Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales

Robert Hardy and Simon C. Harris

March 8, 2004

Abstract

Using the foundations laid down in Hardy and Harris [8], we present new spine proofs of the $\mathcal{L}^p$-convergence of some key additive martingales for three distinct models of branching diffusions, including branching Brownian motion. The spine techniques we develop give clear and simple arguments in the spirit of the conceptual spine proofs found in Kyprianou [19] and Lyons *et al* [22, 21, 18], and they should also extend to more general classes of branching diffusions. Importantly, the techniques in this paper also pave the way for the path large-deviation results for branching diffusions found in Hardy and Harris [9, 10].

1 Overview

In this article we use a change of measure together with spine techniques to analyze the $\mathcal{L}^p$-convergence properties (for $p > 1$) of the strictly-positive martingales for three different models of branching diffusions. It is a common feature of these diffusion models, where there is actually a family of such martingales $\{Z_\lambda : \lambda \in \mathbb{R}\}$, that for all λ within an open interval about 0 the martingale Z_λ is convergent in $\mathcal{L}^p$ for some $p > 1$; for λ outside of this interval the limit of Z_λ is almost surely null. For values of λ at the boundary of this interval, the so-called *critical* values, it has been conjectured (and in some models proven) that the martingales have a null limit – we give a proof for the simpler first model but not for the others since they require different techniques using 'derivative' martingales – see Kyprianou [19] or Harris [13] for examples.

The first model is branching Brownian motion (BBM), for which a proof of $\mathcal{L}^p$-convergence was originally given by Neveu [24] using classical techniques based on the branching decomposition. After introducing the martingale and the measure change via a pathwise construction, in section 2.1 we give a summary of the underlying space and filtrations that we shall use throughout for our spine techniques and show how we can think of the central two measures as being projections of two other measures defined on the larger filtrations, related via a Radon-Nikodym derivative; a foundation article [8] contains fuller details, but in order to make this article self-contained we aim to give enough information here. We actually deal with two variants of BBM: first sections deal with the case of binary-splitting where the fissions produce only two particles, and in section 2.4 we extend the model to allow the fissions to produce a *random* number of offspring.

The second model we look at is a finite-type branching diffusion and is a generalization of a 2-type model treated in the paper by Champneys *et al* [2]. Again, we deal separately with the cases of binary-splitting and random family sizes, and in this second case allow the distribution of the family sizes to be type-dependent.

The third model has a *continuous*-type-space where the type of each particle moves independently as an Orstein-Uhlenbeck process on $\mathbb{R}$. This branching diffusion was first introduced in Harris and Williams [11] and has also been investigated in Harris [12], Harris and Git [14] and Kyprianou and Engländer [7].

Proofs for each of these models each run along similar lines, and it is a credit to the spine approach that this is possible. More classical techniques based on the expectation semigroup are simply not able to generalize easily, since they often require either some *a priori* bounds on the semigroup or involve difficult estimates – for example, in Harris and Williams [11] their important bound of a non-linear term is made possible only by the existence of a good $\mathcal{L}^2$ theory for their operator, and this is not generally available.

Briefly, to prove that the martingale converges in $\mathcal{L}^p$ for some $p > 1$ we use Doob's theorem, and therefore need only to show that the martingale is *bounded* in $\mathcal{L}^p$. The *spine decomposition* is an excellent tool here for showing boundedness of the martingale since it reduces difficult calculations over the whole collection of branching particles to just the single spine process. When $\mathcal{L}^p$-convergence does not hold the martingale limit is almost-surely null, and we prove this by showing that the martingale is almost-surely *not bounded* in a new measure – this approach relies on a measure-theoretic result given below and has become standard in the spine methodology since the important work of Lyons *et al* [22, 21, 18]. We show unboundedness of the martingale just by considering the contribution of the spine, which is shown to be unbounded.

There are a number of reasons why we may be interested in knowing about the $\mathcal{L}^p$ convergence of a martingale: in Neveu's original article [24] it was a means to proving $\mathcal{L}^1$-convergence, whilst Harris and Git [14] and Asmussen and Hering [1] have used it to deduce the almost-sure rate of convergence of the martingale to its limit. Of equal importance are the *techniques* that we use here, and similar ideas have been used in proving a lower bound for a number of problems in the large-deviations theory of branching diffusions – we have used the spine decomposition with Doob's submartingale inequality to get an upper-bound for the growth of the martingale under the new measure which then leads to a lower-bound on the probability that one of the diffusing particles follows an unexpected path – see Hardy and Harris [9] for a spine-based proof of the large deviations principle for branching Brownian motion, and see Hardy and Harris [10] for a proof of a lower bound in the model that we consider in section 4.

2 Branching Brownian motion

Consider a branching Brownian motion (BBM) with constant branching rate r, which is the branching process whereby particles diffuse independently according to a (driftless) 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. It is well known that for any $\lambda \in \mathbb{R}$,

$$Z_\lambda(t) := \sum_{u\in N_t} e^{-rt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} \tag{1}$$

defines a *strictly-positive* P-martingale, so $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ is almost surely finite under each P^x.

We are going to use a change of measure together with the so-called spine decomposition to determine the conditions under which this martingale is $\mathcal{L}^p(P)$-convergent for some $p > 1$. To this end we suppose that $\mathbb{Q}_\lambda$ is a measure on the same filtration $(\mathcal{F}_t)_{t\geq 0}$ which makes the process $\mathbb{X}_t$ equivalent to the following pathwise construction that can also be found in the recent paper of J.Harris, S.C.Harris and Kyprianou [15]:

Definition 2.1 *Under $\mathbb{Q}^x_\lambda$ the point process $\mathbb{X}_t$ evolves as follows:*

- *starting from position x, the original ancestor diffuses according to a Brownian motion on $\mathbb{R}$ with drift λ;*
- *at rate $2r$ the particle undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent;*
- *the other particle initiates, from its birth position, an independent copy of a $P^{\cdot}$ branching Brownian motion with branching rate r.*

In this construction, the individuals that are selected to have a drift of λ make up a (random) line of descent which has come to be referred to as the *spine*. Work carried out by Chauvin and Rouault [5], more recent work by Kyprianou [19] and similar work by Lyons *et al* [21, 18, 22] has shown that

Theorem 2.2 *The change of measure is given by:*

$$\left.\frac{\mathrm{d}\mathbb{Q}_{\lambda}^{x}}{\mathrm{d}P^{x}}\right|_{\mathcal{F}_t} = \frac{Z_{\lambda}(t)}{Z_{\lambda}(0)} = e^{-\lambda x} Z_{\lambda}(t). \tag{2}$$

The paper by Kyprianou [19] used this change of measure and other spine techniques to give necessary and sufficient conditions for the $\mathcal{L}^1$-convergence of this martingale. In fact Kyprianou deals with a slightly more general version of the BBM model in which particle fissions may produce a *random* number of offspring (but always at least one), with that number chosen independently of a particle's position and distributed according to some given distribution on the integers; we look at this more general model in section 2.4. If we were only to consider the binary-splitting case, the result as stated in Kyprianou is:

Theorem 2.3 *Let $\tilde{\lambda} := \sqrt{2r}$. Further convergence properties of the martingale Z_{λ} are determined by the critical value $\tilde{\lambda}$ in that:*

- *if $|\lambda| \geq \tilde{\lambda}$ then $Z_{\lambda}(\infty) = 0$ almost surely;*
- *if $|\lambda| < \tilde{\lambda}$ then $Z_{\lambda}(\infty) > 0$ almost-surely and $Z_{\lambda}(t) \to Z_{\lambda}(\infty)$ in $\mathcal{L}^1(P)$.*

Without loss of generality we throughout suppose that $\lambda \leq 0$, since the cases are symmetrical. The next theorem on $\mathcal{L}^p$-convergence of the martingale for $p > 1$ was originally proven in Neveu [24] by classical techniques, and clearly represents an extension to Kyprianou's above result:

Theorem 2.4 *For each $x \in \mathbb{R}$, and for each $p \in [1, 2]$:*

1. *The martingale Z_{λ} is $\mathcal{L}^p(P^x)$-convergent if $p\lambda^2/2 < r$.*
2. *Almost surely under P^x, $Z_{\lambda}(\infty) = 0$ when $\lambda^2/2 \geq r$.*

Remark: Actually Neveu's [24] result was based on a birth rate of $r = 1$, but the generalization to any $r > 0$ is trivial. We shall give a spine-based proof of this theorem in Section 2.3.

