_t)_{0\le t\le T}$,

$$\zeta_{g_T}(t) := \exp\left\{\int_0^t g_T'(s)\,\mathrm{d}\xi_s - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s\right\}, \tag{14}$$

is a $\tilde{P}$-martingale, where

Definition 4.1. For any fixed $T \ge 0$ and any function $g \in C[0, 1]$ we define

$$g_T(s) := T g(s/T) \quad \forall s \in [0, T]$$

to be the *time-T scaled-up version of g*.

This martingale (14) is well known from the Girsanov theorem, and when used to change the measure it will introduce a drift to the Brownian motion.

Likewise, the process n_t from Definition 2.2 which counts the number of fission times on the spine up to time t is a Poisson process of rate r, therefore

$$t \mapsto \mathrm{e}^{-rt} 2^{n_t}$$

is a $\tilde{P}$-martingale too, which will increase the rate of n_t from r to $2r$ if used to change the measure — see also Kyprianou [12] in addition to Hardy and Harris [7].

We can use the product of these two martingales to define a new measure:

Theorem 4.2. *For each $T \geq 0$ we define a measure $\tilde{\mathbb{Q}}_T$ on the filtration $(\tilde{\mathcal{F}}_t)_{0\leq t\leq T}$ via*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = \mathrm{e}^{-rt}2^{n_t} \times \zeta_{g_T}(t). \tag{15}$$

Under the measure $\tilde{\mathbb{Q}}_T$ we can give a pathwise construction of the branching diffusion $\mathbb{X}_t$ over the time interval $t \in [0, T]$:

- *the spine diffusion $(\xi_t)_{0\leq t\leq T}$ starts at 0 and diffuses so $\xi_t - g_T(t)$ is a $\tilde{\mathbb{Q}}_T$-Brownian motion over the time interval $t \in [0, T]$;*
- *at rate $2r$ the spine undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected to continue the spine;*
- *the other particle initiates, from its birth position, an independent copy of a $P^{\cdot}$ branching Brownian motion with branching rate r.*

We briefly recall that for the third point above, the two particles produced are born at the same location and are therefore spatially indistinguishable, but due to our use of the Ulam–Harris labelling scheme they are distinguishable according to the label that they carry. Therefore the idea of choosing a particle is really a question of choosing between *labels*.

This change of measure gives us an additive martingale:

Definition 4.3. For each $T \geq 0$,

$$Z_{g_T}(t) := \mathrm{e}^{-rt} \sum_{u\in N_t} \mathrm{e}^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s},$$

defines an additive martingale on the filtered probability space $\left(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{0\leq t\leq T}\right)$.

Note, we use the term *additive* martingale in the sense that its value at any time is given by a *sum* of contributions from individual particles currently alive in the branching system. (Of course, *multiplicative* martingales also play an important role in branching systems.)

The fact that Z_{g_T} this is really a martingale is due to the following:

Theorem 4.4. *If we define $\mathbb{Q}_T := \tilde{\mathbb{Q}}_T|_{\mathcal{F}_T}$, then $\mathbb{Q}_T$ is a measure on the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_{g_T}(t).$$

Proof. It is clear from the definition of the conditional expectation that the change of measure (15) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation: for $t \in [0, T]$

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{P}}\right|_{\mathcal{F}_t} = \tilde{P}\left(\mathrm{e}^{-rt}2^{n_t}\zeta_{g_T}|\mathcal{F}_t\right).$$

Bearing in mind that $2^{n_t} = \prod_{v<\xi_t} 2$, if we use the representation (7) we get

$$\tilde{P}\left(\mathrm{e}^{-rt}2^{n_t}\zeta_{g_T}|\mathcal{F}_t\right) = \tilde{P}\left(\mathrm{e}^{-rt}\sum_{u\in N_t} \mathrm{e}^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \mathbf{1}_{(\xi_t=u)} \middle| \mathcal{F}_t\right)$$

$$= \mathrm{e}^{-rt} \sum_{u \in N_t} \mathrm{e}^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \tilde{P}\,(\xi_t = u|\mathcal{F}_t)$$

$$= \mathrm{e}^{-rt} \sum_{u \in N_t} \mathrm{e}^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \prod_{v<u} \frac{1}{2}$$

$$= \mathrm{e}^{-rt} \sum_{u \in N_t} \mathrm{e}^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} = Z_{g_T}(t).$$

