A spine proof of a large-deviations principle for branching Brownian motion

Robert Hardy and Simon C. Harris

March 8, 2004

Abstract

Using the foundations for the use of spines in branching diffusions recently developed in Hardy and Harris [5, 7], we give a new, intuitive and relatively straightforward proof of a path large-deviations principle for branching Brownian motion (BBM) that can be thought of to extend Schilder's theorem for a single Brownian motion. We show how certain additive martingales for BBM can easily be defined in terms of martingales for a single Brownian motion and how these new martingales can be used to change measure so that the spine follows 'close' to the path of interest. We then use the spine decomposition to get an important upper bound on their exponential growth of these martingales under the new measures – this upper bound will then quickly lead to the required large-deviations lower bound. The (easier) large-deviation upper bound is proven using a more familiar 'many-to-one' expectation result that reduces the BBM calculation down to that for a single Brownian motion (the spine). Some of the techniques developed here will also apply in more general branching Markov processes, for example, see Hardy and Harris [6].

1 Overview

Suppose that under a measure $\tilde{P}$ the process $(\xi_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}$. The large-deviations behaviour of ξ_t is controlled by Schilder's theorem, and in order to state this we first define a re-scaling 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 re-scaled path.

Such a scaling of the path is effectively equivalent to supposing that ξ^T is a Brownian motion on $[0,1]$ with variance $1/\sqrt{T}$, and in fact Varadhan's proof of Schilder's theorem (see [17]) deals instead with a Brownian motion on $[0,1]$ whose diffusion coefficient is $\varepsilon > 0$ under a measure P_ε; thus he considers $\varepsilon \to 0$ rather than $T \to \infty$ and we could say heuristically that his ε is our $1/\sqrt{T}$. In our choice of the re-scaling approach we are following the large-deviations work of Git [4].

We use the label $C[0,1]$ to refer to the set of all continuous functions on $[0,1]$, and without losing generality we can suppose that under $\tilde{P}$ the Brownian motion starts at the origin.

Theorem 1.2 (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

$$I(g) := \int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s$$

if $g\in C[0,1]$ with $g(0)=0$ has a square-integrable derivative; otherwise we define $I(g)=\infty$.

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 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 $\mathbb{R}$-valued 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 re-scaling of the paths; we note that a particle u is born at the time $S_u-\sigma_u$, and 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 1.3 *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 re-scaled path*:*

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

Definition 1.4 *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.*

In this article we are going to prove the following theorem concerning the probability that the path of *at least one* of the many particles in the branching diffusion.

Theorem 1.5 *There is a large-deviation principle for 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

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

Note: For the special case when $r=0$ we get a single BM and our result becomes Schilder's theorem.

The BBM large-deviations principle (LDP) was proven by Tzong-Yow Lee [11], where he relied heavily on Friedlin's previous work on rescalings of solutions of reaction-diffusion equations. Our proofs are based on spines, and offer much clearer, neater and independent proofs that can be generalized to cover many different types of branching diffusions – in Hardy and Harris [6] we develop the ideas to deal with the typed branching diffusion originally studied in Harris and Williams [8].

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 a much stronger result has been proven by Git [4] 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*.

1.0.1 Outline of proof

As far as the topological issues in our arguments are concerned, the main reference is Dembo and Zeitouni [3]. It is known that the δ-neighbourhoods make up a base for the topology of $C[0,1]$ induced by the metric $\|f\| := \sup_{w\in[0,1]}\big|f(w)\big|$.

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

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

as the δ-neighbourhood around the function g.

