Exponential growth rates & large deviations for a typed branching diffusion

Y. Git*, J. W. Harris† and S. C. Harris‡

February 26, 2004

Abstract

We find the almost sure rate of exponential growth of particles, $D(\gamma, \kappa)$, which are found simultaneously with spatial positions near γt and type positions near $\kappa\sqrt{t}$ at large times t in the high temperature phase of a typed branching diffusion initially studied in Harris and Williams (1996). Our proofs rest on the study of certain martingales associated with the branching diffusion and a change of measure inducing a spine decomposition.

Contents

*Yoav Git, 22 Mill Street, Cambridge, CB1 2HP. email: Y.Git@statslab.cam.ac.uk
†Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, U.K.. email: j.w.harris@bath.ac.uk
‡*Correspondence:* Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, U.K.. email: s.c.harris@bath.ac.uk

1 Introduction and summary

We shall start this introductory section by describing the family of typed branching diffusions that form the primary study of this paper. Next, some results concerned with the almost sure rate of growth of particles in the spatial dimension only shall be considered before progressing onto the main result which gives the almost sure rates of growth for particles which have both large spatial and large type positions at large times. As a corollary, we gain the almost sure asymptotic shape of the branching diffusion when we view the evolution of particles in the space-type plane and scale appropriately. In all the situations considered, we will see that the 'almost sure' and 'expected' rates of growth coincide exactly on the areas where particles are increasing in numbers. Our proofs of these results are entirely martingale-based, and we make use of a technique that has become popular in the branching process literature recently: a change of measure to known as a 'spine decomposition' of the process.

In the rest of this introductory section we describe the branching model and summarise the main results of the paper. Thereafter, in section 2 we outline how the expected rates of growth may be calculated using the 'many-to-one lemma' of Harris and Williams (1996) and changes of measure for Ornstein-Uhlenbeck process and and drifting Brownian motion. Section 3 contains a heuristic argument that gives a conceptually simple picture to explain why we should expect the almost sure rate of particle growth at large space and type positions should be *at least* $D(\gamma, \kappa)$. In section 4 we give details of the spine decomposition, and then in section 6 we apply this to prove the main step in a lower bound on the asymptotic shape and growth. In sections 7 through 9 we prove the remainder of the necessary martingale results and the almost sure growth results for the branching diffusion. In the final sections we give heuristic derivations of some large deviation results that are applicable to a more general class of branching diffusions; in future work we hope to make these results rigorous by developing further spine arguments.

1.1 The branching model

The typed branching diffusion we consider in this paper has particles which independently move in space according to a Brownian motion with variance controlled by the particle's type process. The type of each particle evolves as an Ornstein-Uhlenbeck process and this type also controls the rate at which births occur. This model was initially studied in Harris and Williams (1996), a paper which forms the foundations for this work.

For time $t \geq 0$; $N(t)$ is the number of particles alive, $X_k(t) \in \mathbb{R}$ is the spatial position of the k^{th}-born particle and $Y_k(t) \in \mathbb{R}$ is the *type* of the k^{th}-born particle. The current state of the particle system is

$$\big(N(t); X_1(t), \ldots, X_{N(t)}; Y_1(t), \ldots, Y_{N(t)}\big) .$$

The type moves on the real line as an Ornstein-Uhlenbeck process associated with the differential operator (generator)

$$\mathcal{Q}_\theta := \frac{\theta}{2}\left(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\right),$$

where θ is a positive real parameter considered as the *temperature* of the system. The spatial motion of a particle of type y is a driftless Brownian motion with variance

$$A(y) := ay^2, \qquad \text{where } a \geq 0.$$

A particle of type y particle undergoes dyadic splitting at a rate

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

i.e. we have one child born at these times. A child inherits its parent's current type and spatial position, and then moves off independently of all others. Particles live forever.

The model has a very different behaviour for low temperature parameter values – see Harris and Williams (1996) – and throughout this paper we consider only values above the critical temperature, that is $\theta > 8r$. All the above parameters of the model are considered as fixed for the rest of this paper, unless otherwise stated. We use $P^{x,y}$ and $E^{x,y}$ with $x, y \in \mathbb{R}$ to represent probability and expectation when the Markov process starts with an initial state $(N; \mathbf{X}; \mathbf{Y}) = (1; x; y)$.

The model was first suggested in Champneys et al. (1995), which studied a typed branching Brownian motion where the type of each particle moved as a two state Markov chain. Analogous martingale techniques were, of course, also fundamental in the probabilistic study of the two-type model. For the inspiration of this work we are in debt to, amongst many others, the initial paper McKean (1975); also work of Biggins (1977), Biggins (1979), and Biggins (1992); and the martingales techniques in Neveu (1988) concerning standard branching Brownian motion. Branching Brownian motions have been long studied and the reader will find many other very important references in the above papers.

The family of models we mainly consider in this paper have some features of fundamental significance, in particular, quadratic breeding is a critical rate in terms of explosions in the population of particles in a branching Brownian motion (see, for example, Itô and McKean (1965)). The Ornstein-Uhlenbeck type motion has exactly the right drift to help counteract the quadratic breeding rates: for large enough temperatures, there is enough ergodicity to keep the particles well behaved but, for low temperatures, the large breeding rate outweighs the pull back toward the origin and the particles behave very differently. These features make the models both interesting in their own right, as well as difficult to study. We hope to consider the low-temperature regime in future works. Although the models in this paper are quite specific, many of the ideas and similar methods will be applicable to yield analogous results in a variety of other typed branching diffusions where a spatial Brownian motion and breeding rates are controlled by a type process moving as an finite state Markov chain or sufficiently ergodic Markov process. We emphasise that study of this family of models motivated some much more general path large deviation conjectures for branching diffusions that are presented in section 10.

1.2 The asymptotic growth rate of particles along spatial rays.

We shall first present a result concerned with the growth rates of particles with large spatial positions at large times. This shall prove to be a useful stepping stone towards a result concerned with growth rates of particles which gain both large spatial and large type positions.

Firstly, suppose we wish to find the *almost sure* rate of growth in numbers of particles along rays in space-time. For $\gamma \geq 0$, define

$$N_t(\gamma) := \sum_{k=1}^{N(t)} \mathbf{1}\{X_k(t) \leq -\gamma t\} \tag{1}$$

The limit giving the *expected* rate of growth,

$$\lim_{t\to\infty} t^{-1} \log E(N_t(\gamma)) \tag{2}$$

can be shown to exist and its value can be calculated to be

$$\Delta(\gamma) = \rho + \frac{\theta}{4} - \frac{1}{4}\sqrt{\theta(\theta - 8r)\left(1 + 4\gamma^2/(\theta a)\right)}. \tag{3}$$

An outline for this calculation is given in the next section. A discussion of this result involving large deviation theory for occupation densities is also presented in Harris and Williams (1996).

It is now tempting to guess that the asymptotic speed of the spatially left-most particle, $\tilde{c}(\theta)$, is given by

$$\begin{aligned} \tilde{c}(\theta) &:= \sup\{\gamma : \Delta(\gamma) > 0\} \\ &= \sqrt{2a\left(r+\rho+\frac{2(2r+\rho)^2}{\theta-8r}\right)}. \end{aligned} \tag{4}$$

In this particular situation, this guess that 'expectation' and 'particle' wave-speeds agree was proved rigorously using a martingale change of measure technique in Harris and Williams (1996). In this paper, we extend this connection and prove that the 'expected' and 'almost sure' rates of growth of particles actually agree.

Theorem 1 *Under each $P^{x,y}$ law, the limit*

$$D(\gamma) := \lim_t t^{-1} \log N_t(\gamma)$$

exists almost surely and is given by

$$D(\gamma) = \begin{cases} \Delta(\gamma) & \text{if } 0 \leq \gamma < \tilde{c}(\theta), \\ -\infty & \text{if } \gamma \geq \tilde{c}(\theta). \end{cases}$$

To prove this result, we will make use of some fundamental additive martingales for the model and a related convergence theorem found in Harris (2000).

We gain the asymptotic speed of the spatially most extreme particle as a corollary to this theorem – an alternative proof of this was given in Harris and Williams (1996) using a martingale change of measure technique:

Corollary 2

$$\lim_{t\to\infty} t^{-1} \sup_{k\leq N(t)} X_k(t) = \tilde{c}(\theta).$$

1.3 The asymptotic shape and growth of the branching diffusion

Next, we wish to give the almost sure rate of growth of particles which are in the vicinity of γt in space and near $\sqrt{\kappa^2 t}$ in type position at large times t.

For $\gamma, \kappa \geq 0$, consider

$$N_t(\gamma,\kappa) := \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \leq -\gamma t, Y_k^2(t) \geq \kappa^2 t\right\}. \tag{5}$$

It can be shown that the limit

$$\lim_{t\to\infty} t^{-1} \log E\left(N_t(\gamma,\kappa)\right) \tag{6}$$

exists and takes the value

$$\Delta(\gamma,\kappa) = \rho + \frac{(\theta-\kappa^2)}{4} - \frac{1}{4\theta a}\sqrt{\theta(\theta-8r)(4a\theta\gamma^2 + a^2(\theta+\kappa^2)^2)}. \tag{7}$$

An outline of this expectation calculation is given in the next section.

Once again, this 'almost sure' rate agrees with the 'expected' rate of growth of particles exactly where there is *growth* in expected particle numbers.