2.1 The underlying space and filtrations

For a clear understanding of the spine techniques that we shall use in all our models, we need a more precise description of spines than the pathwise construction given in definition 2.1. We give fuller details in Hardy and Harris [8], but to make this article self-contained we now briefly lay out the principal elements. The reader who is familiar with the work of Lyons *et al*, or with Kyprianou's paper [19] will notice significant differences in our approach via our use of the filtrations on the single underlying space.

All three models that we consider in this article shall be built on the same underlying space of sample trees with spines, and the measures will all be constructed in analogous ways. Here we lay out the details for the current model of branching Brownian motion, and leave it to the reader to bridge the details when it comes to the other models – [8] contains all the details in a more abstract setting that will cover every model considered in this article.

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 $u \in \tau$, there exists $A_u \in 0, 1, 2, \ldots$ such that $uj \in \tau$ if and only if $1 \le j \le 1 + A_u$, (where $j \in \mathbb{N}$).

We note that in general the underlying trees will not be a binary tree – each node will have $1 + A_u$ branches coming from it. In the binary-branching models that we look at it will actually be the probability measures that exclude $A_u > 1$, but we use this more general case to lay out the details of the underlying spaces.

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 2.5 *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 $N_t := \big\{u \in \tau : S_u - \sigma_u \leq t < S_u\big\}$.

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 $\big\{\emptyset, u_1, u_2, u_3, \ldots\big\}$ 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 [20], 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 $\big\{S_u : u \in \xi\big\}$ 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 2.6 *We define the counting function*

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

or equivalently,

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

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 $\xi_t = \big(\emptyset, u_1, u_2\big)$ 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 [8] 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 2.7 *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 [23] and Chauvin [4, 3].

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

$$\mathrm{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{1+A_v}. \tag{3}$$

(In the binary-branching case we shall necessarily have $\mathrm{Prob}(A_v = 1) = 1$.) This observation (3) is the key to our method for extending the measures, and for this we make use of the following representation found in Lyons [21].

Theorem 2.8 *If f is a $\tilde{\mathcal{F}}_t$-measurable function then we can write:*

$$f = \sum_{u \in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{4}$$

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

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

Definition 2.9 *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*

$$\int_{\tilde{\mathcal{T}}} f \; \mathrm{d}\tilde{P}^x := \int_{\mathcal{T}} \sum_{u \in N_t} f_u \prod_{v<u} \frac{1}{1+A_v} \; \mathrm{d}P^x, \tag{5}$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (4).

The previous approach to spines, exemplified in Lyons [21], 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 (3) 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 2.10 *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 (4) is trivial and therefore by definition

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

However, it can be shown that $\sum_{u\in N_t} \prod_{v<u} \frac{1}{1+A_v} = 1$ 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

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

□

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

Theorem 2.11 *Under $\tilde{P}^x$ the spine diffusion ξ_t is a Brownian motion that starts at x.*

2.2 New measures for BBM

Having now seen the construction of the underlying space and the measure $\tilde{P}^x$, we can define a measure $\tilde{\mathbb{Q}}_\lambda$ via a Radon-Nikodym derivative with respect to $\tilde{P}$:

Definition 2.12 *$\tilde{\mathbb{Q}}_\lambda$ is a measure on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ defined as:*

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}^x_\lambda}{\mathrm{d}\tilde{P}^x}\bigg|_{\tilde{\mathcal{F}}_t} = e^{-\lambda x} \, 2^{n_t} e^{-rt} e^{\lambda \xi_t - \frac{1}{2}\lambda^2 t}. \tag{6}$$

From this definition it is clear that the martingale term $e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}$ is going to give the spine process ξ_t a drift of λ under $\tilde{\mathbb{Q}}_\lambda$. The other term $e^{-rt}2^{n_t}$ is actually a martingale term that will increase the rate of the Poisson process n_t of fission times on the spine from r to $2r$; this can be seen also in Kyprianou [19].

Theorem 2.13 *If we define $\mathbb{Q}^x_\lambda := \tilde{\mathbb{Q}}^x_\lambda|_{\mathcal{F}_\infty}$, then $\mathbb{Q}^x_\lambda$ is a measure on $\mathcal{F}_\infty$ and*

$$\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\bigg|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)},$$

so that under $\mathbb{Q}^x_\lambda$ (and therefore also under $\tilde{\mathbb{Q}}^x_\lambda$) the branching-diffusion point process $\mathbb{X}_t$ has exactly the pathwise construction given in definition 2.1.

There are at least two ways to prove this result: Kyprianou [19] bases his proof on a decomposition of the measure $\tilde{P}$ as a product of measures for the spine's motion, the fission-counting process n_t, and measures on the sub-trees born from the spine. Because of our using filtrations the way we do, we have an alternative.

Proof of Theorem 2.13: It is clear from the definition of the conditional expectation that the change of measure (6) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation:

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^x_\lambda}{\mathrm{d}\tilde{P}^x}\right|_{\mathcal{F}_t} = e^{-\lambda x}\,\tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t}|\mathcal{F}_t\big).$$

Bearing in mind that $2^{n_t} = \prod_{v<\xi_t} 2$, if we use the representation (4) we get

$$\begin{aligned}
\tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t}|\mathcal{F}_t\big) &= \tilde{P}\Big(e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t}\times\prod_{v<u}2\times\mathbf{1}_{(\xi_t=u)}\big|\mathcal{F}_t\Big)\\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t}\times\prod_{v<u}2\times\tilde{P}\big(\xi_t=u|\mathcal{F}_t\big)\\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t}\times\prod_{v<u}2\times\prod_{v<u}\frac{1}{2}\\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t} = Z_\lambda(t).
\end{aligned}$$

Here, the result $\tilde{P}\big(\xi_t=u|\mathcal{F}_t\big) = \prod_{v<u}\frac{1}{2}$ is the case of (3) for binary-splitting. Thus

$$e^{-\lambda x}\,\tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t}|\mathcal{F}_t\big) = \frac{Z_\lambda(t)}{Z_\lambda(0)}.$$

□

2.3 Proof of Theorem 2.4

Just before we proceed to the proof we recall a naturally occurring eigenvalue that will also appear in later models.

Definition 2.14 *For each $\lambda\in\mathbb{R}$ we define $E_\lambda := \frac{1}{2}\lambda^2 + r$.*

Then Z_λ can be re-written as

$$Z_\lambda(t) = \sum_{u\in N_t} e^{\lambda X_u(t)-E_\lambda t},$$

and simple algebra reveals

$$p\lambda^2/2 < r \quad\Leftrightarrow\quad pE_\lambda - E_{p\lambda} > 0.$$

2.3.1 Proof of part 1:

We are going to prove that for every $p\in(1,2]$ the martingale Z_λ is $\mathcal{L}^p(P)$-convergent if $pE_\lambda - E_{p\lambda} > 0$. Furthermore, since $P^x(Z_\lambda(t)) = e^{\lambda x}P^0(Z_\lambda(t))$ we do not lose generality if we suppose that $x=0$, and from now on this is implicit and we do not use the superscript on the measures.

From the change of measure in Theorem 2.2 or Theorem 2.13 it is clear that

$$P(Z_\lambda(t)^p) = P(Z_\lambda(t)^{p-1}Z_\lambda(t)) = \mathbb{Q}_\lambda(Z_\lambda(t)^q),$$

where $q := p-1$. Our aim is to prove that $\mathbb{Q}_\lambda(Z_\lambda(t)^q)$ is bounded for $t\in\mathbb{R}$, since then $Z^p_\lambda(t)$ must be bounded in $\mathcal{L}^p(P)$ and Doob's theorem will then imply that Z_λ is convergent in $\mathcal{L}^p(P)$.