We note that in the above we used $\tilde{P}\,(\xi_t = u|\mathcal{F}_t) = \prod_{v<u} \frac{1}{2}$, which follows easily from (6). □

Note, that we have constructed an additive martingale involving *all* the particles by a straightforward projection from a much simpler martingale that only involves a *single* particle's motion and birth process (the spine).

5. The growth of Z_{g_T} under $\mathbb{Q}_T$

As we shall see in Theorem 5.2 below, the rate at which Z_{g_T} grows under the measure $\mathbb{Q}_T$ is precisely the rate we need in the large deviations. In addition to Hardy and Harris [7], some related techniques developed in [8] were also helpful for this section.

Theorem 5.1. *For any $g \in C_0[0,1]$ and $\alpha \in [0,1]$, $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$-submartingale on $[0,T]$.*

Proof. If $\alpha \in [0,1]$ it follows from Jensen's inequality that $Z_{g_T}(t)^{1+\alpha}$ is a P-submartingale on $[0,T]$. Moreover, Theorem 4.4 revealed

$$\left.\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_{g_T}(t),$$

so it immediately follows that $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$-submartingale on $[0,T]$. □

Next, we will use the *spine decomposition* to get a good estimate of $\mathbb{Q}_T(Z_{g_T}(T)^\alpha)$ that we can use in Doob's submartingale inequality.

Theorem 5.2. *For each $g \in C_0[0,1]$ and for each $\alpha \in [0,1]$,*

$$\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right) \le \mathrm{e}^{\alpha S(g) T} \mathrm{e}^{\frac{1}{2}\alpha^2 T \int_0^1 g'(s)^2\,\mathrm{d}s} (1 + 2rT). \tag{16}$$

Proof. Since it is only the spine that is affected by the change of measure, the so-called *spine decomposition* in which we condition on knowing the spine's behaviour and fission times, is exceptionally useful for dealing with the P-martingale. We recall that the filtration $\tilde{\mathcal{G}}_\infty$ contains all information about the spine and the fission times S_u that occur along it, and therefore obtain the spine decomposition:

$$\tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)\,\middle|\,\tilde{\mathcal{G}}_\infty\right) = \mathrm{e}^{-rT}\mathrm{e}^{\int_0^T g_T'(s)\,\mathrm{d}\xi_s - \frac{1}{2}\int_0^T g_T'(s)^2\,\mathrm{d}s} + \sum_{u<\xi_T} \mathrm{e}^{-rS_u}\mathrm{e}^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\xi_s - \frac{1}{2}\int_0^{S_u} g_T'(s)^2\,\mathrm{d}s}. \tag{17}$$

A proof of this can be found in Hardy and Harris [7], but the intuition relies only on the idea that under $\mathbb{Q}_T$ the sub-trees that leave the spine behave as if under the original measure P for which Z_{g_T} is a martingale.

By definition, $\hat{\xi}_s := \xi_s - g_T(s)$ for $0 \le s \le T$ is a Brownian motion under $\tilde{\mathbb{Q}}_T$, and substituting

$$\mathrm{d}\xi_s = \mathrm{d}\hat{\xi}_s + g_T'(s)\,\mathrm{d}s, \tag{18}$$

into (17) we arrive at:

$$\begin{aligned}
\tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)\,\middle|\,\tilde{\mathcal{G}}_\infty\right) &= \mathrm{e}^{\frac{1}{2}\int_0^T g_T'(s)^2\,\mathrm{d}s - rT}\mathrm{e}^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T}\mathrm{e}^{\frac{1}{2}\int_0^{S_u} g_T'(s)^2\,\mathrm{d}s - rS_u}\mathrm{e}^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\\
&= \mathrm{e}^{\left(\int_0^1\{\frac{1}{2}g'(s)^2 - r\}\,\mathrm{d}s\right)T}\mathrm{e}^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s}\\
&\quad + \sum_{u<\xi_T}\mathrm{e}^{\left(\int_0^{S_u/T}\{\frac{1}{2}g'(s)^2 - r\}\,\mathrm{d}s\right)T}\mathrm{e}^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\\
&\le \mathrm{e}^{\left(\sup_{w\in[0,1]}\int_0^w\{\frac{1}{2}g'(s)^2 - r\}\,\mathrm{d}s\right)T}\left(\mathrm{e}^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T}\mathrm{e}^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right),\\
&= \mathrm{e}^{S(g)T}\left(\mathrm{e}^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T}\mathrm{e}^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right).
\end{aligned} \tag{19}$$

In the above we note that $g_T'(s) = g'(s/T)$.

From the tower property, and since $\mathbb{Q}_T$ and $\tilde{\mathbb{Q}}_T$ agree on $\mathcal{F}_T$,

$$\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right) = \tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)^\alpha\right) = \tilde{\mathbb{Q}}_T\left(\tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)^\alpha\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right),$$

and the conditional form of Jensen's inequality says that for $\alpha \in [0, 1]$,

$$\tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)^\alpha\,\middle|\,\tilde{\mathcal{G}}_\infty\right) \le \tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)\,\middle|\,\tilde{\mathcal{G}}_\infty\right)^\alpha.$$

Since the spine decomposition $\tilde{\mathbb{Q}}_T\left(Z_{g_T}(T)\,\middle|\,\tilde{\mathcal{G}}_\infty\right)$ is a sum, we can use the following result noted by Neveu [19].

Proposition 5.3. *If $\alpha \in (0, 1]$ and $u, v > 0$ then $(u + v)^\alpha \le u^\alpha + v^\alpha$.*

Combining these observations with (19) leads to

$$\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right) \le \mathrm{e}^{\alpha S(g)T}\,\tilde{\mathbb{Q}}_T\left(\mathrm{e}^{\alpha\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T}\mathrm{e}^{\alpha\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right). \tag{20}$$

Under the measure $\tilde{\mathbb{Q}}_T$, the process $(\hat{\xi}_t)_{0\le t\le T}$ is a standard Brownian motion, and therefore

$$\mathrm{e}^{\alpha\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s - \frac{1}{2}\alpha^2\int_0^T g_T'(s)^2\,\mathrm{d}s}$$

is a $\tilde{\mathbb{Q}}_T$-martingale on $t \in [0, T]$. Evaluating this at the bounded stopping times $(S_u : u < \xi_T)$ gives

$$\tilde{\mathbb{Q}}_T\left(\mathrm{e}^{\alpha\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right) = \tilde{\mathbb{Q}}_T\left(\mathrm{e}^{\frac{1}{2}\alpha^2\int_0^{S_u} g_T'(s)^2\,\mathrm{d}s}\right) \le \mathrm{e}^{\frac{1}{2}\alpha^2\int_0^T g_T'(s)^2\,\mathrm{d}s},$$

whence from (20) we obtain

$$\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right) \le \mathrm{e}^{\alpha S(g)T}\mathrm{e}^{\frac{1}{2}\alpha^2\int_0^T g_T'(s)^2\,\mathrm{d}s}\,\tilde{\mathbb{Q}}_T\left(1 + n_T\right).$$

We know that under the measure $\tilde{\mathbb{Q}}_T$ the births on the spine occur as a Poisson process with rate $2r$, whence the expectation grows linearly in T, $\tilde{\mathbb{Q}}_T\,(1+n_T) = 1 + 2rT$, and we arrive at

$$\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right) \le \mathrm{e}^{\alpha S(g)T}\mathrm{e}^{\frac{1}{2}\alpha^2 T\int_0^1 g'(s)^2\,\mathrm{d}s}\,(1+2rT). \quad \square$$

Having established that the martingale grows at the rate we expect, we can prove the following result that is the key to the large-deviations lower bound.