Theorem 3 *Under each $P^{x,y}$ law, the limit*

$$D(\gamma,\kappa) := \lim_{t\to\infty} t^{-1} \log N_t(\gamma,\kappa)$$

exists almost surely and is given by

$$D(\gamma,\kappa) = \begin{cases} \Delta(\gamma,\kappa) & \text{if } \Delta(\gamma,\kappa) > 0, \\ -\infty & \text{otherwise.} \end{cases} \tag{8}$$

To prove the upper bound on the growth rate, we will require knowledge of the exact rate of convergence of some other additive martingales. To prove the trickier lower bound we will exhibit an explicit mechanism by which the branching diffusion can build up at least the required exponential number of particles near to γt in space and $\sqrt{\kappa^2 t}$ in type position by large times t. This mechanism involves certain particles spending almost all of their time building up in sufficient numbers whilst gaining a proportion of the required spatial position, then a small number of this group of particles will succeed in making a very rapid ascent out to the required final position. To prove this mechanism, we combine knowledge of the growth rate $\Delta(\gamma)$ with a lower bound on the probability that a particular particle in the branching diffusion makes the steep ascent. To prove this lower bound we make use of the spine decomposition.

After proving Theorem 3, it becomes straightforward to retrieve the following theorem:

Theorem 4 *For any $F \subset \mathbb{R}^2$, define*

$$N_t(F) := \sum_{k=1}^{N(t)} \mathbf{1}\left\{\left(\frac{X_k(t)}{t}, \frac{Y_k(t)}{\sqrt{t}}\right) \in F\right\}.$$

If $B \subset \mathbb{R}^2$ is any open set and $C \subset \mathbb{R}^2$ is any closed set, then almost surely under any $P^{x,y}$

$$\liminf_{t\to\infty} \frac{1}{t} \log N_t(B) \geq \sup_{(\gamma,\kappa)\in B} D(\gamma, \kappa),$$
$$\limsup_{t\to\infty} \frac{1}{t} \log N_t(C) \leq \sup_{(\gamma,\kappa)\in C} D(\gamma, \kappa),$$

with the growth rate given by

$$D(\gamma, \kappa) = \begin{cases} \Delta(\gamma, \kappa) & \text{if } \Delta(\gamma, \kappa) > 0, \\ -\infty & \text{otherwise}, \end{cases}$$

where

$$\Delta(\gamma, \kappa) = \rho + \frac{(\theta - \kappa^2)}{4} - \frac{1}{4\theta a}\sqrt{\theta(\theta - 8r)(4a\theta\gamma^2 + a^2(\theta + \kappa^2)^2)}$$

As a simple corollary, we now know the almost sure asymptotic shape of the region that the particles in the branching diffusion will occupy after any long period of time.

Corollary 5 *Let $B \subset \mathbb{R}^2$ be any open set. Almost surely, under each $P^{x,y}$ law,*

$$N_t(B) \to \begin{cases} 0 & \text{if } \mathcal{S} \cap B = \emptyset \\ +\infty & \text{if } \mathcal{S} \cap B \neq \emptyset, \end{cases}$$

where $\mathcal{S} \subset \mathbb{R}^2$ is the closed set given by

$$\mathcal{S} := \left\{(\gamma, \kappa) \in \mathbb{R}^2 \middle| \Delta(\gamma, \kappa) \geq 0\right\}.$$

The next section discusses how the various expectation calculations may be performed. In the process, we shall start to gain valuable intuition into how particles within the branching diffusion behave as well as giving hints on the important martingales and large deviation techniques that we will eventually use to prove the almost sure growth rate results.

2 Some expectation calculations

In this section, we outline how the expectation calculations referred to in the previous section can be performed by using the 'many-to-one particle' lemma and changes of measure for drifting Brownian motions and Ornstein-Uhlenbeck processes. See Harris and Williams (1996) for further details of these techniques, or see Rogers and Williams (2000*a*) and Rogers and Williams (2000*b*) for the more general theory.

First, we give some key definitions. Let

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

Let $\lambda \in \mathbb{R}$, with the following convention which we *always* use for λ:

$$\lambda_{\min} < \lambda < 0. \tag{10}$$

Also, define

$$\mu_\lambda := \frac{1}{2}\sqrt{\theta(\theta - 8r - 4a\lambda^2)} \qquad \psi_\lambda^\pm := \frac{1}{4} \pm \frac{\mu_\lambda}{2\theta} \qquad E_\lambda := \rho + \theta\psi_\lambda^- . \tag{11}$$

(Note that $\lambda_{\min}$ is the point beyond which ψ_λ^- is no longer a real number).

For simplicity of notation, we will assume throughout that we have one particle starting at the origin in both space and type at time zero, unless otherwise stated. Using the one particle picture and changing measure to alter the drift of Brownian motion, we see that

$$\begin{aligned}
E\sum_{k=1}^{N(t)} f\big(X_k(t), Y_k(t)\big) &= \mathbb{E}_{\frac{\theta}{2},0}\left\{\exp\left(\int_0^t R(\eta_s)\,\mathrm{d}s\right) f(\xi_t, \eta_t)\right\} \\
&= \mathbb{E}_{\frac{\theta}{2},0}\left\{\exp\left(-\lambda\xi_t + \int_0^t \left\{R(\eta_s) + \frac{\lambda^2}{2}A(\eta_s)\right\}\,\mathrm{d}s\right) f(\xi_t, \eta_t) e^{\lambda\xi_t - \frac{\lambda^2}{2}\int_0^t A(\eta_s)\,\mathrm{d}s}\right\} \\
&= \mathbb{E}_{\frac{\theta}{2},\lambda}\left\{\exp\left(-\lambda\xi_t + \int_0^t \left\{R(\eta_s) + \frac{\lambda^2}{2}A(\eta_s)\right\}\,\mathrm{d}s\right) f(\xi_t, \eta_t)\right\}
\end{aligned}$$

where $\xi_t = B\left(\int_0^t A(\eta_s)\,\mathrm{d}s\right)$ and, under $\mathbb{P}_{\mu,\lambda}$, B is a Brownian motion with drift λ and η is an Ornstein Uhlenbeck process with variance θ and drift μ. To perform a further change of measure on the OU process to get rid of the time integrals in the exponential of the expectation, we recall that

$$\left.\frac{\mathrm{d}\mathbb{P}_{\mu_\lambda,\cdot}}{\mathrm{d}\mathbb{P}_{\frac{\theta}{2},\cdot}}\right|_{\mathcal{F}_t} = M_t^{\mu_\lambda,\frac{\theta}{2}} := \exp\left(\psi_\lambda^-\eta_t^2 - E_\lambda t + \int_0^t \left\{R(\eta_s) + \frac{1}{2}\lambda^2 A(\eta_s)\right\}\,\mathrm{d}s\right)$$

and then

$$\begin{aligned}
E\sum_{k=1}^{N(t)} f\big(X_k(t), Y_k(t)\big) &= \mathbb{E}_{\frac{\theta}{2},\lambda}\left\{\exp\left(-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t\right) f(\xi_t, \eta_t). M_t^{\mu_\lambda,\frac{\theta}{2}}\right\} \\
&= \mathbb{E}_{\mu_\lambda,\lambda}\left\{\exp\left(-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t\right) f(\xi_t, \eta_t)\right\}.
\end{aligned} \tag{12}$$

2.1 The expected rate of growth along spatial rays

We now give the outline of some calculations to find the rate of growth of particles near $-\gamma t$ in space at time t. Using the formula from (12), we find

$$\begin{aligned}
E\sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \sim -\gamma t\right\} &= \mathbb{E}_{\mu_\lambda,\lambda}\left(e^{-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t}\mathbf{1}\left\{\xi_t \sim -\gamma t\right\}\right) \\
&\sim e^{(E_\lambda + \lambda\gamma)t}\mathbb{E}_{\mu_\lambda,\lambda}\left(e^{-\psi_\lambda^-\eta_t^2}; \xi_t \sim -\gamma t\right).
\end{aligned}$$

We can now define

$$\Delta(\gamma) := \inf_{\lambda \in (\lambda_{\min}, 0)} \left\{E(\lambda) + \lambda\gamma\right\} \tag{13}$$

and this infimum is attained at λ_γ when $E'(\lambda_\gamma) = -\gamma$, hence

$$\lambda_\gamma = -\sqrt{\frac{(\theta - 8r)\gamma^2}{\theta a^2 + 4a\gamma^2}} \tag{14}$$

and then

$$\Delta(\gamma) = \rho + \left(\theta - \sqrt{a^{-1}(\theta - 8r)(4\gamma^2 + \theta a)}\right)/4.$$

Of course, choosing this minimising λ_γ value also means that we have maximised the expectation

$$\mathbb{E}_{\mu_\lambda,\lambda}\left(e^{-\psi_\lambda^- \eta_t^2}; \xi_t \sim -\gamma t\right).$$

To confirm that this maximal value is of a reasonable size (in particular, we don't want any exponentially small value) is now easy. Under $\mathbb{P}_{\mu_\lambda,\lambda}$, η is an OU with an invariant measure given by the probability density, ϕ_λ, of a normal distribution with mean zero and variance $\theta/(2\mu_\lambda)$ and, also, $\xi_t = B\left(\int_0^t A(\eta_s)\,\mathrm{d}s\right)$ where B is a BM with drift λ. Almost surely under $\mathbb{P}_{\mu_\lambda,\lambda}$,

$$\frac{\xi_t}{t} \to \lambda \int_{\mathbb{R}} A(y)\phi_\lambda(y)\,\mathrm{d}y = \frac{\lambda a\theta}{2\mu_\lambda} = E'(\lambda)$$

and so when we use the optimal λ_γ value we get exactly the desired drift since $E'(\lambda_\gamma) = -\gamma$. Then,

$$\mathbb{E}_{\mu_{\lambda_\gamma},\lambda_\gamma}\left(e^{-\psi_{\lambda_\gamma}^- \eta_t^2}; \xi_t \sim -\gamma t\right) \to \lim_{t\to\infty} \mathbb{E}_{\mu_{\lambda_\gamma},\lambda_\gamma}\left(e^{-\psi_{\lambda_\gamma}^- \eta_t^2}\right) = \int_{\mathbb{R}} e^{-\psi_{\lambda_\gamma}^- y^2}\phi_{\lambda_\gamma}(y)\,\mathrm{d}y$$

Hence, we have obtained the exact rate of exponential growth for the expectation,

$$\lim_{t\to\infty} t^{-1}\log E(N_t(\gamma)) = \Delta(\gamma). \tag{15}$$

The changes of measure used above are actually suggesting a lot about the dominant particles that are found in the vicinity of a given ray in space. To help get a clearer intuitive picture of the evolution of the branching diffusion see Harris and Williams (1996) where there are discussions involving large deviations for the occupation density of the type process (as well as a dual calculation which is closer to the above arguments).