As we mentioned, the algebra $\tilde{\mathcal{G}}_\infty$ gives us the very important *spine-decomposition* of the martingale Z_λ:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{k=1}^{n_t} e^{\lambda\xi_{S_k}-E_\lambda S_k} + e^{\lambda\xi_t-E_\lambda t},$$

where the sum is taken to equal 0 if $n_t = 0$. Full details of this can be found in [8], but the intuition is quite clear: since the particles that do not make up the spine grow to become independent copies of $\mathbb{X}_t$ distributed *as if under* P, the fact that Z_λ is a P-martingale on these subtrees implies that their contributions to the above decomposition are just equal to their *immediate* contribution on being born at time S_k at location ξ_{S_k}. **Remark:** We emphasize that here we must use the measure $\tilde{\mathbb{Q}}_\lambda$, since $\mathbb{Q}_\lambda$ cannot measure the algebra $\tilde{\mathcal{G}}_\infty \not\subseteq \mathcal{F}_\infty$.

We continue with the conditional form of Jensen's inequality, which says that for $q \in (0,1]$:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)^q | \tilde{\mathcal{G}}_\infty\big) \leq \tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t) | \tilde{\mathcal{G}}_\infty\big)^q. \tag{7}$$

The spine decomposition is a sum, and from Neveu's original proof we use the following simple inequality:

Proposition 2.15 *If* $q \in (0,1]$ *and* $u, v > 0$ *then* $(u+v)^q \leq u^q + v^q$,

to obtain,

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t) | \tilde{\mathcal{G}}_\infty\big)^q \leq \sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k} + e^{q\lambda\xi_t - qE_\lambda t}.$$

Taking $\tilde{\mathbb{Q}}_\lambda$-expectations of this and using (7),

$$\begin{aligned}
\mathbb{Q}_\lambda(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)^q) &\leq \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k} + e^{q\lambda\xi_t - qE_\lambda t}\Big) \\
&= \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) + \tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big),
\end{aligned} \tag{8}$$

and the proof of $\mathcal{L}^p$-boundedness will be complete once we show that this RHS is bounded in t.

As written, (8) is made up of two terms, and since they play a central role from here on we name them explicitly: on the far right we have the **spine term** $\tilde{\mathbb{Q}}_\lambda\big(e^{q\lambda\xi_t - qE_\lambda t}\big)$, the other being the **sum term** $\tilde{\mathbb{Q}}_\lambda\big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\big)$.

The spine term: Changing from $\tilde{P}$ to $\tilde{\mathbb{Q}}_\lambda$ gives the spine a drift of λ, and therefore the change-of-measure for just the spine's motion (i.e. on the algebra $\mathcal{G}_t$) is carried out by the martingale $e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}$:

$$\begin{aligned}
\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big) &:= \tilde{P}\Big(e^{q\lambda\xi_t - qE_\lambda t} \times e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}\Big), \\
&= e^{-(pE_\lambda - E_{p\lambda})t}\, \tilde{P}\Big(e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}\Big), \\
&= e^{-(pE_\lambda - E_{p\lambda})t},
\end{aligned} \tag{9}$$

since the second-line term $e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}$ is also a $\tilde{P}$-martingale.

The sum term: Under the measure $\tilde{\mathbb{Q}}_\lambda$ we know that the fission times S_u on the spine occur as a Poisson process of rate $2r$. Appealing to standard results from Poisson theory (see [16] for example) we can therefore write the sum term as an integral:

$$\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) = \tilde{\mathbb{Q}}_\lambda\Big(\int_0^t 2r\, e^{q\lambda\xi_s - qE_\lambda s}\, \mathrm{d}s\Big).$$

In this integral all the terms are positive and so Fubini's theorem can be used, which along with the equality (9) above gives

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k}-qE_\lambda S_k}\Big) &= \int_0^t 2r\,\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_s-qE_\lambda s}\Big)\,\mathrm{d}s\\ &= 2r\int_0^t e^{-(pE_\lambda-E_{p\lambda})s}\,\mathrm{d}s\\ &= \frac{2r}{pE_\lambda-E_{p\lambda}}\Big[1-e^{-(pE_\lambda-E_{p\lambda})t}\Big], \qquad \text{if } pE_\lambda\neq E_{p\lambda}. \qquad (10)\end{aligned}$$

Thus we have found an explicit upper-bound:

$$P(Z_\lambda(t)^p)\leq \frac{2r}{pE_\lambda-E_{p\lambda}}\Big[1-e^{-(pE_\lambda-E_{p\lambda})t}\Big]+e^{-(pE_\lambda-E_{p\lambda})t}. \qquad (11)$$

If $pE_\lambda - E_{p\lambda} > 0$ this clearly implies that $P(Z_\lambda(t)^p)$ will remain bounded as $t\to\infty$, which together with Doob's theorem will complete the proof of the first part of Theorem 2.4. □

2.3.2 Proof of Part 2:

The following proof was first given by Kyprianou [19], in the more general case of a BBM where fissions may produce a random number of offspring (but always at least one). The second part of the following theorem is the key element in using the measure change (2) to determine properties of the martingale Z_λ:

Theorem 2.16 *Suppose that P and $\mathbb{Q}$ are two probability measures on a space $(\Omega,\mathcal{F}_\infty)$ with filtration $(\mathcal{F}_t)_{t\geq 0}$, such that for some positive martingale Z_t,*

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\bigg|_{\mathcal{F}_t}=Z_t.$$

The limit $Z_\infty := \limsup_{t\to\infty} Z_t$ therefore exists and is finite almost surely under P. Furthermore, for any $F\in\mathcal{F}_\infty$

$$\mathbb{Q}(F)=\int_F Z_\infty\,\mathrm{d}P+\mathbb{Q}\big(F\cap\{Z_\infty=\infty\}\big), \qquad (12)$$

and consequently $P(Z_\infty=0)=1$ if and only if $\mathbb{Q}(Z_\infty=\infty)=1$.

A proof of the decomposition (12) can be found in Durrett [6], at page 241. In order to show that the martingale limit of our $Z_\lambda(t)$ is null we therefore intend to show that

$$\mathbb{Q}_\lambda\Big(\limsup_{t\to\infty} Z_\lambda(t)=\infty\Big)=1.$$

Because one of the individuals $u\in N_t$ *must* be the spine, it is immediate that

$$Z_\lambda(t)=\sum_{u\in N_t} e^{\lambda X_u(t)-E_\lambda t} > e^{\lambda\xi_t-E_\lambda t}=e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t-rt}.$$

Under the measure $\tilde{\mathbb{Q}}_\lambda$ the spine has a linear drift equal to λ, whence under $\tilde{\mathbb{Q}}_\lambda$,

$$e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t-rt}=e^{\lambda B_t+(\frac{1}{2}\lambda^2-r)t},$$

where B_t is a $\tilde{\mathbb{Q}}_\lambda$-Brownian motion. Thus $\lambda^2\geq 2r$ will force $\limsup_{t\to\infty} Z_\lambda(t)=\infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda$.

□

2.4 Random family sizes

In the binary-branching model of BBM that we just considered, at each fission time two particles are produced. Kyprianou [19] deals with a model in which at a fission time an individual u may split into a $1 + A$ particles where A is an integer-valued random variable chosen independently of the position $X_u(t)$ of the individual u, with general distribution:

$$P(A = i) = p_i, \qquad i \in \{0, 1, \ldots\},$$

giving an average size of $m := P(A) = \sum_{i=0}^{\infty} i\, p_i$. This introduction of random family sizes implies a small change in the form of the branching martingale:

Theorem 2.17 *For any $\lambda \in \mathbb{R}$,*

$$Z_\lambda(t) := \sum_{u \in N_t} e^{-rmt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} = \sum_{u \in N_t} e^{\lambda X_u(t) - E_\lambda t}$$

is a P-martingale, where $E_\lambda := \frac{1}{2}\lambda^2 + rm$.

We remind the reader that the filtration $(\tilde{\mathcal{G}}_t)_{t \geq 0}$ includes knowledge of the sizes of the families produced by all fissions on the spine:

$$\tilde{\mathcal{G}}_t = \sigma(\mathcal{G}_t, \mathrm{node}_s(\xi) : s \leq t, A_u : u < \xi_t).$$

In the binary-splitting model it was *always* the case that $A_u = 1$ and so we didn't really notice this extra information in $\tilde{\mathcal{G}}_\infty$, but here it is important in the spine decomposition.