Theorem 5.4. *For each $\varepsilon > 0$,*

$$\mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) \le \mathrm{e}^{(S(g)+\varepsilon)T}\right) \to 1, \quad \text{as } T \to \infty. \tag{21}$$

Proof. Recall, Theorem 5.1 revealed that, for each $\alpha \in [0, 1]$, $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$ submartingale on $t \in [0, T]$ and we can now prove a probability bound on its growth by combining the estimate (16) with Doob's submartingale inequality. For any small $\varepsilon > 0$ and for any fixed $T > 0$,

$$\begin{aligned}\mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) > \mathrm{e}^{(S(g)+\varepsilon)T}\right) &= \mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s)^\alpha > \mathrm{e}^{\alpha(S(g)+\varepsilon)T}\right)\\ &\le \frac{\mathbb{Q}_T\left(Z_{g_T}(T)^\alpha\right)}{\mathrm{e}^{\alpha(S(g)+\varepsilon)T}}.\end{aligned}$$

Using (16) this gives

$$\mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) > \mathrm{e}^{(S(g)+\varepsilon)T}\right) \le \mathrm{e}^{\left(\alpha\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s-\varepsilon\right)\alpha T}\,(1+2rT).$$

Bearing in mind that $\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s$ is just a finite number, we can choose $\alpha > 0$ small enough so that $\alpha\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s - \varepsilon < 0$, whence the exponential decay dominates the linear growth in the above, and we have proven that

$$\mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) > \mathrm{e}^{(S(g)+\varepsilon)T}\right) \to 0, \quad \text{as } T \to \infty. \quad \square$$

6. A local lower bound

We note that in the case of the lim inf we do not deal immediately with the limit as $\delta \to 0$, since without the monotonicity that we had for the lim sup we do not a priori know that the limit exists.

Theorem 6.1. *Let $g \in C[0, 1]$. For any fixed $\delta > 0$, we have*

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \ge -S(g). \tag{22}$$

Proof. First we note that if g is not in $C_0[0, 1]$ then $S(g) = \infty$ and the result holds trivially. Therefore we assume that throughout we have $g \in C_0[0, 1]$.

Importantly, the event we are considering is $\mathcal{F}_T$-measurable, and on this algebra the change of measure is carried out by Z_{g_T}, as stated in Theorem 4.4. Therefore,

$$P\left(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\right) = \mathbb{Q}_T\left(\frac{1}{Z_{g_T}(T)}; \exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\right). \tag{23}$$

The upper bound that we have derived for Z_{g_T} will now serve as a lower bound for $1/Z_{g_T}(T)$, so that for any $\varepsilon > 0$,

$$\begin{aligned}
&P\left(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\right)\\
&\quad\geq \mathrm{e}^{-(S(g)+\varepsilon)T}\mathbb{Q}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) \leq \mathrm{e}^{(S(g)+\varepsilon)T}; \exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\right)\\
&\quad\geq \mathrm{e}^{-(S(g)+\varepsilon)T}\tilde{\mathbb{Q}}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) \leq \mathrm{e}^{(S(g)+\varepsilon)T}; \xi^T \in \mathbb{B}_\delta(g)\right).
\end{aligned} \tag{24}$$