2.2 The expected asymptotic shape

We give a rough outline of calculations that yield the correct exponential growth in the expected number of particles both near $-\gamma t$ in space *and* $\sqrt{\kappa^2 t}$ in type at large times t. Once again, using the formula from (12),

$$\begin{aligned} E\sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \sim -\gamma t; Y_k(t)^2 \geq \kappa^2 t\right\} \\ &= \mathbb{E}_{\mu_\lambda,\lambda}\left(e^{-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t}\mathbf{1}\left\{\xi_t \sim -\gamma t; \eta_t^2 \geq \kappa^2 t\right\}\right) \\ &\sim e^{\left(E_\lambda + \lambda\gamma - \kappa^2\psi_\lambda^-\right)t}\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t; \eta_t^2 \geq \kappa^2 t\right) \end{aligned}$$

Now, continuing with our approximate calculations,

$$\begin{aligned} \mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t; \eta_t^2 \geq \kappa^2 t\right) &= \mathbb{P}_{\mu_\lambda,\lambda}\left(\eta_t^2 \geq \kappa^2 t\right)\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t \middle| \eta_t^2 \geq \kappa^2 t\right) \\ &\sim e^{-\frac{\mu_\lambda}{\theta}\kappa^2 t}\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t \middle| \eta_t^2 \geq \kappa^2 t\right) \end{aligned}$$

and recall $\psi_\lambda^- + \mu_\lambda/\theta = \psi_\lambda^+$, yielding

$$\begin{aligned} &E\sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \sim -\gamma t; Y_k(t)^2 \geq \kappa^2 t\right\} \sim \\ &\qquad e^{\left(E_\lambda + \lambda\gamma - \kappa^2\psi_\lambda^+\right)t}\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t \,\middle|\, \eta_t^2 \geq \kappa^2 t\right). \end{aligned} \tag{16}$$

Define

$$\begin{aligned} \Delta(\gamma,\kappa) &:= \inf_{\lambda\in(\lambda_{\min},0)}\left\{E_\lambda + \lambda\gamma - \kappa^2\psi_\lambda^+\right\} \\ &= \rho + (\theta - \kappa^2)/4 - \sqrt{\theta(\theta - 8r)(4a\theta\gamma^2 + a^2(\theta+\kappa^2)^2)}/4\theta a \end{aligned} \tag{17}$$

where the infimum is attained with a λ value of

$$\bar{\lambda}(\gamma,\kappa) = -\sqrt{\frac{\gamma^2\theta(\theta-8r)}{a^2(\kappa^2+\theta)^2+4a\gamma^2\theta}} \in (\lambda_{\min},0) \tag{18}$$

Then equation (16) certainly leads to the inequality

$$\limsup_{t\to\infty} t^{-1}\log\sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t)\sim -\gamma t; Y_k(t)^2 \geq \kappa^2 t\right\} \leq \Delta(\gamma,\kappa)$$

It is clear from equation (16) that when minimising $E_\lambda + \lambda\gamma - \kappa^2\psi_\lambda^+$, we simultaneously maximise the probability $\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t \sim -\gamma t \,\middle|\, \eta_t^2 \geq \kappa^2 t\right)$. In fact, this probability must reach a value well away from 0 when we choose the optimal parameter for λ.

It can be shown that

$$\mathbb{P}_{\mu_{\bar\lambda},\bar\lambda}\left(\xi_t\sim -\gamma t; \eta_t^2\geq \kappa^2 t\right) \sim \exp\left(-\frac{\mu_{\bar\lambda}}{\theta}\kappa^2 t\right)$$

using large deviations arguments. Indeed, the adaptations of the following heuristics are discussed more fully in the next section as they will guide our later proof of the almost sure rate of growth of particles.

We immediately gain the required upper bound since

$$\mathbb{P}_{\mu_{\bar\lambda},\bar\lambda}\left(\xi_t\sim -\gamma t\,; \eta_t^2\geq \kappa^2 t\right) \leq \mathbb{P}_{\mu_{\bar\lambda},\bar\lambda}\left(\eta_t^2\geq \kappa^2 t\right) \sim \exp\left(-\frac{\mu_{\bar\lambda}}{\theta}\kappa^2 t\right)$$

For the lower bound, we break paths into two sections: normal ergodic behaviour over large time period $[0,t]$ followed by a rapid ascent out to type position $\kappa\sqrt{t}$ over a much shorter period $[t,t+\tau]$.

(i) *Ergodic behaviour.* Over a large time t, the occupation density of η will most likely have settled close to the invariant measure. Hence for large t, almost surely under $\mathbb{P}_{\lambda,\mu_\lambda}$,

$$\int_0^t \eta_s^2 \,\mathrm{d}s \sim t\left(\frac{\theta}{2\mu_\lambda}\right)$$

(ii) *Rapid ascent.* Over a large fixed time τ, but where $t \gg \tau$, the probability that η starts close to the origin and ends near to $\kappa\sqrt{t}$ having followed close to the path y over the entire time period τ is roughly given by

$$\exp\left(-\frac{1}{2\theta}\int_0^\tau \left\{\dot{y}(s)+\mu_\lambda y(s)\right\}^2 \,\mathrm{d}s\right)$$

under the $\mathbb{P}_{\lambda,\mu_\lambda}$ law. See, for example, Varadhan (1984). Considering the path

$$y(s) = \kappa\sqrt{t}\left\{\frac{\sinh\mu_\lambda s}{\sinh\mu_\lambda\tau}\right\}$$

gives $\int_0^\tau y(s)^2\,\mathrm{d}s \approx \kappa^2 t/2\mu_\lambda$ with the probability of this path being roughly $\exp -(\mu_\lambda/\theta)\kappa^2 t$.

Combining these two types of behaviour, we can find paths with final positions

$$\eta_{t+\tau}\sim\kappa\sqrt{t}, \qquad \xi_{t+\tau}\sim\lambda a\int_0^{t+\tau}\eta_s^2\,\mathrm{d}s \approx \lambda a\left(\frac{\theta}{2\mu_\lambda}+\frac{\kappa^2}{2\mu_\lambda}\right)t$$

and, moreover, when using the optimal λ value of $\bar\lambda(\gamma,\kappa)$, this actually gives $\xi_{t+\tau}\sim\gamma t$. Further, one of these paths occurs with a probability of roughly $\exp -(\mu_\lambda/\theta)\kappa^2 t$ and note that $t+\tau\sim t$ since $t\ll\tau$, whence the probability $\mathbb{P}_{\mu_{\bar\lambda},\bar\lambda}\left(\xi_t\sim-\gamma t;\eta_t^2\geq\kappa^2 t\right)$ must be at least as large as required.

A tightening up of this argument proves, as required, that

$$\lim_{t\to\infty} t^{-1}\log E\left(\sum_{k=1}^{N(t)}\mathbf{1}\left\{X_k(t)\leq -\gamma t; Y_k(t)^2\geq\kappa^2 t\right\}\right) = \Delta(\gamma,\kappa). \tag{19}$$

If we scale all spatial coordinates by t^{-1} and all type coordinates by $(\sqrt{t})^{-1}$ at time t, the expected asymptotic shape can be considered to be the region $\mathcal{S} := \{(\gamma,\kappa) : \Delta(\gamma,\kappa)\geq 0\}$ where, on average, we have growth in the numbers of (scaled) particles.

3 Large deviation heuristics

In this section we try to present an intuitive picture that better explains why the almost-sure asymptotic exponential growth of $N_t(\gamma, \kappa)$ is at *least* of size $\Delta(\gamma, \kappa)$ whenever $\Delta(\gamma, \kappa) > 0$. This intuitive large deviation picture will provide a crucial insight into the behaviour of the particles, clearly revealing the mechanism that facilitates the growth, although the spine change of measure used to prove rigorously the difficult lower bound will be somewhat different in its approach and martingales will be used to prove the same rate also gives an almost sure upper bound. Nevertheless, the intuitive large deviation picture of this section will also motivate the conjectured results in the final section which we hope to prove rigourously and elegantly using spines in some future work.