We aim to prove Theorem 1.5 in two stages. First we use spine techniques to prove the following local results:

Theorem 1.7 *For any fixed $g \in C_0[0,1]$ we have*

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

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

For the local upper bound (3) we use the Many-to-One theorem to reduce the question to just the spine and then use Schilder's theorem. Our proof of the local lower bound (4) is based on getting an upper-bound for a new additive martingale for BBM, and we can illustrate the principle behind this simple idea: suppose that for some $\mathcal{F}_t$-measurable martingale $Z(t)$ we have a measure $\mathbb{Q}$ defined via $\mathrm{d}\mathbb{Q}/\mathrm{d}P = Z(t)$ on $\mathcal{F}_t$; this means that for a set $F \in \mathcal{F}_t$

$$P(F) = \mathbb{Q}\Big(\frac{1}{Z(t)}; F\Big),$$

and therefore an upper-bound on the growth of $Z(t)$ under $\mathbb{Q}$ (where it is a submartingale as we shall see) will here give us a lower bound for the probability $P(F)$. The main work of the local lower bound is therefore to define the correct martingale to give the appropriate new measure and to obtain a suitable upper-bound for it under this new measure – here the techniques also developed in Hardy and Harris [7] are very useful.

A topological-type theorem from Dembo and Zeitouni means that these two local results imply the existence of (at worst) a *weak* LDP for the probabilities of Theorem 1.5, which is to say that 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*. Once Theorem 1.7 has been proved, in section 7 we use the Many-to-One theorem with the concept of *exponential tightness* for a single Brownian motion (the spine) to improve these local results to the full large-deviations principle of Theorem 1.5.

2 The spine-approach foundations

Before moving on to the proofs, we briefly review the formal constructions on which our spine analysis is based – full details are laid out in foundation article [5]. The reader who is familiar with the work of Lyons *et al* [13, 9, 14], or with Kyprianou's paper [10] 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, ie. 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.*

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, it is always necessary to 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 $X_u : [S_u - \sigma_u, S_u) \to \cdot$ for example). 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 $\mathcal{T}$:

- $\mathcal{T} := \Big\{(\tau, X, \sigma) : \tau \in \mathbb{T} \text{ and for each } u \in \tau, \sigma_u \in \mathbb{R}^+, X_u : [S_u - \sigma_u, S_u) \to \mathbb{R}\Big\}.$
- 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 that is alive at time t.

In a more formal definition, which can for example be found in the paper by Rouault and Liu [12], 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

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

It is natural to speak of the *position of the spine at time* t which 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 to emphasize, we shall give the node a longer name:

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

which may also be written as $\text{node}_t(\xi)$.

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 subtrees 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 $\emptyset$ is considered to be the 0th generation)

Definition 2.2 *We define the counting function*

$$n_t = \big|node_t(\xi)\big|,$$

or equivalently,

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

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 $\xi_t = (\emptyset, u_1, u_2)$ *then both* $\emptyset$ *and* u_1 *have died and so* $n_t = 2$.

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 [5] 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 in J 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 [15] and Chauvin [2, 1].

All spine approaches rely on building a measure $\tilde{P}^x$ under which the spine is a single genealogical line of descent chosen uniformly from the underlying tree. If we are given a sample tree (τ, M) for the branching process it can be verified that a uniform choice of which line of descent makes up the spine ξ implies that if $u \in \tau$ then

$$\mathrm{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{2}. \tag{5}$$

This observation is the key to our method for extending the measures, and for this we make use of the following representation found in Lyons [13].

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

$$f = \sum_{u\in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{6}$$

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*

$$\int_{\tilde{\mathcal{T}}} f \,\mathrm{d}\tilde{P}^x := \int_{\mathcal{T}} \sum_{u\in N_t} f_u \prod_{v<u} \frac{1}{2} \,\mathrm{d}P^x, \tag{7}$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (6)*.*

The previous approach to spines, exemplified in Lyons [13], used the idea of *fibres* to get a measure analogous to our $\tilde{P}$ that could measure the spine. However, a weakness in this approach was that the corresponding measure did not have a finite mass and therefore could not be normalized to become a probability measure like our $\tilde{P}$. Our new idea of using the down-weighting term of (5) in the definition of $\tilde{P}$ is crucial in ensuring that we do not get an infinite-mass measure, and 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 $f \in m\mathcal{F}_t$ then the representation (6) is trivial and therefore by definition

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

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

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

□

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.*

In Harris and Williams [8] they used a many-to-one lemma to reduce additive expectation calculations of the whole collection of branching particles to an expectation calculation depending on just a single particle (the spine in our model). We shall need to use this theorem in the following section and state it here; a proof is given in Hardy and Harris [5].