Since $\xi_s^T - g(s)$ is a $\tilde{\mathbb{Q}}_T$-Brownian motion on $[0, 1]$ with diffusion coefficient $1/\sqrt{T}$, it follows that

$$\tilde{\mathbb{Q}}_T\left(\xi^T \in \mathbb{B}_\delta(g)\right) \to 1, \quad \text{as } T \to \infty,$$

and this combines with the result of Theorem 5.4 to give:

$$\tilde{\mathbb{Q}}_T\left(\sup_{s\in[0,T]} Z_{g_T}(s) \leq \mathrm{e}^{(S(g)+\varepsilon)T};\ \xi^T \in \mathbb{B}_\delta(g)\right) \to 1, \quad \text{as } T \to \infty.$$

Thus from (24) we have

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \geq -S(g) - \varepsilon$$

which proves (22) since ε was arbitrary. $\square$

Corollary 6.2. *For each $g \in C[0, 1]$ we have*

$$\lim_{\delta\to 0}\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g). \tag{25}$$

Proof. The case where $g \notin C_0[0, 1]$, so $S(g) = +\infty$, is immediate from the upper bound of Theorem 3.1, and therefore we assume that $g \in C_0[0, 1]$. Using Theorem 3.1, for each $\delta > 0$ we can choose $\varepsilon_\delta > 0$ such that $\varepsilon_\delta \to 0$ as $\delta \to 0$ and

$$\begin{aligned}
-S(g) + \varepsilon_\delta &> \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g))\\
&\geq \liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \geq -S(g),
\end{aligned}$$

where we have also used a trivial inequality between the limsup and liminf combined with the lower bound of (22). Hence, the $\delta \to 0$ limit exists for the lim inf as required. $\square$

Together with Theorem 3.1 we have now completed the proof of the local limit result Theorem 1.7.

7. Improving the 'weak' large-deviations result

As mentioned, the local results of Theorem 1.7 and the fact that the δ-neighbourhoods $\mathbb{B}_\delta(g)$ form a base for the topology of $C[0,1]$ means that we have at least a *weak* large-deviations result: the lower bound of Theorem 1.5 holds, but the upper bound is proven only for *compact* sets (as opposed to *closed* sets). The main ideas for the following proof of this come from Theorem 4.1.11 of Dembo and Zeitouni [4] where, as discussed with the comments following Theorem 1.5, we need to take particular care with the weakening of the additivity property to mere sub-additivity.

Theorem 7.1. *The local results of Theorem 1.7 imply that the upper bound of our main result Theorem 1.5 holds certainly for all $C \subset C[0,1]$ that are closed and compact:*

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \le -\inf_{g\in C} S(g),$$

whilst the lower bound holds in full for all open subsets $V \subset C[0,1]$:

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in V) \ge -\inf_{g\in V} S(g).$$

Proof. First of all we consider the lower bound. The open δ-neighbourhoods $\{\mathbb{B}_\delta(g) : g \in C[0,1], \delta > 0\}$ form a base for the topology of $C[0,1]$ which we shall call $\mathcal{A}$. Therefore if $V \subset C[0,1]$ is an open set then for each $g \in V$ we can be sure that for some small enough $\delta > 0$ we shall have $g \in \mathbb{B}_\delta(g) \subset V$, and therefore

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in V) \ge \liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)).$$

Furthermore, since the δ-neighbourhoods sit inside one another as $\delta \to 0$ we actually have a limit result which combines with result (25) to say that for each $g \in V$:

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in V) \ge \lim_{\delta\to 0} \liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g).$$

Since this holds for all $g \in V$ it will hold for the supremum:

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in V) \ge \sup_{g\in V} -S(g) = -\inf_{g\in V} S(g),$$

concluding the proof of the lower bound.