Our interest is in the number of particles, $N_t(\gamma, \kappa)$, that have space-type positions *near* to $(\gamma t, \kappa\sqrt{t})$ at a large time t. Consider the case when $\kappa \neq 0$, so we require particles at a large distance from the type origin. However, when $\theta > 8r$ the attraction to the type origin is much stronger than the quadratic growth rate and it is very unlikely that any particle satisfies $Y_k^2(t) > \varepsilon t$ for a prolonged period of time. The majority of particles that contribute to $N.(\gamma, \kappa)$ at time t will fail to contribute to $N.(\gamma, \kappa)$ at times soon after; particles come and go very quickly from the wave front near $(\gamma t, \kappa\sqrt{t})$. There will be many possible trajectories these particles have travelled along to get to a position $(\gamma t, \kappa\sqrt{t})$ by large time t, but we need to find the form of a trajectory that a dominant number of particles will follow. It turn out that such dominant particles will more than likely have had a history made up of two distinct phases:

(i) the *long tread* taking up almost all of the available time. During $[0, t]$, their type has the modified occupation measure which allows them to drift spatially with speed $\gamma\theta/(\theta + \kappa^2)$. This is followed by

(ii) the *short climb* when, over a fixed time $[t, t + \tau]$ (where $\tau \ll t$ is a fixed number) some particles make a rapid ascent to type position $\kappa\sqrt{t}$ whilst additionally gaining $\{\gamma\kappa^2/(\theta + \kappa^2)\}t$ in spatial positioning. These particles will then stay near $(\gamma t, \kappa\sqrt{t})$ for only a short period of time before their type decays back to 0 again.[*Note:* If $\kappa = 0$, only the first phase is applicable.]

3.1 The long tread

Consider particles that are found near αt in space at large time t, earlier expected numbers calculations have already suggested that there will be of order

$$\exp \Delta(\alpha)t$$

of such particles where

$$\Delta(\alpha) = \rho + \left(\theta - \sqrt{a^{-1}(\theta - 8r)(4\alpha^2 + \theta a)}\right)/4$$

The particles alive at this time will subsequently evolve independently. We refer the reader to the heuristics in Harris and Williams (1996), based on large-deviations of the occupation density for the type process. Roughly speaking, the quadratic breeding enables a large number of particles to behave *as if* their type positions moved as Ornstein-Uhlenbeck processes with a *weaker* drift of μ_{λ_α} back towards the type origin, which then increases the spatial variance and so facilitates $\exp \Delta(\alpha)t$ of these also gaining a final spatial position of αt. The ancestors of the majority of particles that contribute to $N_t(\alpha)$ will have also contributed to $N_s(\alpha)$ at earlier times $s \leq t$, that is, the vast majority of those particles near αt in space at large time t will have got there by steadily drifting spatially with speed α. We can track these particles because they contributed most to the martingale $Z^-_{\lambda_\alpha}(t)$ and we later prove that $\Delta(\alpha)$ is indeed the almost sure growth rate when $\alpha < \tilde{c}(\theta)$ via martingale methods.

3.2 The short climb

We will now suggest why the probability a single particle starting near to the origin manages to make the rapid ascent to have at least one descendent in the vicinity of $(\beta t, \kappa\sqrt{t})$ during a small interval of time just before fixed time of length τ turns out to be roughly

$$\exp -\Theta(\beta, \kappa)t$$

as t gets very large, where

$$\Theta(\beta,\kappa) := \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta-8r)(a^2\kappa^4+4a\theta\beta^2)}}{4a\theta}$$

Suppose we start the branching diffusion with a single particle at $(0,0)$, we wish to know the probability that there is at least one particle at a fixed time τ that has a spatial position near βt having followed close to the path $x(s)$ for $0 \le s \le \tau$ *and* a type position near $\kappa\sqrt{t}$ having closely followed the path $y(s)$ for $0 \le s \le \tau$ for t arbitrarily large.

We recall from large deviation theory that the probability that a *single* particle manages to closely follow both the type path $y(s)$ and the spatial path $x(s)$ for $0 \le s \le \tau$ is roughly given by

$$\exp -\frac{1}{2\theta}\int_0^\tau \left\{\dot{y}(s) + \frac{\theta}{2}y(s)\right\}^2 \mathrm{d}s - \frac{1}{2}\int_0^\tau \frac{\dot{x}(s)^2}{ay(s)^2}\,\mathrm{d}s$$

when $x(0)=0, x(\tau)=\beta t$, $y(0)=0, y(\tau)=\kappa\sqrt{t}$ and t is very large. This probability will typically be very small, but if such paths are followed by particles in the branching diffusion, we have to also take account of the large breeding rates that are found far from the type origin.

If we let $X(s)$ represent the numbers of particles in the branching diffusion that are alive at time s and have travelled 'close' to the path $\big(x(u),y(u)\big)$ for $0 \le u \le s$, then we can get a rough idea of how X might behave by considering the following:

3.3 A birth-death process

Let M be a time-dependent birth-death process where at time s particles either give birth to single offspring with *breeding rate* λ_s given by

$$\lambda_s = \rho + ry(s)^2$$

or particles die with *death rate* μ_s given by

$$\mu_s = \frac{1}{2\theta}\left\{\dot{y}(s) + \frac{\theta}{2}y(s)\right\}^2 + \frac{1}{2}\frac{\dot{x}(s)^2}{ay(s)^2}.$$

The *effective death rate* up to time t is defined by $\rho_t := \int_0^t(\mu_s - \lambda_s)\,\mathrm{d}s$, so here

$$\rho_t = J(x,y,t) := \int_0^t \left\{\frac{1}{2\theta}\left(\dot{y}(s)+\frac{\theta}{2}y(s)\right)^2 + \frac{1}{2}\frac{\dot{x}(s)^2}{ay(s)^2} - ry(s)^2 - \rho\right\}\mathrm{d}s.$$

The distribution for total number of offspring surviving, $M(\tau)$, for the time-dependent birth-death process is well known. Defining

$$W_\tau := e^{-\rho_\tau}\left\{1 + \int_0^\tau \mu_s\, e^{\rho_s}\, ds\right\}, \qquad \xi_\tau := 1 - e^{-\rho_\tau}W_\tau^{-1}, \qquad \eta_\tau := 1 - W_\tau^{-1}$$

we have

$$\mathbb{P}(M(\tau)=0) = \xi_\tau, \quad \mathbb{P}(M(\tau)=n) = (1-\xi_\tau)(1-\eta_\tau)\eta_\tau^{n-1} \qquad n=1,2,\ldots,$$

with $\mathbb{E}M(\tau) = e^{-\rho_\tau}$ and $\mathbb{E}(M(\tau)|M(\tau)\ge 1) = W_\tau$.

In our particular case, we have

$$\mathbb{E}\big(M(\tau)\big) = \exp -J(x,y,\tau).$$

Define the largest effective death rate prior to time τ by

$$L(x,y,\tau) := \sup_{s\in[0,\tau]} J(x,y,s) \ge 0.$$

If we are in a case when $L(x,y,\tau)$ is very large (suggesting a high chance of extinction), then

$$\mathbb{P}\big(M(\tau)\ge 1\big) \sim K_\tau \exp -L(x,y,\tau),$$

where $K_\tau^{-1} := \int_0^\tau \mu_s \exp -\{L(x,y,\tau)-\rho_s\}\,ds$ and, if there is at least one particle alive, we expect to have

$$\mathbb{E}(M(\tau)|M(\tau)\geq 1) \sim K_\tau^{-1}\exp L(x,y,\tau) - J(x,y,\tau).$$

Thus, we would guess the probability any particles in the branching diffusion manage to make the difficult, rapid ascent along path (x,y) to finish up near $(\beta t, \kappa\sqrt{t})$ can, very roughly, be estimated by $\exp -L(x,y,\tau)$.

We further guess that the chance any particles manage to stay near position $(\beta t, \kappa\sqrt{t})$ during a very small interval of time close to τ should roughly look like

$$\exp - \inf_{x,y} L(x,y,\tau),$$

where we permit all possible paths x and y satisfying $x(0)=0, x(\tau)=\beta t$ and $y(0)=0, y(\tau)=\kappa\sqrt{t}$ for the fixed time τ.

3.4 Finding the optimal path and probability

We proceed to calculate

$$\inf_{x,y} L(x,y,\tau)$$

over paths x and y satisfying $x(0)=0, x(\tau)=\beta t$ and $y(0)=0, y(\tau)=\kappa\sqrt{t}$ for the fixed time τ.

We first note that

$$\inf_{x,y} L(x,y,\tau) = \inf_{x,y}\sup_{s\in[0,\tau]} J(x,y,s) \geq \inf_{x,y} J(x,y,\tau) \tag{20}$$

and we now proceed to calculate $\inf_{x,y} J(x,y,\tau)$.

We can easily optimise over the choice of function x given y, finding that

$$\dot{x}(s) \propto \lambda a y(s)^2 \quad \Rightarrow \quad x(s) = \lambda a \int_0^\tau y(u)^2\,\mathrm{d}u$$

where

$$\lambda = \frac{\beta t}{a\int_0^\tau y(s)^2\,\mathrm{d}s}$$

yielding

$$\frac{1}{2}\int_0^\tau \frac{\dot{x}(s)^2}{ay(s)^2}\,\mathrm{d}s = \frac{\beta^2 t^2}{2a\int_0^\tau y(s)^2\,\mathrm{d}s}.$$

This is as anticipated since, when following the path y in type space, the spatial position of a particle is following a Brownian motion with total amount of variance over period τ given by

$$a\int_0^\tau y(s)^2\,\mathrm{d}s$$

and so the probability that a particle following the path y in type space will also be found near to βt in space at time τ is roughly

$$\exp\frac{-\beta^2 t^2}{2a\int_0^\tau y(s)^2\,\mathrm{d}s}$$

Then we are left to find

$$\begin{aligned}
&\inf_y\left\{\int_0^\tau \left(\frac{1}{2\theta}\Big(\dot{y}(s)+\frac{\theta}{2}y(s)\Big)^2 + \frac{\beta^2t^2}{2a\int_0^\tau y(s)^2\,\mathrm{d}s} - ry(s)^2\right)\,\mathrm{d}s\right\}\\
&\quad = \inf_y\sup_\lambda\left\{\int_0^\tau \left(\frac{1}{2\theta}\Big(\dot{y}(s)+\frac{\theta}{2}y(s)\Big)^2 - ry(s)^2 - \frac{1}{2}a\lambda^2 y(s)^2\right)\,\mathrm{d}s + \lambda\beta t\right\}\\
&\quad \geq \sup_\lambda\inf_y\left\{\int_0^\tau \left(\frac{1}{2\theta}\Big(\dot{y}(s)+\frac{\theta}{2}y(s)\Big)^2 - ry(s)^2 - \frac{1}{2}a\lambda^2 y(s)^2\right)\,\mathrm{d}s + \lambda\beta t\right\}
\end{aligned}$$