Theorem 2.8 (Many-to-One) *If $g(t) \in m\mathcal{G}_t$ has the representation*

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

where $g_u(t) \in m\mathcal{F}_t$, then

$$e^{rt}\tilde{P}\big(g(t)\big) = P\Big(\sum_{u\in N_t} g_u(t)\Big).$$

3 A local upper bound

Theorem 3.1 *Let $g \in C_0[0,1]$. Then,*

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in B_\delta(g)) \le -S(g). \tag{8}$$

Proof: We first note that a monotonocity holds:

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

and therefore the $\delta \to 0$ limit (8) exists (though it could potentially be $-\infty$). We also point out that the supremum in the definition of the rate functional $S(g)$ will be reached at some point $\hat{w} \in [0,1]$, whence

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

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

$$P(\exists u \in N_T : X_u^T \in \mathrm{B}_\delta(g)) \le P\Big(\sum_{u\in N_t} \mathbf{1}\{X_u^T \in \mathrm{B}_\delta(g)\}\Big),$$

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

$$P\Big(\sum_{u\in N_T}\mathbf{1}\big\{X_u^T\in \mathrm{B}_\delta(g)\big\}\Big)=\tilde P\big(e^{rT}\mathbf{1}\{\xi^T\in \mathrm{B}_\delta(g)\}\big)$$
$$=e^{rT}\tilde P\big(\xi^T\in \mathrm{B}_\delta(g)\big).$$

In fact, this result 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 $w\in[0,1]$,

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

In particular, choosing $w=\hat w$ gives

$$\begin{aligned}&\limsup_{T\to\infty}T^{-1}\log P(\exists u\in N_T : X_u^T\in \mathrm{B}_\delta(g))\\ &\qquad\le r\hat w+\limsup_{T\to\infty}T^{-1}\log\tilde P\big(\big|\xi^T(s)-g(s)\big|<\delta,\forall s\in[0,\hat w]\big).\end{aligned}\tag{9}$$

Since the spine diffusion ξ_t is just a Brownian motion, Schilder's theorem says that over the time interval $[0,\hat w]$, its re-scaled path $\xi^T(s)$ will satisfy a large-deviations principle with rate functional $I^{\hat w}(g):=\frac12\int_0^{\hat w}g'(s)^2\,\mathrm{d}s$, and therefore,

$$\lim_{\delta\to0}\limsup_{T\to\infty}T^{-1}\log\tilde P\big(\big|\xi^T(s)-g(s)\big|<\delta,\forall s\in[0,\hat w]\big)=-\frac12\int_0^{\hat w}g'(s)^2\,\mathrm{d}s.\tag{10}$$

Our local upper bound for the BBM now follows directly from (9) and (10):

$$\begin{aligned}\lim_{\delta\to0}\limsup_{T\to\infty}T^{-1}\log P(\exists u\in N_T : X_u^T\in \mathrm{B}_\delta(g)) &\le r\hat w-\frac12\int_0^{\hat w}g'(s)^2\,\mathrm{d}s\\ &=-S(g).\end{aligned}$$

□

4 A new martingale Z_{g_T} for BBM

Let $g\in C_0[0,1]$ be a fixed path. Given that the spine diffusion ξ_t is itself is a $\tilde P$-Brownian motion, it follows that on the sub-filtration $(\mathcal{G}_t)_{0\le t\le T}$,

$$\zeta_{g_T}(t):=\exp\Big\{\int_0^t g_T'(s)\,\mathrm{d}\xi_s-\frac12\int_0^t g_T'(s)^2\,\mathrm{d}s\Big\},\tag{11}$$

is a $\tilde P$-martingale, where

Definition 4.1 *for any fixed $T\ge0$ and any function $g\in C[0,1]$ we define*

$$g_T(s):=Tg(s/T)\qquad\forall s\in[0,T].$$

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

This martingale (11) 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 of Definition 2.2 which counts the number of fission times on the spine up to time t is a Poisson process of rate r and therefore

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

is a $\tilde{P}$-martingale too which will increase the rate of n_t from r to $2r$ if used to change the measure – Kyprianou [10] also uses this martingale.