The most significant alteration for the measure change is that the distribution of the family sizes produced by fissions on the spine (but not off it) is *tilted*:

Theorem 2.18 *If we define the measure $\mathbb{Q}_\lambda^x$ via*

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^x}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = e^{-\lambda x} Z_\lambda(t),$$

then it follows that under $\mathbb{Q}_\lambda^x$ the point process $\mathbb{X}_t$ evolves as follows:

- *starting from position x, the original ancestor diffuses according to a Brownian motion on $\mathbb{R}$ with drift λ;*
- *at an accelerated rate $(1 + m)r$ the particle undergoes fission producing $1 + \tilde{A}$ particles, where the distribution of $\tilde{A}$ is still independent of the spine's motion but is* size-biased*:*

$$\tilde{\mathbb{Q}}_\lambda(\tilde{A} = i) = \frac{(i + 1)p_i}{m + 1}, \qquad i \in \{0, 1, \ldots\}.$$

- *with equal probability, one of these offspring particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent with the size-biased offspring distribution;*
- *the other particles initiate, from their birth position, an independent copy of a $P^\cdot$ branching Brownian motion with branching rate r and family-size distribution given by A (which is without the size-biasing).*

This phenomena of size-biasing was noted in the Lyons *et al* papers, and is a common feature of such measure changes in the spine approach.

As we have seen, the measures P and $\mathbb{Q}_\lambda$ on the $\mathcal{F}_\infty$ can be extended to $\tilde{P}$ and $\tilde{\mathbb{Q}}_\lambda$ on $\tilde{\mathcal{F}}_\infty$; or equivalently, we can define P and $\mathbb{Q}_\lambda$ as the projections onto $\mathcal{F}_\infty$ of the measures $\tilde{P}$ and $\tilde{\mathbb{Q}}_\lambda$ defined on $\tilde{\mathcal{F}}_\infty$.

The theorem on the $\mathcal{L}^p$-convergence of Z_λ is now slightly modified to take into account the random distribution of the family sizes:

Theorem 2.19 *For each $x \in \mathbb{R}$, and for each $p \in [1,2]$:*

1. *The martingale Z_λ is $\mathcal{L}^p(P^x)$-convergent if $p\lambda^2/2 < rm$ and $P(A^p) < \infty$.*
2. *Almost surely under P^x, $Z_\lambda(\infty) = 0$ when $\lambda^2/2 \geq rm$.*

The proof is not very different from the binary-splitting case since the spine decomposition is different only in the sum term:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{k=1}^{n_t} A_u e^{\lambda\xi_{S_k} - E_\lambda S_k} + e^{\lambda\xi_t - E_\lambda t},$$

and therefore we go rather more quickly here. As before we can use Jensen's inequality and Proposition 2.15 to arrive at

$$\mathbb{Q}_\lambda(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)^q) \leq \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k} + e^{q\lambda\xi_t - qE_\lambda t}\Big).$$

The spine term can be dealt with as at (9), and the sum term can be written as an integral, but in order to deal with the random number of offspring we first use conditioning (without knowledge of the family sizes) to replace the term A_u^q with an expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\big|\mathcal{G}_t\Big) &= \sum_{k=1}^{n_t} \tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big) e^{q\lambda\xi_{S_k} - qE_\lambda S_k} && \text{by independence,}\\ &= \tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big)\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}. && (13)\end{aligned}$$

The term $\tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big)$ is guaranteed to be finite if $P(A^p)$ is:

Lemma 2.20 *If $P(A^p) < \infty$ then $\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) < \infty$.*

Proof:

$$\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) = \sum_{i=1}^{\infty} i^q \frac{i+1}{m+1} p_i = \frac{1}{m+1}\Big(\sum_{i=1}^{\infty} i^p p_i + \sum_{i=1}^{\infty} i^q p_i\Big) = \frac{P(A^p) + P(A^q)}{m+1} < \frac{2P(A^p)}{m+1}.$$

$\square$

Taking expectations of both sides of (13), converting the sum to an integral and then using Fubini's theorem gives:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}\Big) &= \tilde{\mathbb{Q}}_\lambda\big(\tilde{A}^q\big)\int_0^t 2r\,\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s\\ &= 2r\,\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q)\int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s && \text{by (9).}\end{aligned}$$

Thus the condition $P(A^p) < \infty$ in the first part of Theorem 2.19 means that $\mathbb{Q}_\lambda(Z_\lambda(t)^q)$ will be bounded if $pE_\lambda - E_{p\lambda} > 0$, and it can be shown that with $E_\lambda = \frac{1}{2}\lambda^2 + rm$,

$$pE_\lambda - E_{p\lambda} > 0 \quad \Leftrightarrow \quad p\lambda^2/2 < rm,$$

completing the proof. The proof of the second part of Theorem 9 does not need to be changed. $\square$

3 A typed branching diffusion

We move on to consider a second model of a branching diffusion in which diffusing *typed* particles branch at random times. The diffusion coefficient and birth rate of each particle will depend on its current type, which will be any one of a finite number and will change in time according to an independent Markov chain. Precisely, the finite-set $I := \{1, \ldots, n\}$ is to be the space of *types*, so that $J := \mathbb{R} \times I$ is the underlying space for all the branching particles. We suppose that under the probabilities $\{\tilde{P}^{x,y} : (x, y) \in J\}$ defined on $(\tilde{\mathcal{F}}_t)_{t\geq 0}$, the natural filtration with a spine, the process evolves according to the following description:

The spine process (ξ_t, η_t) in J is such that, under $\tilde{P}^{x,y}$, its *type-location* η_t starts at $y \in I$ and is an irreducible, time-reversible Markov chain on I with Q-matrix θQ (θ is a strictly positive constant) and invariant measure $\pi = (\pi_1, \ldots, \pi_n)$; its *spatial-location* $\xi_t \in \mathbb{R}$ moves as a Brownian motion, starting from $x \in \mathbb{R}$, with zero drift and diffusion coefficient $a(y) > 0$ whenever η_t is in state y:

$$\mathrm{d}\xi_t = a(\eta_t)^{\frac{1}{2}}\,\mathrm{d}B_t, \qquad B_t \text{ a } \mathbb{P}\text{-Brownian motion.} \tag{14}$$

The formal generator of this process (ξ_t, η_t) is:

$$\mathcal{H}F(x, y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + \theta\sum_{j\in I} Q(y, j)F(x, j), \qquad (F : J \to \mathbb{R}). \tag{15}$$

We often use matrix calculations, and it is convenient to gather the diffusion coefficients together in a diagonal matrix $A := \mathrm{diag}[a(1), \ldots, a(n)]$.

The branching diffusion. Under $P^{x,y}$, the projection of $\tilde{P}^{x,y}$ onto the natural filtration *without spines*, each branching particle $\big(X_u(t), Y_u(t)\big)$ moves like the above description, and undergoes fission – to produce a number of children at its own current position and of its own current type – at a rate $r(y)$, where y represents the particle's current type. We gather together the birth rates in a diagonal matrix $R := \mathrm{diag}[r(1), \ldots, r(n)]$.

It should be noted that the condition of time-reversibility on the Markov chain is not absolutely necessary, and is really just a simplifying assumption that gives us an easier $\mathcal{L}^2$ theory for the matrices and eigenvectors. If we were to drop this assumption we would still have an $\mathcal{L}^2$ theory since we are working with finite vectors, but it would not be so immediate and our aim is really to show how the spine techniques work – lessening the geometric complexity of the model serves a good purpose.

In the next sections we are assuming that fission times produce two particles, and then later extend this to a model in which the number of offspring is determined by a type-dependent distribution.

3.1 The martingale

Via generators it is possible to show that for any $\lambda \in \mathbb{R}$, any function (vector) $v_\lambda : I \to \mathbb{R}$ and any number $E_\lambda \in \mathbb{R}$, the expression

$$Z_\lambda(t) := \sum_{k=1}^{N(t)} v_\lambda(Y_k(t))\, e^{\lambda X_k(t) - E_\lambda t}, \tag{16}$$

will be a martingale if and only if v_λ and E_λ satisfy:

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + R\Big) v_\lambda = E_\lambda v_\lambda, \tag{17}$$

which is to say that v_λ must be an eigenvector of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$, with eigenvalue E_λ.