Some further Euler-Lagrange optimisation now gives the optimal path as

$$y_\lambda(s) = \kappa\sqrt{t}\left\{\frac{\sinh\mu_\lambda s}{\sinh\mu_\lambda \tau}\right\} \qquad (0 \le s \le \tau) \tag{21}$$

where

$$\mu_\lambda = \frac{\sqrt{\theta(\theta - 8r - 4a\lambda^2)}}{2}$$

and then

$$\begin{aligned}&\sup_\lambda \inf_y \left\{\int_0^\tau \left(\frac{1}{2\theta}\Big(\dot{y}(s) + \frac{\theta}{2}y(s)\Big)^2 - ry(s)^2 - \frac{1}{2}a\lambda^2 y(s)^2\right) \mathrm{d}s + \lambda\beta t\right\}\\ &\qquad = \sup_\lambda \left\{\lambda\beta t + \kappa^2 t\left(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth\mu_\lambda\tau\right)\right\}\end{aligned}$$

The optimal choice $\hat{\lambda}$ satisfies

$$\frac{\beta t}{a\hat{\lambda}} = \kappa^2 t\left(\frac{\coth\mu_{\hat{\lambda}}\tau}{2\mu_{\hat{\lambda}}} - \frac{\tau}{2\sinh^2\mu_{\hat{\lambda}}\tau}\right) = \int_0^\tau y_{\hat{\lambda}}(s)^2\,\mathrm{d}s$$

and it is now easy to check that the supremum and infimum could have been freely interchanged, maintaining equality in the previous expression.

Then, with the optimal spatial path

$$x_\lambda(s) := \lambda a\int_0^s y_\lambda(u)^2\,\mathrm{d}u = \beta t\left\{\frac{\sinh 2\mu_\lambda s - 2\mu_\lambda s}{\sinh 2\mu_\lambda\tau - 2\mu_\lambda\tau}\right\} \tag{22}$$

and defining $\hat{x} := x_{\hat{\lambda}}, \hat{y} := y_{\hat{\lambda}}$ we have

$$\begin{aligned}\inf_{x,y} J(x,y,\tau) &= J(\hat{x},\hat{y},\tau)\\ &= t\ \sup_\lambda\left\{\lambda\beta + \kappa^2\left(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth\mu_\lambda\tau\right)\right\} - \rho\tau\\ &= t\left\{\hat{\lambda}\beta + \kappa^2\left(\frac{1}{4} + \frac{\mu_{\hat{\lambda}}}{\theta}\coth\mu_{\hat{\lambda}}\tau\right)\right\} - \rho\tau\end{aligned}$$

Finally, it is easy to check that $J(\hat{x},\hat{y},\tau) = L(\hat{x},\hat{y},\tau)$, whence we have

$$\inf_{x,y} J(x,y,\tau) \ge \inf_{x,y} L(x,y,\tau),$$

and, combining with equation (20), we have found that

$$\inf_{x,y} L(x,y,\tau) = J(\hat{x},\hat{y},\tau).$$

3.5 A note on the optimal paths

As $\tau \to \infty$, we have

$$\sup_\lambda\left\{\lambda\beta + \kappa^2\left(\frac{1}{4} + \frac{\mu_\lambda}{\theta}\coth\mu_\lambda\tau\right)\right\} \uparrow \sup_\lambda\left\{\lambda\beta + \kappa^2\psi_\lambda^+\right\} = \overline{\lambda}\beta + \kappa^2\psi_{\overline{\lambda}}^+$$

where

$$\overline{\lambda} = -\sqrt{\frac{\beta^2\theta(\theta - 8r)}{a^2\kappa^4 + 4a\theta\beta^2}} = \overline{\lambda}\left(\left\{\frac{\kappa^2+\theta}{\kappa^2}\right\}\beta, \kappa\right), \qquad \mu_{\overline{\lambda}} = \frac{\kappa^2\sqrt{\theta(\theta - 8r)}}{2\sqrt{\kappa^4 + 4\theta\beta^2/a}}$$

$$\Theta(\beta,\kappa) := \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta - 8r)(a^2\kappa^4 + 4a\theta\beta^2)}}{4a\theta} = \overline{\lambda}\beta + \kappa^2\psi_{\overline{\lambda}}^+$$

Then, for all $\tau > 0$, we have

$$\exp - \inf_{x,y} J(x,y,\tau) \geq \exp -\Theta(\beta,\kappa)t + \rho\tau$$

and writing $\overline{x} := x_{\overline{\lambda}}, \overline{y} := y_{\overline{\lambda}}$ we further note that, for all $\varepsilon > 0$ and $\delta > 0$, there exist $\tilde{\tau} > 0$ and $\eta > 0$ such that for all $t > 0$ and $\tau > \tilde{\tau}$

$$\begin{aligned}\exp - \inf_{x,y} J(x,y,\tau) &\geq \exp -J(\overline{x},\overline{y},\tau)\\ &= \exp -t\left\{\overline{\lambda}\beta + \kappa^2\left(\frac{1}{4} + \frac{\mu_{\overline{\lambda}}}{2\theta}\coth \mu_{\overline{\lambda}}\tau\right)\right\} + \rho\tau\\ &\geq \exp -t\left\{\Theta(\beta,\kappa) + \varepsilon\right\}\end{aligned}$$

and (when, say, $\kappa > 0$) for all $s \in [\tau - \eta, \tau]$

$$\overline{y}(s) \geq \kappa\sqrt{t}(1-\delta), \qquad \overline{x}(s) \geq \beta t(1-\delta)$$

so that the paths stay close to the required positions for a certain fixed length of time.

3.6 The successful deviant particles

Then the number of particles that are near $(\alpha t, 0)$ at time large time t which then proceed to have *at least one descendant* alive near position $((\alpha+\beta)t, \kappa\sqrt{t})$ during the small time interval $[t+\tau-\eta, t+\tau]$ will be *approximately* Poisson with mean

$$\exp\left\{\Delta(\alpha) - \Theta(\beta,\kappa)\right\}t$$

The actual number found will be remain sufficiently close to this large mean with a very high probability, so we will eventually keep finding large enough numbers of particles close to the required positions.

Optimising over $\alpha + \beta = \gamma$, some simple calculus reveals that with

$$\overline{\alpha} = \gamma\theta/(\theta+\kappa^2), \qquad \overline{\beta} = \gamma\kappa^2/(\theta+\kappa^2)$$

we have $\Delta(\alpha) - \Theta(\beta,\kappa) \leq \Delta(\overline{\alpha}) - \Theta(\overline{\beta},\kappa) = \Delta(\gamma,\kappa)$. [*Note:* In agreement with the optimal parameter found in the expectation calculations of section 2, we note that $\lambda_{\overline{\alpha}} = \overline{\lambda}(\gamma,\kappa)$ from equations (14) and (18).]

Finally, we have the number of particles found in the vicinity of $(\gamma t, \kappa\sqrt{t})$ to be roughly of order

$$\exp \Delta(\gamma,\kappa)\, t$$

for all sufficiently large times t, hence we get a lower bound on the exponential growth rate of $\Delta(\gamma,\kappa)$, in agreement with the growth rate of expected number of particles.

4 Martingales, filtrations and changes of measure

In this section we recall some martingales for both single-particle processes and the OU BBM that we will use for the spine change of measure: a technique that has become popular in the branching-process literature because it has yielded conceptually simple proofs of quite major results.

First the martingales: from Harris and Williams (1996) and Harris (2000) we have two families of additive martingales, Z_λ^- and Z_λ^+, defined by:

$$Z_\lambda^{\pm}(t) := \sum_{u\in N(t)} v_\lambda^{\pm}(Y_u(t))e^{\lambda X_u(t) - E_\lambda^{\pm}t}, \qquad t \geq 0, \lambda \in (\lambda_{\min}, 0) \tag{23}$$

where v_λ^+ and v_λ^- are strictly-positive eigenfunctions of the self-adjoint operator

$$\frac{1}{2}\lambda^2 A + R + Q_\theta,$$

with corresponding eigenvalues $E_\lambda^- < E_\lambda^+$. These martingales are well studied – see Harris and Williams (1996), Harris (2000) and Hardy (2004) for details of their $\mathcal{L}^p$-convergence and limits. When we start the branching diffusion from a single particle at space-type position $(x, y) \neq (0, 0)$ we will also require the normalised martingale

$$\widehat{Z}_\lambda^+(t) := \frac{Z_\lambda^+(t)}{Z_\lambda^+(0)}.$$

Additive martingales for branching processes such as those in (23) can be used to change measure on the probability space of the branching process, and under this new measure the behaviour of a single particle (the *spine*) is altered, whilst its offspring behave as if under the original measure. This was first seen for branching Brownian motion in Chauvin and Rouault (1988) and developed for Galton-Watson processes in Lyons et al. (1995), Lyons (1997) and Kurtz et al. (1997). Kyprianou (2003) and Engländer and Kyprianou (2002) developed the technique for some families of branching diffusions, and more recently the spine approach has been significantly improved for a very general class of branching diffusions in Hardy and Harris (2004*a*). We use this improved approach to construct our change of measure, and, following Hardy and Harris (2004*a*), we will first outline the notation and the describe the changes of measure. The reader that is familiar with Kyprianou (2003) will recognise much of the notation, but the differences between Kyprianou's approach to the spine construction and the method below are described in detail in Hardy and Harris (2004*a*) and Hardy (2004).