For the upper bound we use a finite-covering argument: supposing that $C \subset C[0,1]$ is closed and compact we shall cover it with a finite number of open sets from $\mathcal{A}$ to deduce the result. Eq. (9) states that

$$\lim_{\delta\to 0} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g).$$

Since each open set $A \in \mathcal{A}$ contains at least one δ-neighbourhood, this result implies that

$$\inf_{\{A\in\mathcal{A}, g\in A\}} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) = -S(g),$$

and for the following argument we rearrange this as

$$\sup_{\{A\in\mathcal{A}, g\in A\}} \left[-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) \right] = S(g). \tag{26}$$

If we choose and fix $\delta > 0$ and define

$$S^{\delta}(g) := \min\{S(g) - \delta, 1/\delta\},$$

then for each $g \in C[0, 1]$ the above (26) implies that there is some open set $A_g \in \mathcal{A}$ (which may depend on δ) such that

$$-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A_g) \geq S^{\delta}(g). \tag{27}$$

For the given compact set C we can extract a finite cover from the covering $\bigcup_{g\in C} A_g$, which we denote $\{A_{g_1}, \ldots, A_{g_n}\}$; then by the *sub-additivity property* discussed earlier in this chapter we have

$$P(\exists u \in N_T : X_u^T \in C) \leq \sum_{i=1}^{n} P(\exists u \in N_T : X_u^T \in A_{g_i}). \tag{28}$$

Here we are dealing with a finite sum and can use a standard result of Laplace on the growth rate of finite sums of exponentials (see Dembo and Zeitouni's Lemma 1.2.15):

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in C) \leq \max_{i=1,\ldots,n} \limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A_{g_i}).$$

From (27) we have

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A_{g_i}) \leq -S^{\delta}(g_i)$$

and therefore

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \leq -\min_{i=1,\ldots,n} S^{\delta}(g_i) \leq -\inf_{g\in C} S^{\delta}(g).$$

The proof of the upper bound is completed by considering the limit as $\delta \to 0$. □

It is clear how the compactness property was the connecting link between the local properties of Theorem 1.7 and the above *weak* large-deviations result. We now wish to improve this to get the full large-deviations result of Theorem 1.5, and a standard approach here is to use *exponential tightness* of measures. This approach is particularly suitable for spines since the question of exponential tightness of the BBM probabilities can be reduced to that of the single Brownian-motion probabilities using the Many-to-One Theorem 2.8.

Definition 7.2. A family of probability measures $\{\mu_T\}$ on a set $\mathcal{X}$ is said to be exponentially tight if for each $\alpha < \infty$ there exists a compact $K \subset \mathcal{X}$ such that

$$\limsup_{T\to\infty} T^{-1} \log \mu_T(K^{\complement}) < -\alpha,$$

where $K^{\complement}$ denotes the set complement.

We recall without proof a standard result that on the set of paths $C[0, 1]$, the measures

$$P(\xi_T \in A), \quad \text{for } A \subset C[0, 1],$$

concerning the paths of a single Brownian motion ξ_T are exponentially tight (see, for example, Dembo and Zeitouni [4], p. 120).

Theorem 7.3. *The fact that the above path measures of a single Brownian motion are exponentially tight implies that the (sub-additive) measures* $\left\{P(\exists u \in N_T : X_u^T \in \cdot)\right\}_{T\geq 0}$ *for the branching Brownian motion are also exponentially tight.*

Proof. For any set $K \subset C[0,1]$, we have an expectation bound that combines with the Many-to-One property to give:

$$P(\exists u \in N_T : X_u^T \in K^{\complement}) \leq P\left(\sum_{u\in N_T} \mathbf{1}\left\{X_u^T \in K^{\complement}\right\}\right) = \mathrm{e}^{rT} P(\xi^T \in K^{\complement}),$$

whence

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) \leq r + \limsup_{T\to\infty} T^{-1} \log P(\xi^T \in K^{\complement}). \tag{29}$$

Let $\alpha < \infty$ be given. Since the spine is a Brownian motion, for which it is known that the probabilities $P(\xi^T \in \cdot)$ are exponentially tight, we can find some compact $K \subset C[0,1]$ such that

$$\limsup_{T\to\infty} T^{-1} \log P(\xi^T \in K^{\complement}) < -r - \alpha,$$

and therefore from (29),

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) < -\alpha. \quad \square$$