Definition 3.1 *For two vectors u, v on I, we define*

$$\big\langle u, v \big\rangle_\pi := \sum_{i=1}^{n} u_i v_i \pi_i,$$

which gives us a Hilbert space which we refer to as $\mathcal{L}^2(\pi)$. We suppose that the eigenvector v_λ is normalized so that $\|v_\lambda\|_\pi := \big\langle v_\lambda, v_\lambda \big\rangle_\pi = 1$.

The fact that the Markov chain is time-reversible implies that the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ is self-adjoint with respect to this inner product. This in itself is enough to guarantee the existence of eigenvectors in $\mathcal{L}^2(\pi)$, but the fact that we are dealing with a finite-state Markov chain means that we also have the *Perron-Frobenius* theory to hand, which allows us to suppose that v_λ is a *strictly positive* eigenvector whose eigenvalue E_λ is real and the farthest to the right of all the other eigenvalues – see Seneta [25] for details. This implies a useful representation for the eigenvalue:

Theorem 3.2

$$E_\lambda = \sup_{\|v\|_\pi = 1} \Big\langle \big((\lambda^2/2)A + \theta Q + R\big)\, v, v \Big\rangle_\pi, \tag{18}$$

since it is the rightmost eigenvalue.

A proof can be found in Kreyzig [17]. From this it is not difficult to show that E_λ is a strictly-convex function of λ, which incidentally is also a feature of the third model that we look at later. Interestingly, it will be seen in our proofs that it is the geometry of the eigenvalue E_λ that determines the interval that gives rise to martingales $Z_\lambda(t)$ that are $\mathcal{L}^p$-convergent.

Corollary 3.3 *As a function of λ, E_λ is strictly-convex and infinitely differentiable with*

$$E'_\lambda = \lambda \big\langle A v_\lambda, v_\lambda \big\rangle_\pi. \tag{19}$$

If we define the speed function

$$c_\lambda := -E_\lambda / \lambda, \tag{20}$$

then on $(-\infty, 0)$ the function c_λ has just one minimum at a single point $\tilde{\lambda}(\theta)$, and is strictly increasing to $+\infty$ as $\lambda \downarrow -\infty$ or as $\lambda \uparrow 0$.

Furthermore, and importantly for our later proofs of $\mathcal{L}^p$-convergence of the martingale, for each $\lambda \in (\tilde{\lambda}(\theta), 0]$ there is some $p > 1$ such that $pE_\lambda - E_{p\lambda} > 0$; on the other hand, if $\lambda < \tilde{\lambda}(\theta)$ there is no such $p > 1$.

We refer to the function c_λ as the speed function since it relates to the asymptotic speed of the travelling waves associated with the martingale $Z_\lambda(t)$; see Harris [13] or Champneys *et al* [2] for details of the relationship between branching-diffusion martingales and travelling waves.

Since $Z_\lambda(t)$ is a strictly-positive martingale it is immediate that $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ exists and is finite almost-surely under $P^{x,y}$. We shall prove

Theorem 3.4 *Suppose that $\lambda \le 0$.*

1. *For every $p \in (1,2]$ the martingale Z_λ is $\mathcal{L}^p$-bounded provided that $pE_\lambda - E_{p\lambda} > 0$. In fact this inequality holds for some $p \in (1,2]$ if and only if $\lambda \in (\tilde{\lambda}(\theta), 0]$ (see Lemma 3.3).*
2. *Almost surely (under P), $Z_\lambda(\infty) = 0$ if $\lambda < \tilde{\lambda}(\theta)$.*

Remark: Both parts together imply that the martingale converges in $\mathcal{L}^1$ only if $\lambda \in (\tilde{\lambda}(\theta), 0]$. The question of what happens at the critical $\lambda = \tilde{\lambda}(\theta)$ is not considered, but based on the work by Harris [13] or Kyprianou [19] for BBM, we conjecture (but do not prove) that the martingale limit $Z_\lambda(\infty)$ is null for λ at the critical value.

Our spine approach means that we should like to make a change of measure with $Z_\lambda(t)$ as the Radon-Nikodym derivative, as we did for BBM at Theorem 2.13. Here a pathwise construction could me made like the one laid out in Definition 2.1, but we have seen the better alternative:

Definition 3.5 *For each $\lambda \le 0$ we define a measure $\tilde{\mathbb{Q}}^{x,y}_\lambda$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^{x,y}_\lambda}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}}\, e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} 2^{n_t} \times v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}. \tag{21}$$

This Radon-Nikodym derivative is going to introduce drift to the spine via the martingale term $v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}$, as we see in the following section. The other martingale term $e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} 2^{n_t}$ is a well-known martingale for the Cox process n_t that will increase the rate at which fission times occur on the spine – see Kyprianou [19] for more details; the terms $v_\lambda(y)^{-1}e^{-\lambda x}$ are just normalizing constants. A proof like that for Theorem 2.13 with the conditional expectation works here to give:

Theorem 3.6 *If we define $\mathbb{Q}^{x,y}_\lambda := \tilde{\mathbb{Q}}^{x,y}_\lambda|_{\mathcal{F}_\infty}$, then $\mathbb{Q}^{x,y}_\lambda$ is a measure on $\mathcal{F}_\infty$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}^{x,y}_\lambda}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

3.2 The spine process (ξ_t, η_t) under $\tilde{\mathbb{Q}}_\lambda$

In the BBM model it was clear to see that the spine process ξ_t received a drift under the measure $\tilde{\mathbb{Q}}$; something similar happens here:

Lemma 3.7 *Under $\tilde{\mathbb{Q}}_\lambda$ the spine process (ξ_t, η_t) has generator:*

$$\mathcal{H}_\lambda F\,(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \sum_{j\in I} \theta Q_\lambda(y,j)F(x,j), \tag{22}$$

where Q_λ is an honest Q-matrix:

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \ne j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + r(i) & \text{if } i = j \end{cases}$$

with invariant measure $\pi_\lambda = v_\lambda^2 \pi$.

Thus under $\tilde{\mathbb{Q}}_\lambda$ the Q-matrix (generator) of η_t is changed, and the process $\xi_t \in \mathbb{R}$ is given an instantaneous drift of $a(\eta_t)\lambda$. The form of this above generator can be obtained from the theory of Doob's h-transforms, due to the fact that on the algebra $\mathcal{G}_t$ the change of measure is given by:

$$\left.\frac{\mathrm{d}\mathbb{Q}^{x,y}_\lambda}{\mathrm{d}P^{x,y}}\right|_{\mathcal{G}_t} = \frac{1}{v_\lambda(y)e^{\lambda x}} v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}. \tag{23}$$

The long-term behaviour under $\tilde{\mathbb{Q}}_\lambda$ of the spine diffusion ξ_t can now be retrieved from the generator (22) and the properties of E_λ stated in Lemma 3.3:

Corollary 3.8 *The long-term drift of the spine is given explicitly as*

$$\lim_{t\to\infty} t^{-1}\xi_t = E'_\lambda.$$

Proof: From the generator stated at (22) we can write:

$$\xi_t = B\left(\int_0^t a(\eta_s)\mathrm{d}s\right) + \lambda \int_0^t a(\eta_s)\mathrm{d}s,$$

where $B(t)$ is a $\tilde{\mathbb{Q}}_\lambda$-Brownian motion. Then by the ergodic theorem and the fact that $\pi_\lambda = v_\lambda^2\pi$:

$$t^{-1}\xi_t \to \lambda \sum_{y\in I} a(y)\pi_\lambda(y) = \lambda \sum_{y\in I} a(y)v_\lambda^2(y)\pi(y) = \lambda\big\langle Av_\lambda, v_\lambda\big\rangle_\pi = E'_\lambda.$$

$\square$

Direct calculation from (20) gives $E'_\lambda = -c_\lambda - \lambda c'_\lambda$, and therefore $t^{-1}(\xi_t + c_\lambda t) \to -\lambda c'_\lambda$, whence

- $\xi_t + c_\lambda t$ drifts off to $+\infty$ if $\lambda \in (\tilde{\lambda}(\theta), 0]$,
- $\xi_t + c_\lambda t$ drifts off to $-\infty$ if $\lambda < \tilde{\lambda}(\theta)$.

This second fact will be important in showing that the martingale limit is null when $\lambda < \tilde{\lambda}(\theta)$.