All probability measures are to be defined on the space $\tilde{\mathcal{T}}$ of *marked Galton-Watson trees with spines*; before defining precisely what this space is we need to set up some other notation. We recall the set of Ulam-Harris labels, Ω, defined by

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

where $\mathbb{N} := \{1, 2, 3, \ldots\}$. For two words $u, v \in \Omega$, uv denotes the concatenated word, where we take $u\emptyset = \emptyset u = u$. So Ω contains elements such as '$\emptyset 412$', which represents 'the individual being the 2nd child of the 1st child of the 4th child of the initial ancestor $\emptyset$'. For two labels $u, v \in \Omega$ the notation $v < u$ means that v is an ancestor of u, and $|u|$ denotes the length of u.

We define a *Galton-Watson tree* to be a set $\tau \subset \Omega$ such that:

(i) $\emptyset \in \tau$: there is the unique initial ancestor;

(ii) if $u, v \in \Omega$, then $vu \in \tau \Rightarrow v \in \tau$: this means that τ contains all of its ancestors of its nodes, and by $v < u$ we mean that v is an ancestor of u;

(iii) for all $u \in \tau$, $uj \in \tau$ if and only if $j \in \{1, 2\}$ – so each particle produces two offspring.

The set of all such trees is $\mathbb{T}$, and we will use the symbol τ for a particular tree. Of course, as our work concerns branching diffusions we shall often refer to the labels of τ as particles. A Galton-Watson tree by itself only records the family structure of the individuals, so to each individual $u \in \tau$ we give a mark (X_u, Y_u, σ_u) which contains the following information:

- $\sigma_u \in [0, \infty)$ is the *lifetime* of particle u, which also determines the *fission time* of the particle as $S_u := \sum_{v \leq u} \sigma_v$ (with $S_\emptyset := \sigma_\emptyset$). We may also refer to the S_u as *death times*;
- the function $X_u(t) : [S_u - \sigma_u, S_u) \to \mathbb{R}$ describes the particle's *spatial* motion in $\mathbb{R}$ during its lifetime;
- the function $Y_u(t) : [S_u - \sigma_u, S_u) \to \mathbb{R}$ describes the evolution of the particle's *type* in $\mathbb{R}$ during its lifetime. Y_u maps into $\mathbb{R}$ because the type evolves as an Ornstein-Uhlenbeck process in our model – this would be different if the type evolved as, say, a Markov chain.

For clarity we must decide whether or not a particle is in existence at its death time: our convention will be that a particle dies 'infinitesimally before' its death time – this is why X_u and Y_u are defined on $[S_u - \sigma_u, S_u)$ and not $[S_u - \sigma_u, S_u]$ – so that at time S_u the particle u has disappeared and has been replaced by its two children.

We denote a particular marked tree by (τ, X, Y, σ), or the abbreviation (τ, M), and the set of all marked Galton-Watson trees by $\mathcal{T}$:

$$\mathcal{T} := \Big\{(\tau, X, Y, \sigma) : \tau \in \mathbb{T}, \text{ and for each } u \in \tau \text{ we have } \sigma_u \in \mathbb{R}^+ \\ \text{and } X_u, Y_u : [S_u - \sigma_u, S_u) \to \mathbb{R}\Big\}.$$

For each $(\tau, X, Y, \sigma) \in \mathcal{T}$, the set of particles alive at time t is defined as

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

For any given marked tree $(\tau, M) \in \mathcal{T}$ we can distinguish individual lines of descent from the initial ancestor: $\emptyset, u_1, u_2, u_3, \ldots \in \tau$ where u_i is a child of u_{i-1} for all $i \in \{2, 3, \ldots\}$ and u_1 is a child of the initial individual $\emptyset$. We call such a line of descent a *spine* and denote it by ξ. In a slight abuse of notation we refer to ξ_t as the unique node in ξ that is alive at time t, and also for the position of the particle that makes up the spine at time t; that is $\xi_t := X_u(t)$, where $u \in \xi \cap N_t$. However, although the interpretation of ξ_t should always be clear from the context, we introduce the following notation for use where some ambiguity may still arise:

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

It is natural to think of the spine as a single diffusing particle ξ_t, or, strictly speaking, the pair (ξ_t, η_t), where η_t is the type of the spine at time t, because the type process of the spine could also be a diffusion: in our situation it is an Ornstein-Uhlenbeck process.

We define n_t to be a counting function that tells us which generation of the spine is currently alive, or equivalently the number of fission times there have been on the spine:

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

The collection of all marked trees with a distinguished spine is the space $\tilde{\mathcal{T}}$ on which our probability measures will eventually be defined, but first we define four filtrations on this space that contain different levels of information about the branching diffusion.

Filtration $(\mathcal{F}_t)_{t\geq 0}$

We define a filtration of $\tilde{\mathcal{T}}$ made up of the σ-algebras

$$\mathcal{F}_t := \sigma\Big((u, X_u, Y_u, \sigma_u) : S_u \leq t; (u, X_u(s), Y_u(s) : s \in [S_u - \sigma_u, t]) \\ : t \in [S_u - \sigma_u, S_u)\Big),$$

which means that $\mathcal{F}_t$ is generated by the information concerning all particles that have lived and died before time t, and also those that are still alive at time t. Each of these σ-algebras is a subset of the limit

$$\mathcal{F}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{F}_t\Big).$$

Filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$

We define the filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ simply by augmenting the filtration $\mathcal{F}_t$ with the knowledge of which node is the spine at time t:

$$(\tilde{\mathcal{F}}_t)_{t\geq 0} := \sigma(\mathcal{F}_t, \text{node}_t(\xi)), \qquad \tilde{\mathcal{F}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{F}}_t\Big),$$

so that this filtration knows everything about the branching diffusion and everything about the spine.

Filtration $(\mathcal{G}_t)_{t\geq 0}$

$(\mathcal{G}_t)_{t\geq 0}$ is a filtration of $\tilde{\mathcal{T}}$ defined by

$$\mathcal{G}_t := \sigma(\xi_s : 0 \leq s \leq t), \qquad \mathcal{G}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{G}_t\Big).$$

These σ-algebras are generated only by the spine's motion and so do not contain the information about which nodes of the tree τ make up the spine.

Filtration $(\tilde{\mathcal{G}}_t)_{t\geq 0}$

As we did in going from $\mathcal{F}_t$ to $\tilde{\mathcal{F}}_t$ we create $(\tilde{\mathcal{G}}_t)_{t\geq 0}$ from $(\mathcal{G}_t)_{t\geq 0}$ by including knowledge of which nodes make up the spine:

$$(\tilde{\mathcal{G}}_t)_{t\geq 0} := \sigma(\mathcal{G}_t, \mathrm{node}_t(\xi)), \qquad \tilde{\mathcal{G}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{G}}_t\Big).$$

This means that $\tilde{\mathcal{G}}_t$ also knows when the fission time on the spine occurred, whereas $\mathcal{G}_t$ does not.

The relationships between the filtrations of $\tilde{\mathcal{T}}$ may be summarised thus:

$$\mathcal{G}_t \subset \tilde{\mathcal{G}}_t \subset \tilde{\mathcal{F}}_t$$

$$F_t \subset \tilde{\mathcal{F}}_t$$

Importantly, we have $\mathcal{G}_t \not\subseteq \mathcal{F}_t$, since $\mathcal{F}_t$ does not know which line of descent makes up the spine and so it cannot know the spine's motion.

Now that we have defined the underlying space and the filtrations of it that we require, we define the probability measures on that space that give us the typed branching diffusion. The measures $\{P^{x,y} : x, y \in \mathbb{R}\}$ on $(\mathcal{T}, \mathcal{F}_\infty)$ give rise to the typed branching diffusion described in section 1, and the following result of Lyons (1997) allows us to extend the measure $P^{x,y}$ to a measure that also keeps track of the spine.

Theorem 6 *If f is an $\tilde{\mathcal{F}}_t$-measurable function then we can write*

$$f = \sum_{u\in N_t} f_u \mathbf{1}_{(\xi_t = u)}, \tag{24}$$

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

Now we extend $P^{x,y}$ to a measure $\tilde{P}^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by choosing the spine uniformly over all lines of descent in the tree; more precisely, for any $f \in m\tilde{\mathcal{F}}_t$ with representation like (24), we have:

$$\int_{\tilde{\mathcal{T}}} f \,\mathrm{d}\tilde{P}^{x,y}(\tau, M, \xi) := \int_{\mathcal{T}} \sum_{u\in N_t} f_u \prod_{v<u} \frac{1}{2} \,\mathrm{d}P^{x,y}(\tau, M).$$

Remark: Empty products are taken as equal to 1.