Dembo and Zeitouni [4] state at Lemma 1.2.18, that when an exponentially tight family of measures satisfy a weak large-deviation principle (LDP) then in fact the LDP holds in full. For completeness we modify this result as it applies to our particular context:

Theorem 7.4. *The weak large-deviation result of Theorem 7.1 together with the exponential tightness proven in Theorem 7.3 imply that the large-deviations result Theorem 1.5 of this chapter holds in full.*

Proof. The lower bound of Theorem 1.5 is exactly the same as that proven in the weak version of Theorem 7.1 and therefore here we are only looking to extend the upper bound of Theorem 7.1 for closed and compact sets to hold for the large class of all closed sets. That is, we want to show that for each closed subset $C \subset C[0,1]$ we have

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \leq -\inf_{g\in C} S(g).$$

From the fact that we have exponential tightness of the probabilities, we know that there is some compact subset $K \subset C[0,1]$ such that

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) < -\inf_{g\in C} S(g); \tag{30}$$

we point out that we have just taken $\alpha = \inf_{g\in C} S(g)$ in Definition 7.2. The covering of C as

$$C = (C \cap K) \cup (C \cap K^{\complement}) \subset (C \cap K) \cup K^{\complement},$$

together with the *sub-additivity property* of our probabilities gives

$$P(\exists u \in N_T : X_u^T \in C) \leq P(\exists u \in N_T : X_u^T \in C \cap K) + P(\exists u \in N_T : X_u^T \in K^{\complement}). \tag{31}$$

Now we are in a position to apply the 'weak' upper bound of Theorem 7.1 to the *compact* set $C \cap K$ to obtain

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C \cap K) \le -\inf_{g\in C\cap K} S(g) \le -\inf_{g\in C} S(g).$$

Applying the simple Laplace bound (Dembo and Zeitouni Lemma 1.2.15 as mentioned previously) to this above and (30) we obtain the desired result:

$$\begin{aligned}
&\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \\
&\quad \le \max \left\{ \begin{array}{c} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C \cap K) \\ \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) \end{array} \right\} \\
&\quad \le -\inf_{g\in C} S(g). \quad \square
\end{aligned}$$

Thus we conclude that the full large-deviation result holds, and the proof of Theorem 1.5 is concluded. We however include a final brief section on the rate function $S(g)$ for BBM.

8. A brief discussion of $S(g)$

$S(g)$ has been defined earlier as

$$S(g) := \begin{cases} \sup_{w\in[0,1]} \left(\int_0^w \frac{1}{2} g'(s)^2 - r \, \mathrm{d}s \right) & \text{if } g \in C_0[0,1], \\ \infty & \text{otherwise.} \end{cases}$$

In this section we give short proofs that this $S(g)$ is actually a so-called *good rate function*. Such facts were proven in Lee's [13] by arguments involving more heavy analytic estimates.

Theorem 8.1. *$S(g)$ is a good rate function. This is to say two things:*

- *it is a rate function, which is defined as stating that it is non-negative and that its level sets $\{g \in C[0,1] : S(g) \le \alpha\}$ are closed subsets of $C[0,1]$ for each α;*
- *it is a good rate function, which means that its level sets are actually compact subsets of $C[0,1]$.*

The ideas used in the following come from Dembo and Zeitouni [4].

Proof. First of all we trivially have $S(g) \ge 0$ for all $g \in C[0,1]$. Let $\alpha \ge 0$ be given. If $S(g) > \alpha$ then from (26) which says

$$S(g) = \sup_{\{A\in\mathcal{A}, g\in A\}} \left[-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) \right]$$

it follows that there is an open neighbourhood $A \in \mathcal{A}$ of g for which

$$-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) > \alpha.$$

This would therefore imply from (26) that for each $f \in A$ we also have

$$S(f) \ge -\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) > \alpha,$$

which is to say that $S(f) > \alpha$ for all $f \in A$. Hence the set $\{g : S(g) > \alpha\}$ is open and S really is a rate function.