Before moving on to our proof of Theorem 3.4, and just to drive home the point we state the pathwise construction under $\tilde{\mathbb{Q}}^x_\lambda$:

- starting from position x, the original ancestor diffuses according to the generator (22);
- at rate $2R(\eta_t)$ the particle undergoes fission producing two particles;
- with equal probability, one of these two particles is selected to form the next node of the spine;
- this chosen particle repeats stochastically the behaviour of the parent;
- the other particle initiates, from its birth position, an independent copy of a $P^\cdot$ branching Brownian motion with branching rate $R(\cdot)$.

3.3 Proof of Theorem 3.4

Proof of Part 1 of Theorem 3.4: Suppose $p \in (1,2]$, then with $q := p-1$ a slight modification of the BBM proof arrives at

$$\begin{aligned} P^{x,y}(Z_\lambda(t)^p) &= \mathbb{Q}_\lambda^{x,y}(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda(t)^q) \\ &\leq \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\Big) + \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big) \end{aligned}$$

and the proof of $\mathcal{L}^p$-boundedness will be complete once we show that this RHS expectation is bounded in t.

The spine term. It is always useful to first focus on the spine term, since we can change the measure with (23) to get

$$\begin{aligned} \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t-qE_\lambda t}\Big) &= \tilde{P}^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t-qE_\lambda t} \times \frac{v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}}{v_\lambda(y)e^{\lambda x}}\Big) \\ &= e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times e^{-(pE_\lambda - E_{p\lambda})t}\ \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big). \end{aligned} \tag{24}$$

Bearing in mind that η_t is a finite-state irreducible Markov chain and therefore ergodic, and given $\pi_\lambda(y) = v_\lambda(y)^2\pi(y)$, the following result is immediate,

Lemma 3.9 *In the finite-type model, for any $\lambda, p \in \mathbb{R}$, the expectation*

$$\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big)$$

is positive, bounded and convergent with

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) = \frac{v_\lambda^p}{v_{p\lambda}} \cdot \pi_{p\lambda} = \big\langle v_\lambda^p, v_{p\lambda}\big\rangle_\pi,$$

With boundedness and convergence of $\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\big)$ ensured by the above lemma, it follows from (24) that in the long term the growth or decay of the spine term is determined by the term $e^{-(pE_\lambda - E_{p\lambda})t}$.

The sum term. We know that under $\tilde{\mathbb{Q}}_\lambda$ the fission times S_u on the spine occur as a *Cox* process – that is, conditional on knowing η, the times occur as a Poisson process of rate $2R(\eta_s)$. Therefore, if we condition on $\mathcal{G}_t$ which knows about η_s at all times $0 \leq s \leq t$ we can transform the sum into an integral and use Fubini's theorem:

$$\begin{aligned} \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\Big) &= \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\big|\mathcal{G}_t\Big)\Big) \\ &= \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\int_0^t 2R(\eta_s)\, v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\,\mathrm{d}s\Big) \\ &= \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)v_\lambda(\eta_s)^q e^{q\lambda\xi_s-qE_\lambda s}\Big)\,\mathrm{d}s. \end{aligned} \tag{25}$$

The change of measure used in (24) can be used on the sum term in its integral form,

$$\begin{aligned} \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)v_\lambda(\eta_s)^q e^{q\lambda\xi_s-qE_\lambda s}\Big)\,\mathrm{d}s &= e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)}\int_0^t \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(2R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big)e^{-(pE_\lambda-E_{p\lambda})s}\,\mathrm{d}s \\ &\leq K(p,\lambda,y)\int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s \end{aligned}$$

where

$$K(p,\lambda,x,y) := e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \sup_{s\geq 0} \tilde{\mathbb{Q}}^{0,y}_{p\lambda}\Big(2R(\eta_s)\,\frac{v^p_\lambda}{v_{p\lambda}}(\eta_s)\Big), \tag{26}$$

is finite by a simple adaptation of the above lemma since the birth rates $R(\cdot)$ are clearly bounded.

Having dealt with the spine term and the sum term, we have therefore obtained an explicit upper-bound, (if $pE_\lambda - E_{p\lambda} \neq 0$)

$$P^{x,y}\big(Z_\lambda(t)^p\big) \leq K(p,\lambda,x,y)\bigg\{\frac{2\bar{r}}{pE_\lambda - E_{p\lambda}}\Big[1 - e^{(pE_\lambda - E_{p\lambda})t}\Big] + e^{-(pE_\lambda - E_{p\lambda})t}\bigg\},$$

where $\bar{r} = \max r(i)$ and hence:

$$pE_\lambda - E_{p\lambda} > 0 \quad \Rightarrow \quad P^{x,y}\big(Z_\lambda(t)^p\big) \text{ is bounded for all } t.$$

Together with the facts laid out in Lemma 3.3 this completes the proof of the first part of Theorem 3.4.

□

Proof of Part 2 of Theorem 3.4: For the second part we again use the spine term as a lower bound to $Z_\lambda(t)$:

$$Z_\lambda(t) \geq v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t} = v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)}.$$

We aim to show that under $\tilde{\mathbb{Q}}_\lambda$ this spine term is almost-surely unbounded whenever $\lambda < \tilde{\lambda}(\theta)$, leading to the result

$$\limsup_{t\to\infty} Z_\lambda(t) = \infty, \qquad \mathbb{Q}_\lambda\text{-a.s.}$$

which with Theorem 2.16 will imply

$$Z_\lambda(\infty) = 0 \qquad P\text{-a.s.}$$

It was seen that as a consequence of Corollary 3.8, $\xi_t + c_\lambda t \to -\infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda$ whenever $\lambda < \tilde{\lambda}(\theta) < 0$, and therefore

$$\limsup_{t\to\infty} v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)} = \infty$$

because $v(\eta_t) \geq \min_{y\in I} v_\lambda(y) > 0$. This concludes the proof of part 2.

□

3.4 Random family sizes

As we did for the BBM model earlier, we now consider extending our typed branching diffusion to allow a type-dependent randomness in the number of offspring that each individual produces of its own type when it undergoes fission. Thus we suppose that under the probabilities $\{\tilde{P}^{x,y} : (x,y) \in J\}$ defined on $(\tilde{\mathcal{F}}_t)_{t\geq 0}$, if a type-$w$ particle undergoes fission then the number $1 + A_u$ of type-w offspring it produces is distributed like a random variable $1 + A(w) \in \{1, 2, \ldots\}$ with

$$P\big(A(w) = i\big) = p_i(w), \qquad i \in \{0, 1, \ldots\},$$

and mean $P\big(A(w)\big) = m(w)$. For the reasons seen at (13) and then in Lemma 2.20, we suppose that there is some $p > 1$ such that for each $w \in I$,

$$P\big(A(w)^p\big) < \infty.$$

In the spine composition below we shall need to refer to the tilted distribution $\tilde{\mathbb{Q}}_\lambda(\tilde{A}(w)^q)$, and from Lemma 2.20 we know what this value is:

Definition 3.10 *For $p \in (1,2]$ and $q = p-1$, we define*

$$M_q(w) := \frac{P(A^p(w)) + P(A^q(w))}{m(w)+1}.$$

The form of the martingale is unchanged by the random offspring numbers, but the relationship between v_λ and E_λ is altered to account for the average family size being $m(w)$ when a fission occurs for a type-w particle:

Definition 3.11

$$Z_\lambda(t) := \sum_{k=1}^{N(t)} v_\lambda(Y_k(t))\, e^{\lambda X_k(t) - E_\lambda t}, \tag{27}$$

where v_λ is a vector on I normalized so that $\langle v_\lambda, v_\lambda \rangle_\pi = 1$, and $E_\lambda \in \mathbb{R}$ satisfying

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + mR\Big) v_\lambda = E_\lambda v_\lambda,$$

in which mR is the diagonal matrix

$$mR = \mathrm{diag}[m(1)R(1), m(2)R(2), \dots, m(n)R(n)].$$

We have seen in three examples that for the change of measure (on sub-filtration $(\mathcal{F}_t)_{t\geq 0}$) to have $Z_\lambda(t)$ as its Radon-Nikodym derivative, we should expect the new measure to introduce three changes:

- it should *affect the motion* of the spine (ξ_t, η_t) in some way;
- it should *increase the rate* at which the fissions occur on the spine;
- it should cause the distributions of families produced from the spine to be *size-biased.*

These three features will be brought about by three martingales.