The changes of measure leading to the spine decomposition will use the Z_λ^+ additive martingale from (23). From the existing spine literature we expect that there will be 3 changes to the spine's behaviour under the new measure: the spine will gain a spatial drift, the spine's type process will be modified and the birth rate on the spine will increase. We now list the martingales that will perform these changes:

Proposition 7 *The process*

$$2^{n_t} e^{-\int_0^t R(\eta_s)\,\mathrm{d}s}$$

is a $\tilde{P}$-martingale that will increase the rate at which births occur on the spine from $R(\eta_t)$ to $2R(\eta_t)$:

$$\frac{\mathrm{d}\mathbb{L}_t^{2R(\eta_t)}}{\mathrm{d}\mathbb{L}_t^{R(\eta_t)}} = 2^{n_t} e^{-\int_0^t R(\eta_s)\,\mathrm{d}s}.$$

Proposition 8 *The process*

$$e^{\lambda \xi_t - \frac{1}{2}\lambda^2 \int_0^t A(\eta_s)\,\mathrm{d}s}$$

is a $\tilde{P}$-martingale that will give the spine an instantaneous spatial drift of $\lambda A(\eta_t)$ under the measure $\mathbb{P}_\lambda$:

$$\left.\frac{\mathrm{d}\mathbb{P}^x_\lambda}{\mathrm{d}\mathbb{P}^x}\right|_{\tilde{\mathcal{F}}_t} = e^{\lambda \xi_t - \frac{1}{2}\lambda^2 \int_0^t A(\eta_s)\,\mathrm{d}s}$$

Proposition 9 *The process*

$$v^+_\lambda(\eta_t) e^{\int_0^t \left\{R(\eta_s) + \frac{1}{2}\lambda^2 A(\eta_s)\right\}\mathrm{d}s - E^+_\lambda t}$$

is a $\tilde{P}$-martingale that will change the drift direction of η_t. More precisely, we have:

$$\frac{\mathrm{d}\mathbb{P}_t^{(\theta,-\mu_\lambda)}}{\mathrm{d}\mathbb{P}_t^{(\theta,\frac{\theta}{2})}} = v^+_\lambda(\eta_t) e^{\int_0^t \left\{R(\eta_s) + \frac{1}{2}\lambda^2 A(\eta_s)\right\}\mathrm{d}s - E^+_\lambda t},$$

where, under $\mathbb{P}_t^{(\delta,\theta)}$, η_t moves on $\mathbb{R}$ as an OU-process with generator

$$\frac{\theta}{2}\frac{\partial^2}{\partial y^2} - \delta y \frac{\partial}{\partial y},$$

and we will refer to this as an OU(θ,δ) process.

These product of these 3 martingales allows us to construct the $\tilde{\mathcal{F}}_t$-measurable martingale $\tilde{\zeta}(t)$ defined by

$$\begin{aligned}\tilde{\zeta}(t) &:= v^+_\lambda(\eta_t) e^{\int_0^t \left\{R(\eta_s) + \frac{1}{2}\lambda^2 A(\eta_s)\right\}\mathrm{d}s - E^+_\lambda t} \times 2^{n_t} e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} \\ &\qquad \times e^{\lambda \xi_t - \frac{1}{2}\lambda^2 \int_0^t A(\eta_s)\,\mathrm{d}s} \\ &= v^+_\lambda(\eta_t) 2^{n_t} e^{\lambda \xi_t - E^+_\lambda t}.\end{aligned}$$

Via the martingale $\tilde{\zeta}(t)$ we may construct a measure $\tilde{\mathbb{Q}}^{x,y}_\lambda$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^{x,y}_\lambda}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}(t)}{\tilde{\zeta}(0)} = \frac{e^{-\lambda x}}{v^+_\lambda(y)} v^+_\lambda(\eta_t) 2^{n_t} e^{\lambda \xi_t - E^+_\lambda t}. \tag{25}$$

We can now give decompositions of the measures $\tilde{P}$ and $\tilde{\mathbb{Q}}_\lambda$ that enables us to see what the path-wise construction of the branching diffusion should be under the two measures; similar constructions were first seen in Chauvin and Rouault (1988) and the form that we use here is taken from Kyprianou (2003):

Theorem 10 *With the notation of Propositions 7, 8 and 9 we have:*

$$\mathrm{d}\tilde{P}(\tau, M, \xi) = \mathrm{d}\mathbb{P}(\xi)\,\mathrm{d}\mathbb{P}^{(\theta,\frac{\theta}{2})}(\eta)\,\mathrm{d}\mathbb{L}^{R(\eta_t)} \prod_{v<\xi_t} \frac{1}{2} \times \mathrm{d}P((\tau, M)^v).$$

From this decomposition we can recover the path-wise construction of the branching diffusion under $\tilde{P}$:

- the spine diffuses spatially as a driftless Brownian motion with variance $A(\eta_t)$;
- the type of the spine diffuses as an OU(θ,$\frac{\theta}{2}$) process;
- fission-times along the spine occur as an independent Poisson process of rate $R(\eta_t)$;
- at each fission-time on the spine two particles are produced: the spine is chosen uniformly at random between these two particles and the other particle gives rise to the subtree $(\tau, M)^v$, which is not part of the spine and behaves like an independent copy of the measure $P^{\cdot,\cdot}$.

If we now factor in the Radon-Nikodým derivatives from the martingale $\tilde{\zeta}(t)$ we will obtain a similar decomposition for $\tilde{\mathbb{Q}}$. On $\tilde{\mathcal{F}}_t$ we have:

$$\begin{aligned}
\mathrm{d}\tilde{\mathbb{Q}} &= \tilde{\zeta}(t)\,\mathrm{d}\tilde{P} \\
&= v_\lambda^+(\eta_t) e^{\int_0^t \left\{R(\eta_s)+\frac{1}{2}\lambda^2 A(\eta_s)\right\}\mathrm{d}s - E_\lambda^+ t} \times 2^{n_t} e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} \\
&\qquad \times e^{\lambda\xi_t - \frac{1}{2}\lambda^2\int_0^t A(\eta_s)\,\mathrm{d}s} \times \mathrm{d}P((\tau,M)^v) \\
&= \mathrm{d}\mathbb{P}_\lambda(\xi)\,\mathrm{d}\mathbb{P}^{(\theta,-\mu_\lambda)}(\eta)\,\mathrm{d}\mathbb{L}^{2R(\eta_t)} \prod_{v<\xi_t}\frac{1}{2} \times \mathrm{d}P((\tau,M)^v),
\end{aligned}$$

and, as with the measure $\tilde{P}$, from this decomposition we see that the branching diffusion may be constructed path-wise under $\tilde{\mathbb{Q}}$ as below:

Definition 11

- *starting from spatial position x and type y the spine (ξ_t,η_t) diffuses spatially as a Brownian motion with infinitesimal variance $A(\eta_t)$ and infinitesimal drift $\lambda A(\eta_t)$;*
- *the type of the spine, η_t, begins at y and moves in type space as an Ornstein-Uhlenbeck process with generator*
$$\frac{\theta}{2}\frac{\partial^2}{\partial y^2} + \mu_\lambda y \frac{\partial}{\partial y}$$
(notice that η_t has a drift driving it away *from the origin in type space);*
- *the spine branches at rate $2R(\eta_t)$, producing 2 particles;*
- *one of these particles is selected uniformly at random;*
- *the chosen offspring repeats stochastically the behaviour of its parent;*
- *the other offspring particle initiates a $P^{\cdot,\cdot}$-BBM from its birth position and type.*

We note here that our use of this change of measure is motivated by the fact that the expected position of the outward drifting OU-process follows the optimal path for a single particle to reach large positions in type space (21).

By definition of conditional expectation, the change of measure (25) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation:

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

To evaluate this conditional expectation we use the representation (24) and the fact that

$$2^{n_t} = \prod_{v<\xi_t} 2,$$

and we obtain

$$\begin{aligned}
\tilde{P}^{x,y}\left(v_\lambda^+(\eta_t)2^{n_t}e^{\lambda\xi_t - E_\lambda^+ t}\middle|\mathcal{F}_t\right) &= \tilde{P}^{x,y}\Bigg(\sum_{u\in N_t} v_\lambda^+(Y_u(t))2^{n_t}e^{\lambda X_u(t) - E_\lambda^+ t} \\
&\qquad \times \prod_{v<u} 2 \times \mathbf{1}_{(\xi_t = u)}\Bigg|\mathcal{F}_t\Bigg) \\
&= \sum_{u\in N_t} v_\lambda^+(Y_u(t))2^{n_t}e^{\lambda X_u(t) - E_\lambda^+ t} \times \prod_{v<u} 2 \\
&\qquad \times \tilde{P}^{x,y}(\xi_t = u|\mathcal{F}_t) \\
&= \sum_{u\in N_t} v_\lambda^+(Y_u(t))2^{n_t}e^{\lambda X_u(t) - E_\lambda^+ t} \times \prod_{v<u} 2 \times \prod_{v<u}\frac{1}{2} \\
&= \sum_{u\in N_t} v_\lambda^+(Y_u(t))2^{n_t}e^{\lambda X_u(t) - E_\lambda^+ t} = Z_\lambda^+(t).
\end{aligned}$$

So we have proved:

Theorem 12 *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* $\mathcal{F}_\infty$ *that satisfies*

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

Moreover under $\mathbb{Q}_\lambda^{x,y}$ *and* $\tilde{\mathbb{Q}}_\lambda^{x,y}$*, the path-wise construction of the branching diffusion is that given in definition 11.*

Remark: Although the path-wise construction of the branching diffusion is the same under $\mathbb{Q}_\lambda$ and $\tilde{\mathbb{Q}}_\lambda$, only the measure $\tilde{\mathbb{Q}}_\lambda$ 'knows' about the spine. It is clear, however, that we have $\tilde{\mathbb{Q}}_\lambda(A) = \mathbb{Q}_\lambda(A)$ for any $A \in \mathcal{F}_\infty$.