We 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*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = e^{-rt} 2^{n_t} \times \zeta_{g_T}(t). \tag{12}$$

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.*

This change of measure gives us an additive martingale over the branching particles:

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

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

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

That this is really a martingale is due to the following:

Theorem 4.4 *If we define $\mathbb{Q}_{g_T} := \tilde{\mathbb{Q}}_{g_T}|_{\mathcal{F}_T}$, then $\mathbb{Q}_{g_T}$ 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 (12) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation: for $t \in [0, T]$

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

The foundations article [5] gives more detail on this point. Bearing in mind that $2^{n_t} = \prod_{v<\xi_t} 2$, if we use the representation (6) we get

$$\begin{aligned}
\tilde{P}\big(e^{-rt}2^{n_t}\zeta_{g_T}|\mathcal{F}_t\big) &= \tilde{P}\Big(e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2\times \mathbf{1}_{(\xi_t=u)}\big|\mathcal{F}_t\Big)\\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2\times \tilde{P}\big(\xi_t = u|\mathcal{F}_t\big)\\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2\times \prod_{v<u}\frac{1}{2}\\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} = Z_{g_T}(t).
\end{aligned}$$

We note that in this above chain we used $\tilde{P}\big(\xi_t = u|\mathcal{F}_t\big) = \prod_{v<u}\frac{1}{2}$, which follows easily from (5). □

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

As we shall see in Theorem 5.2 below, the rate at which Z_{g_T} grows under the measure $\mathbb{Q}_{g_T}$ is precisely the rate we need in the large-deviations.

Theorem 5.1 *If $\alpha\in[0,1]$ then $Z_{g_T}(t)^\alpha$ 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 $[0,T]$. This means that for $t\geq s$,

$$P\big(Z_{g_T}(t)^{1+\alpha}|\mathcal{F}_s\big) \geq Z_{g_T}(s)^{1+\alpha}, \qquad P\text{-a.s.}$$

or equivalently, for all $F\in\mathcal{F}_s$,

$$P\big(Z_{g_T}(t)^{1+\alpha};F\big) \geq P\big(Z_{g_T}(s)^{1+\alpha};F\big). \tag{13}$$

But this inequality is exactly the same as:

$$\text{for all } F\in\mathcal{F}_s,\quad \mathbb{Q}_T\big(Z_{g_T}(t)^{\alpha};F\big) \geq \mathbb{Q}_T\big(Z_{g_T}(s)^{\alpha};F\big),$$

and therefore we conclude that $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$-submartingale on $[0,T]$. □

We are going to 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 $\alpha\in[0,1]$,*

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \leq e^{\alpha S(g)T} e^{\frac{1}{2}\alpha^2 T\int_0^1 g'(s)^2\,\mathrm{d}s}\big(1+2rT\big). \tag{14}$$

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:

$$\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big) = e^{-rT}e^{\int_0^T g_T'(s)\,\mathrm{d}\xi_s-\frac{1}{2}\int_0^T[g_T'(s)]^2\,\mathrm{d}s} + \sum_{u<\xi_T} e^{-rS_u}e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\xi_s-\frac{1}{2}\int_0^{S_u}[g_T'(s)]^2\,\mathrm{d}s}. \tag{15}$$