Lemma 3.12

$$v_\lambda(\eta_t)\, e^{\int_0^t mR(\eta_s)\,\mathrm{d}s} e^{\lambda \xi_t - E_\lambda t}$$

is a P-martingale that will introduce a drift to the spine's spatial motion ξ_t and will change the Q-matrix (generator) of the spine's type motion η_t.

Lemma 3.13

$$e^{-\int_0^t mR(\eta_s)\,\mathrm{d}s} \prod_{v<\xi_t} \big(1 + m(\eta_{S_v})\big)$$

is a P-martingale that will increase the rate at which fission times occur on the spine from $R(\eta_t)$ to $(1+m(\eta_t))R(\eta_t)$.

Lemma 3.14

$$\prod_{v<\xi_t} \frac{1 + A_v}{1 + m(\eta_{S_v})}$$

is a P-martingale that will cause the family distribution on the spine to be size-biased to the distribution

$$Prob(\tilde{A} = i) = \frac{(i+1)p_i}{m+1}, \qquad i \in \{0, 1, \dots\}.$$

Therefore we change the measure $\tilde{P}$ by the product of these three martingales – for which some of the terms cancel:

Definition 3.15 *For each $\lambda \leq 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via*

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\bigg|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}} \prod_{v<\xi_t} (1 + A_v) \times v_\lambda(\eta_t) e^{\lambda \xi_t - E_\lambda t}. \tag{28}$$

Once again, a proof using the conditional expectation of this measure-change martingale confirms:

Theorem 3.16 *If we define $\mathbb{Q}_\lambda^{x,y} := \tilde{\mathbb{Q}}_\lambda^{x,y}|_{\mathcal{F}_\infty}$, then $\mathbb{Q}_\lambda^{x,y}$ is a measure on $\mathcal{F}_\infty$ and*

$$\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\bigg|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

The geometry of E_λ is not significantly changed by the introduction of random offspring numbers, although $\tilde{\lambda}(\theta)$ will probably now be different. In fact, as far as E_λ is concerned, the introduction of random offspring numbers acts just like a multiplying term on the birth rates $R(y)$.

Theorem 3.17 *Suppose that $\lambda \leq 0$.*

1. *For every $p \in (1, 2]$ the martingale Z_λ is $\mathcal{L}^p$-bounded provided that $pE_\lambda - E_{p\lambda} > 0$.*
2. *Almost surely (under P), $Z_\lambda(\infty) = 0$ if $\lambda < \tilde{\lambda}(\theta)$.*

Proof of part 1: We quickly cover the main points since they are very similar to the previous proof for the binary-splitting case. The spine decomposition leads us to:

$$P^{x,y}(Z_\lambda(t)^p) \leq \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) + \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big),$$

By conditioning on which nodes make up the spine ξ and on the spine's motion (information in $\mathcal{G}_t$), we can replace the term A_u^q by its expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\big|\xi, \mathcal{G}_\infty\Big) &= \sum_{u<\xi_t} \tilde{\mathbb{Q}}_\lambda^{x,y}(A_u^q) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u} \\ &= \sum_{u<\xi_t} M_q(\eta_{S_u}) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}.\end{aligned}$$

Taking expectations again,

$$\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) = \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} M_q(\eta_{S_u}) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big).$$

The remainder of the proof now goes through as before when we write this sum as an integral, use Fubini's theorem and change the measure like at (24):

$$\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) \leq \Big(\max_{i\in I} M_q(i)\Big)\ K(p, \lambda, x, y) \int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\, \mathrm{d}s,$$

where

$$K(p, \lambda, x, y) := e^{q\lambda x} \frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \sup_{s\geq 0} \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(2R(\eta_s)\, \frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) < \infty$$

as at (26). Thus if $p > 1$ is such that $P(A(w)^p) < \infty$ for all $w \in I$ we know that $M_q(w) < \infty$ also, whence $\max_{i\in I} M_q(i) < \infty$, and the growth of the sum term is once again dependent on the term $e^{-(pE_\lambda - E_{p\lambda})t}$, completing the proof of part 1. □

Proof of part 2: The fact that the geometry interprets random offspring numbers much as it would handle an increase in the birth rates is plain to see in the way that the Q-matrix Q_λ is defined:

Lemma 3.18 *Under $\tilde{\mathbb{Q}}_\lambda$ the spine process (ξ_t, η_t) has generator:*

$$\mathcal{H}_\lambda F\,(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \sum_{j\in I}\theta Q_\lambda(y,j)F(x,j), \tag{29}$$

where Q_λ is an honest Q-matrix:

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \neq j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + m(i)r(i) & \text{if } i = j \end{cases}$$

with invariant measure $\pi_\lambda = v_\lambda^2\pi$.

The proof for the binary-splitting model easily adapts:

$$\limsup_{t\to\infty} Z_\lambda(t) > \limsup_{t\to\infty} v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)} = \infty, \quad \text{if } \lambda < \tilde{\lambda}(\theta).$$

□

4 An continuous-typed branching-diffusion

The previous finite-type model was originally inspired by the model that we now turn to, originally laid out in Harris and Williams [11]. In this model the type moves on the real line as an Orstein-Uhlenbeck process associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{with } \theta > 0 \text{ considered as the } \textit{temperature},$$

which has the standard normal density as its invariant distribution:

$$\phi(y) := (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}y^2}.$$

The spatial movement of a particle of type y is a driftless brownian motion with instantaneous variance

$$A(y) := ay^2, \qquad \text{for some fixed } a > 0,$$

and fission of a particle of type y occurs at a rate

$$R(y) := ry^2 + \rho, \qquad \text{where } r, \rho > 0 \text{ are fixed},$$

to produce two particles at the same type-space location as the parent (we consider only binary splitting). The model has very different behaviour for low temperature values (i.e. low θ), but most studies have considered the high temperature regime where $\theta > 8r$. Also, the parameter λ must be restricted to an interval $(\lambda_{\min}, 0)$ in order for some of the model's parameters to remain in $\mathbb{R}$, where

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}.$$

Generally, *unboundedness* in a model's rates is a serious obstacle to classical proofs since they often depend on the expectation semigroup of the branching process, and unbounded rates tend to lead to unbounded *eigenfunctions*. Here this is the case, but the existence of a spectral theory for their particular expectation operator allowed Harris and Williams to get a sufficiently good bound in particular for a non-linear term (see Theorem 5.1 of [11]), and therefore to prove $\mathcal{L}^p$-convergence of the martingale.

We use the same notation as previously $\mathbb{X}_t = \big\{\big(X_u(t), Y_u(t)\big) : u \in N_t\big\}$ to denote the point process of space-type locations in $\mathbb{R}\times\mathbb{R}$, and suppose that the measures $\big\{\tilde{P}^{x,y} : (x,y) \in \mathbb{R}^2\big\}$ on the natural filtration with a spine $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ are such that the initial ancestor starts at (x,y) and $\big(\mathbb{X}_t, (\xi_t, \eta_t)\big)$ becomes the above-described branching diffusion with a spine.

4.1 The measure change

Although there are some significant differences, this model is similar in flavour to our finite-type model – it was in fact the inspiration for that finite model. There is a strictly-positive martingale Z_λ^- defined as

$$Z_\lambda^-(t) := \sum_{u \in N_t} v_\lambda(Y_u(t)) e^{\lambda X_u(t) - E_\lambda^- t}$$

where v_λ^- and E_λ^- are the eigenvector and eigenvalue associated with the self-adjoint (in $\mathcal{L}^2(\phi)$) operator:

$$Q_\theta + \frac{1}{2}\lambda^2 A(y) + R(y).$$

The eigenfunction v_λ is normalizable against the $\mathcal{L}^2(\phi)$ norm, and can be found explicitly as

$$v_\lambda(y) = e^{\psi_\lambda^- y^2}$$

where

$$\psi_\lambda^- := \frac{1}{4} - \frac{\mu_\lambda}{2\theta}, \quad \mu_\lambda := \frac{1}{2}\sqrt{\theta^2 - \theta(8r + 4a\lambda^2)},$$

are both positive for all $\lambda \in (\lambda_{\min}, 0)$; another important parameter is $\psi_\lambda^+ := \frac{1}{4} + \frac{\mu_\lambda}{2\theta}$. The eigenvalue E_λ^- is then given by

$$E_\lambda^- = \rho + \theta\psi_\lambda^-.$$

We again define the speed function $c_\lambda^- := -E_\lambda^-/\lambda$, and $\tilde{\lambda}(\theta) < 0$ is the unique point (on the negative axis) at which c_λ^- hits its minimum – further details are given in Harris and Williams [11]. We are going to use spines to prove:

Theorem 4.1 *Suppose that* $\lambda \in (\lambda_{\min}, 0)$.