To prove that S is a good rate function we use the lower bound of our large-deviations result together with the property of exponential tightness which says that for any given α there is some compact subset $K \subset C[0, 1]$ for which

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) < -\alpha.$$

At the same time, since $K^{\complement}$ is an open set we can apply the lower bound to get

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) \geq -\inf_{g\in K^{\complement}} S(g).$$

These two together force

$$\inf_{g\in K^{\complement}} S(g) > \alpha,$$

from which we deduce that $\{g : S(g) \leq \alpha\} \subset K$. The level set is therefore a closed subset of a compact set K, whence it too is compact, and whence S is a good rate function. $\square$

Acknowledgements

We thank the referees for their careful reading and suggestions. R. Hardy was supported by an EPSRC research studentship.

References

[1] Brigitte Chauvin, Arbres et processus de Bellman–Harris, Ann. Inst. H. Poincaré Probab. Statist. 22 (2) (1986) 209–232.
[2] Brigitte Chauvin, Alain Rouault, Supercritical branching Brownian motion and K-P-P equation in the critical speed-area, Math. Nachr. 149 (1990) 41–59.
[3] Brigitte Chauvin, Product martingales and stopping lines for branching Brownian motion, Ann. Probab. 19 (3) (1991) 1195–1205.
[4] Amir Dembo, Ofer Zeitouni, Large Deviations Techniques and Applications, Springer, 1998.
[5] Yoav Git, Almost Sure Path Properties of Branching Diffusion Processes, in: Séminaire de Probabilités, vol. XXXII, Springer, 1998.
[6] Robert Hardy, Simon C. Harris, A spine approach to branching diffusions with applications to martingale convergence, 2006 (submitted for publication).
[7] Robert Hardy, Simon C. Harris, A new formulation of the spine approach to branching diffusions, 2004, no. 0404, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/˜massch/Research/Papers/spine-foundations.pdf.
[8] Robert Hardy, Simon C. Harris, Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales, 2004, no. 0405, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/˜massch/Research/Papers/spine-Lp-cgce.pdf.
[9] Robert Hardy, Simon C. Harris, A spine proof of a lower-bound for a typed branching diffusion, 2004, no. 0408, Mathematics Preprint, University of Bath. http://www.bath.ac.uk/˜massch/Research/Papers/spine-oubbm.pdf.
[10] Simon C. Harris, David Williams, Large deviations and martingales for a typed branching diffusion. I, Astérisque (236) (1996) 133–154. Hommage à P. A. Meyer et J. Neveu.
[11] Thomas Kurtz, Russell Lyons, Robin Pemantle, Yuval Peres, A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), in: IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185.
[12] Andreas Kyprianou, Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris's probabilistic analysis, Ann. Inst. H. Poincaré Probab. Statist. 40 (1) (2004) 53–72.
[13] Tzong-Yow Lee, Some large-deviation theorems for branching diffusions, Ann. Probab. 20 (3) (1992) 1288–1309.

[14] Quansheng Liu, Alain Rouault, On two measures defined on the boundary of a branching tree, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), in: IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 187–201.

[15] Russell Lyons, A simple path to Biggins' martingale convergence for branching random walk, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), in: IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221.

[16] Russell Lyons, Robin Pemantle, Yuval Peres, Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes, Ann. Probab. 23 (3) (1995) 1125–1138.

[17] Henry P. McKean, Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskounov, Comm. Pure Appl. Math. 28 (1975) 323–331.

[18] Jacques Neveu, Arbres et processus de Galton–Watson, Ann. Inst. H. Poincaré Probab. Statist. 22 (2) (1986) 199–207.

[19] Jacques Neveu, Multiplicative martingales for spatial branching processes, in: E. Çinlar, K.L. Chung, R.K. Getoor (Eds.), Seminar on Stochastic Processes, Birkhäuser, 1987, pp. 223–241.

[20] Srinivasa R.S. Varadhan, Large Deviations and Applications, SIAM, 1984.