A proof of this can be found in Hardy and Harris [5], but the intuition relies only on the idea that under $\mathbb{Q}_{g_T}$ the subtrees 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

$$\mathrm{d}\xi_s = \mathrm{d}\hat{\xi}_s + g_T'(s)\mathrm{d}s, \tag{16}$$

into (15) we arrive at:

$$\begin{aligned}
\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big) &= e^{\frac{1}{2}\int_0^T g_T'(s)^2\,\mathrm{d}s - rT} e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\frac{1}{2}\int_0^{S_u} g_T'(s)^2\,\mathrm{d}s - rS_u} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s} \\
&= e^{\left(\int_0^1 \frac{1}{2} g'(s)^2 - r\,\mathrm{d}s\right)T} e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\left(\frac{1}{2}\int_0^{S_u/T} [g'(s)]^2 - r\,\mathrm{d}s\right)S_u} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s} \\
&\le e^{\left(\sup_{w\in[0,1]}\int_0^w \frac{1}{2} g'(s)^2 - r\,\mathrm{d}s\right)T} \Big(e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\Big), \\
&= e^{S(g)T} \Big(e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\Big).
\end{aligned} \tag{17}$$

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

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

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

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

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

Since the spine decomposition $\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big)$ is a sum, we can use the following result noted by Neveu [16]

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

Continuing from (17) these lead to

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \le e^{\alpha S(g)T}\, \tilde{\mathbb{Q}}_T\Big(e^{\alpha\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\alpha\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\Big). \tag{18}$$

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

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

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

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

whence from (18) we obtain

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

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:

$$\tilde{\mathbb{Q}}_T\big(1+n_T\big) = 1 + 2rT,$$

and we arrive at

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

□

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 $\varepsilon > 0$,*

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \le e^{(S(g)+\varepsilon)T}\Big) \to 1, \qquad \text{as } T\to\infty. \tag{19}$$

Proof: First of all, for each $\alpha \in [0,1]$, a proof very similar to that given in Theorem 5.1 implies that $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$ submartingale on $t\in[0,T]$, and we now prove a probability bound on its growth.

The estimate (14) works well with Doob's submartingale inequality: for any small $\varepsilon > 0$ and for any fixed $T > 0$,

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_\lambda(s) > e^{(S(g)+\varepsilon)T}\Big) = \mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_\lambda(s)^\alpha > e^{\alpha(S(g)+\varepsilon)T}\Big) \le \frac{\mathbb{Q}_T\big(Z_\lambda(T)^\alpha\big)}{e^{\alpha(S(g)+\varepsilon)T}}.$$

Using (14) this gives

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

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

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

□

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 $g \in C_0[0,1]$. For any fixed $\delta > 0$, we have*

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in B_\delta(g)) \ge -S(g). \tag{20}$$

Proof: 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,

$$P\big(\exists u\in N_T : X_u^T \in \mathrm{B}_\delta(g)\big) = \mathbb{Q}_T\Big(\frac{1}{Z_{g_T}(T)}; \exists u \in N_T : X_u^T \in \mathrm{B}_\delta(g)\Big). \tag{21}$$

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 $\varepsilon > 0$,

$$\begin{aligned} &P\big(\exists u \in N_T\colon X_u^T \in \mathrm{B}_\delta(g)\big) \\ &\geq e^{-(S(g)+\varepsilon)T}\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \leq e^{(S(g)+\varepsilon)T}; \exists u \in N_T\colon X_u^T \in \mathrm{B}_\delta(g)\Big) \\ &\geq e^{-(S(g)+\varepsilon)T}\tilde{\mathbb{Q}}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \leq e^{(S(g)+\varepsilon)T};\, \xi^T \in \mathrm{B}_\delta(g)\Big). \end{aligned} \tag{22}$$

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

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

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

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

Thus from (22) we have

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

which proves (20) since ε was arbitrary. □

Corollary 6.2

$$\lim_{\delta\to 0}\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in B_\delta(g)) = -S(g). \tag{23}$$

Proof: We have proved Theorem 3.1 which can be interpreted as saying that for each $\delta > 0$ there is an $\varepsilon_\delta > 0$ such that

$$-S(g) + \varepsilon_\delta > \limsup_{t\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathrm{B}_\delta(g)),$$

with $\varepsilon_\delta \to 0$ as $\delta \to 0$. Now a trivial inequality between the limsup and liminf combines with this and with the lower bound of Theorem 20 to give