1. *Let* $p \in (1, 2]$. *The martingale* Z_λ *is* $\mathcal{L}^p$*-bounded if both* $pE_\lambda - E_{p\lambda} > 0$ *and* $p\psi_\lambda^- < \psi_{p\lambda}^+$.
2. *Almost surely (under* P*),* $Z_\lambda(\infty) = 0$ *if* $\lambda < \tilde{\lambda}(\theta)$.

Once again, for each $\lambda \leq 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}} 2^{n_t} v_\lambda(\eta_t) e^{\lambda\xi_t - E_\lambda^- t}, \tag{30}$$

so that with $\mathbb{Q}_\lambda := \tilde{\mathbb{Q}}_\lambda|_{\mathcal{F}_\infty}$ we have

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

The facts are that under $\tilde{\mathbb{Q}}_\lambda$:

- the spine diffusion ξ_t has instantaneous drift $a\eta_t^2\lambda$;
- the type process η_t has generator $\frac{\theta}{2}\big(\frac{\partial^2}{\partial y^2} - \frac{2\mu_\lambda}{\theta} y \frac{\partial}{\partial y}\big)$ and invariant measure $N(0, \frac{\theta}{2\mu_\lambda})$;
- fission times on the spine occur at the accelerated rate of $2R(\eta_t)$;
- all particles not in the spine behave as if under the original measure P.

4.2 Proof of Theorem 4.1

Proof of Part 1: Suppose $p \in (1, 2]$. Then using the spine decomposition with Jensen's inequality and Proposition 2.15 we find,

$$P^{x,y}(Z_\lambda^-(t)^p) \le \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) + \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big).$$

The spine term. On the algebra $\mathcal{G}_t$ the change of measure takes the form

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\bigg|_{\mathcal{G}_t} = v_\lambda(y)^{-1} e^{-\lambda x}\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_s) e^{\lambda\xi_t - E_\lambda^- t},$$

which we can use on the spine term to arrive at

$$\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big) = e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) e^{-(pE_\lambda - E_{p\lambda})t}. \tag{31}$$

The sum term. As for the finite-type model the fission times S_u on the spine occur as a Cox process and therefore

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) &= \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)\, v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s \\ &= e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)}\int_0^t \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(2R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s.\end{aligned}$$

Bringing together the results for the sum and spine terms, we have an upper bound

$$\begin{aligned}P^{x,y}(Z_\lambda(t)^p) \le e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \Big\{&\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) e^{-(pE_\lambda - E_{p\lambda})t} \\ &+ \int_0^t \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(2R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s\Big\}.\end{aligned} \tag{32}$$

Once again the role of $pE_\lambda - E_{p\lambda} > 0$ is clear, but the new condition $p\psi_\lambda^- < \psi_{p\lambda}^+$ in part 1 of Theorem 4.1 is due to the term $\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\big(R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\big)$, which could potentially be unbounded.

Harris and Williams have shown in [11] that under $\tilde{\mathbb{Q}}_{p\lambda}^{0,y}$, the random variable η_s has the normal distribution

$$\mathrm{N}\Big(e^{-\mu_{p\lambda}s}y, \frac{\theta(1 - e^{-2\mu_{p\lambda}s})}{2\mu_{p\lambda}}\Big) =: \mathrm{N}(\alpha, \beta)$$

and from the known form of the eigenfunctions we have the explicit result:

$$\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(R(\eta_s)\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) = \frac{1}{\sqrt{2\pi\alpha^2}}\int_{-\infty}^{\infty}(ru^2 + \rho)e^{(p\psi_\lambda^- - \psi_{p\lambda}^-)u^2} e^{-(u^2-\alpha)/2\beta}\,\mathrm{d}u. \tag{33}$$

which will be finite and bounded for all $s \ge 0$ if $p\psi_\lambda^- - \psi_{p\lambda}^- - \frac{\mu_{p\lambda}}{\theta} < 0$; here the term $\frac{\mu_{p\lambda}}{\theta}$ comes from the term $(2\beta)^{-1}$ in the distribution of η_s. But

$$p\psi_\lambda^- - \psi_{p\lambda}^- - \frac{\mu_{p\lambda}}{\theta} < 0 = p\psi_\lambda^- - \psi_{p\lambda}^+$$

and therefore the condition $p\psi_\lambda^- < \psi_{p\lambda}^+$ ensures that the expectation (33) stays bounded for all $s \geq 0$. Thus we deduce that with $p \in (1,2]$ and $\lambda \in (\lambda_{\min}, 0)$,

$$pE_\lambda - E_{p\lambda} < 0 \text{ and } p\psi_\lambda^- < \psi_{p\lambda}^+ \quad \text{implies} \quad P^{x,y}(Z_\lambda(t)^p) \text{ is bounded for all } t \geq 0,$$

□

Proof of Part 2: The proof that we have seen in the finite-type model will work here unchanged, since under $\tilde{\mathbb{Q}}_\lambda$ the spatial motion is

$$\xi_t = B\left(\int_0^t a(\eta_s)\mathrm{d}s\right) + \lambda a \int_0^t \eta_s^2 \mathrm{d}s,$$

and the type process η_s has invariant distribution $N(0, \frac{\theta}{2\mu_\lambda})$, whence

$$t^{-1}\xi_t \to \lambda a\theta/\mu_\lambda = E'_\lambda.$$

Therefore it follows that under $\tilde{\mathbb{Q}}_\lambda$ the diffusion $\xi_t + c_\lambda t$ drifts off to $-\infty$ if $\lambda < \tilde{\lambda}(\theta)$ whence

$$\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\limsup_{t\to\infty} Z_\lambda(t) = \infty\Big) = 1.$$

□

References

[1] Søren Asmussen and Heinrich Hering, *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 3, 195–212. MR 54 #8900

[2] A. Champneys, S. Harris, J. Toland, J. Warren, and D. Williams, *Algebra, analysis and probability for a coupled system of reaction-diffusion equations*, Philosophical Transactions of the Royal Society of London **350** (1995), 69–112.

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

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

[5] Brigitte Chauvin and Alain Rouault, *KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees*, Probab. Theory Related Fields **80** (1988), no. 2, 299–314. MR 90b:60113

[6] Richard Durrett, *Probability: Theory and examples*, 2nd ed., Duxbury Press, 1996.

[7] János Engländer and Andreas E. Kyprianou, *Markov branching diffusions: martingales, Girsanov-type theorems and applications to the long-term behaviour*, preprint, 2001.

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

[9] ———, *A spine proof of a large-deviations principle for branching Brownian motion*, (2004), no. 0407, Mathematics Preprint, University of Bath.

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

[11] 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

[12] Simon C. Harris, *Branching diffusions*, Ph.D. thesis, University of Cambridge DPMMS, 1995.

[13] ______, *Travelling-waves for the FKPP equation via probabilistic arguments*, Proceedings of the Royal Society of Edinburgh **129A** (1999), 503–517.

[14] Simon C. Harris and Yoav Git, *Large deviations and martingales for a typed branching diffusion: II*, University of Bath Preprint, 2001.

[15] J.W.Harris, S.C.Harris, and A.E.Kyprianou, *Further probabilistic analysis of the Fisher-Kolmogorov-Petrovskii-Piscounov equation: one-sided travelling waves*, (2003), no. 0403, Mathematics Preprint, University of Bath.

[16] O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, 2002.

[17] Erwin Kreyszig, *Introductory functional analysis with applications*, John Wiley and Sons, 1989.

[18] 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

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

[20] 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

[21] 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

[22] 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

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

[24] 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.

[25] E. Seneta, *Non-negative matrices and Markov chains*, Springer-Verlag, 1981.