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

which implies that the $\delta \to 0$ limit exists also for the $\liminf$. □

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

7 Improving the weak LDP

As mentioned, the local results of Theorem 1.7 and the fact that the δ-neighbourhoods $\mathrm{B}_\delta(g)$ form a base for the topology of $C_0[0,1]$ means that we have at least a *weak* large-deviations principle: the lower bound of Theorem 1.5 holds, but the upper bound is proven only for *compact* sets (as opposed to *closed* sets). A proof of this can be found at Theorem 4.1.11 of Dembo and Zeitouni [3]: the lower bound is immediate from the definition of an open set as a union of δ-neighbourhoods, and the upper bound for a compact set follows once we cover it with a finite number of δ-neighbourhoods.

We now wish to improve this to get the full LDP of Theorem 1.5. A neat way to do this is to prove that the probabilities are *exponentially tight*.

Definition 7.1 *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.*

As Dembo and Zeitouni [3] state a Lemma 1.2.18, when an exponentially tight family of probability measures satisfy a weak LDP with a rate function $I(\cdot)$, then I is a good rate function and the LDP holds in full.

This approach is particularly adapted to the spines since the question of exponential tightness of the BBM probabilities can be reduced to that of the plain BM probabilities, by the Many-to-One theorem. Exponential tightness for plain Brownian motion is a known result, and therefore the following result will complete our proof of Theorem 1.5.

Theorem 7.2 *The probabilities* $\{P(\exists u \in N_T : X_u^T \in \cdot\ )\}_{T\geq 0}$ *are exponentially tight.*

Proof: For any set $K \subset C_0[0,1]$, we have:

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

whence

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) \leq r + \limsup_{T\to\infty} T^{-1} \log P(\xi^T \in K^{\complement}). \tag{24}$$

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 $K \subset C_0[0,1]$ such that

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

and therefore from (24),

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

□

Thus we conclude that the full LDP holds, and the proof of Theorem 1.5 is concluded.

References

[1] B. Chauvin, *Arbres et processus de Bellman-Harris*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 209–232. MR 88a:60151

[2] Brigitte Chauvin, *Product martingales and stopping lines for branching Brownian motion*, Ann. Probab. **19** (1991), no. 3, 1195–1205. MR 93b:60188

[3] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Springer, 1998.

[4] Yoav Git, *Almost sure path properties of branching diffusion processes*, Séminaire de probabilités, vol. XXXII, Springer, 1998.

[5] Robert Hardy and Simon C. Harris, *A new formulation of the spine approach to branching diffusions*, (2004), no. 0404, Mathematics Preprint, University of Bath.

[6] ———, *A spine proof of a lower-bound for a typed branching diffusion*, (2004), no. 0408, Mathematics Preprint, University of Bath.

[7] ———, *Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales*, (2004), no. 0405, Mathematics Preprint, University of Bath.

[8] S. C. Harris and D. Williams, *Large deviations and martingales for a typed branching diffusion. I*, Astérisque (1996), no. 236, 133–154, Hommage à P. A. Meyer et J. Neveu. MR 98c:60110

[9] Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres, *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185. MR 98j:60122

[10] Andreas Kyprianou, *Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris's probabilistic analysis*, **40** (2004), no. 1, 53–72.

[11] Tzong-Yow Lee, *Some large-deviation theorems for branching diffusions*, Ann. Probab. **20** (1992), no. 3, 1288–1309. MR 93i:60052

[12] Quansheng Liu and Alain Rouault, *On two measures defined on the boundary of a branching tree*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 187–201. MR 98j:60123

[13] Russell Lyons, *A simple path to Biggins' martingale convergence for branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221. MR 1 601 749

[14] Russell Lyons, Robin Pemantle, and Yuval Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), no. 3, 1125–1138. MR 96m:60194

[15] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR 88a:60150

[16] J. Neveu, *Multiplicative martingales for spatial branching processes*, Seminar on Stochastic Processes (E. Çinlar, K.L.Chung, and R.K.Getoor, eds.), Birkhäuser, 1987, pp. 223–241.

[17] S. R. S. Varadhan, *Large deviations and applications*, SIAM, 1984.