Under the measure $\tilde{\mathbb{Q}}$ only the behaviour of the spine is altered, and combining this observation with conditioning on the spine's path and fission-times gives us a very useful representation for $Z_\lambda^+(t)$ under $\tilde{\mathbb{Q}}_\lambda$ that we shall refer to as the *spine decomposition*:

$$\tilde{\mathbb{Q}}(Z_\lambda^+(t)|\tilde{\mathcal{G}}_\infty) = \sum_{u<\xi_t} v_\lambda^+(\eta_{S_u})e^{\lambda\xi_{S_u} - E_\lambda^+ S_u} + v_\lambda^+(\eta_t)e^{\lambda\xi_t - E_\lambda^+ t}. \tag{26}$$

Throughout the rest of this article we will refer to the two pieces of this decomposition as the 'sum term' and the 'spine term'. This decomposition is discussed in detail for a wide variety of branching diffusions in Hardy (2004), but to derive it we simply note that the contributions to $Z_\lambda^+(t)$ from the subtrees that branch off the spine have *constant* $\tilde{\mathbb{Q}}_\lambda$-expectation because they behave as if under the original measure P; and we know that $Z_\lambda^+(t)$ is a P-martingale. The spine decomposition reduces many calculations about the behaviour of $Z_\lambda^+(t)$ under $\tilde{\mathbb{Q}}_\lambda$ to one-particle calculations about the spine, and this observation is exploited in the spine proofs of $\mathcal{L}^p$-boundedness for some families of additive martingales in Hardy and Harris (2004*d*).

The following measure-theoretic result shows the relationships between the limits of $Z_\lambda^+(t)$ under the measures P and $\mathbb{Q}$.

Theorem 13 *Let* P *and* $\mathbb{Q}$ *be two probability measures on a filtered space* $\big(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}\big)$ *that are related by the Radon-Nikodým derivative*

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

for some strictly-positive martingale Z_t*. Define* $Z_\infty := \limsup_{t\to\infty} Z_t$*, then for any* $A \in \mathcal{F}_\infty$ *we have*

$$\mathbb{Q}(A) = \int_A Z_\infty \,\mathrm{d}P + \mathbb{Q}\big(A \cap \{Z_\infty = \infty\}\big), \tag{27}$$

and hence:

(i) $\mathbb{Q}$ *is absolutely continuous with respect to* P *if and only if* $\int_\Omega Z_\infty \,\mathrm{d}P = 1$ *if and only if* $\mathbb{Q}(Z_\infty = \infty) = 0$*;*

(ii) $\mathbb{Q}$ *is singular with respect to* P *if and only if* $P(Z_\infty = 0) = 1$ *if and only if* $\mathbb{Q}(Z_\infty = \infty) = 1$*.*

A proof of (27) can be found in Athreya (2000).

5 Heuristics with spines

In this section we present some heuristics that motivate our use of the spine decomposition in section 6. We also state the main intermediate result in the proof of the almost-sure asymptotic shape, and sketch the main steps of the proof that follows in section 6. Fundamental to our later proofs is the fact that, with the optimal choice for λ, under $\tilde{\mathbb{Q}}_\lambda$ the space-type motion of the spine gives it a high probability of making the short climb.

So for $t \ll \tau$, let

$$A_\tau := \{\exists k : X_k(\tau) \sim \beta t; Y_k(\tau) \sim \kappa\sqrt{t}\} \qquad \text{and} \qquad A^\xi_\tau := \{\xi_\tau \sim \beta t; \eta_\tau \sim \kappa\sqrt{t}\}.$$

The major step in proving the lower bound for the asymptotic shape and growth (Theorem 3) is to prove the following result, which is a lower bound on the probability that a single particle makes the short climb:

Theorem 14 *For some fixed $\varepsilon > 0$,*

$$P^{x,y}(A_\tau) \geq \varepsilon \exp\{-\Theta(\beta,\kappa)t\},$$

where

$$\Theta(\beta,\kappa) := \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta - 8r)(a^2\kappa^4 + 4a\theta\beta^2)}}{4a\theta}.$$

The key step in the proof of this is to use the spine change of measure; for some function $g(\tau)$ we have

$$\begin{aligned} P^{x,y}(A_\tau) &= \mathbb{Q}^{x,y}_\lambda\left(\frac{\mathbf{1}_{A_\tau}}{\widehat{Z}^+_\lambda(\tau)}\right) = \tilde{\mathbb{Q}}^{x,y}_\lambda\left(\frac{\mathbf{1}_{A_\tau}}{\widehat{Z}^+_\lambda(\tau)}\right) \geq \tilde{\mathbb{Q}}^{x,y}_\lambda\left(\frac{\mathbf{1}_{A^\xi_\tau}}{\widehat{Z}^+_\lambda(\tau)}\right) \\ &\geq \tilde{\mathbb{Q}}^{x,y}_\lambda\left(\frac{\mathbf{1}_{A^\xi_\tau}}{\widehat{Z}^+_\lambda(\tau)}; \sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g(\tau)\right) \\ &\geq g(\tau)^{-1}\tilde{\mathbb{Q}}^{x,y}_\lambda\left(A^\xi_\tau; \sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g(\tau)\right). \end{aligned} \tag{28}$$

This is where we see the usefulness of the spine decomposition: the rest of the proof is technical, but essentially we just have to make the 'correct' choice for g in expression (28), for which we need estimates on $\tilde{\mathbb{Q}}^{x,y}_\lambda(Z^+_\lambda(t)^\alpha)$, for $\alpha \in (0,1)$. In the first line of the above system of inequalities it looks like we are losing a lot of mass when we replace the event A_τ with A^ξ_τ; but under $\tilde{\mathbb{Q}}^{x,y}_\lambda$ we have given the spine a high probability of being at large space and type positions at time τ, so this inequality is not as severe as it seems at first sight.

We emphasize that, although this paper is concerned only with a particular family of typed branching diffusions, adaptation of the above argument to other branching diffusions using different spine behaviour also yields very neat proofs of both new and some classical results: for a new proof of $\mathcal{L}^p$-boundedness of additive martingales see Hardy and Harris (2004*d*), see Hardy and Harris (2004*b*) for a proof of a large-deviations principle for standard branching Brownian motion using spine techniques, or the forthcoming work in Hardy and Harris (2004*c*) for an alternative spine approach to Theorem 14.

We first want to find the growth rate of $Z^+_\lambda(s)$ under $\mathbb{Q}_\lambda$, and for this we use the decomposition of $\widehat{Z}^+_\lambda$ in equation (26). By Jensen's inequality we have that, for $\alpha \in (0,1)$, $\widehat{Z}^+_\lambda(t)^\alpha$ is a submartingale. Then by the tower property and the conditional form of Jensen we have

$$\mathbb{Q}_\lambda(\widehat{Z}^+_\lambda(t)^\alpha) = \tilde{\mathbb{Q}}_\lambda(\widehat{Z}^+_\lambda(t)^\alpha) = \tilde{\mathbb{Q}}_\lambda\left(\tilde{\mathbb{Q}}_\lambda(\widehat{Z}^+_\lambda(t)^\alpha)\middle|\tilde{\mathcal{G}}_\infty\right) \leq \tilde{\mathbb{Q}}_\lambda\left(\tilde{\mathbb{Q}}_\lambda(\widehat{Z}^+_\lambda(t)|\tilde{\mathcal{G}}_\infty)^\alpha\right). \tag{29}$$

Conditioning on $\tilde{\mathcal{G}}_\infty$ allows us to use the decomposition of $Z^+_\lambda(t)$ given in (26), which simplifies the calculations regarding $Z^+_\lambda(t)$ significantly. Explicit calculation shows that the left-most expectation in (29) is finite for sufficiently small α, but we must have $\alpha \to 0$ as $t \to \infty$ for the expectation to remain finite. Before continuing with our calculations we will state a useful inequality from Neveu (1988):

Lemma 15 *For $\alpha \in (0,1)$ and real $x_i \geq 0$, we have*

$$\left(\sum_i x_i\right)^\alpha \leq \sum_i x_i^\alpha.$$

Then, applying this to decomposition of $Z^+_\lambda(t)$ given in (26) and using (29), we have

$$\mathbb{Q}_\lambda(\widehat{Z}^+_\lambda(t)^\alpha) \le \tilde{\mathbb{Q}}_\lambda \left\{ e^{-\alpha(\psi^+_\lambda y^2+\lambda x)} \left(\sum_{u\in\xi_t} e^{\alpha(\psi^+_\lambda \eta(S_u)^2+\lambda\xi(S_u)-E^+_\lambda S_u)} + e^{\alpha(\psi^+_\lambda \eta_t^2+\lambda\xi_t-E^+_\lambda t)} \right) \right\}$$

$$= e^{-\alpha(\psi^+_\lambda y^2+\lambda x)} \tilde{\mathbb{Q}}_\lambda \left(\sum_{u\in\xi_t} f(S_u)^\alpha + f(t)^\alpha \right), \tag{30}$$

where $f(t) := e^{\psi^+_\lambda \eta(t)^2+\lambda\xi(t)-E^+_\lambda t}$.

The $f(t)^\alpha$ term grows extremely rapidly under $\tilde{\mathbb{Q}}^{x,y}_\lambda$ because of the outward drifting Ornstein-Uhlenbeck process, however calculations similar to those of Harris and Williams (1996, Section 3) show that the expectations above are finite for sufficiently small α. If we condition on the path of the spine, $\mathcal{G}_\infty$, we see $f(t)^\alpha$ appearing in the sum term also:

$$\tilde{\mathbb{Q}}_\lambda \left(\sum_{u\in\xi_t} f(S_u) \right) = \tilde{\mathbb{Q}}_\lambda \left\{ \tilde{\mathbb{Q}}_\lambda \left(\sum_{u\in\xi_t} f(S_u) \middle| \mathcal{G}_\infty \right) \right\} = \tilde{\mathbb{Q}}_\lambda \left(\int_0^t R(\eta_s) f(s)^\alpha \, \mathrm{d}s \right),$$

and then by Fubini's theorem

$$\tilde{\mathbb{Q}}_\lambda \left(\int_0^t R(\eta_s) f(s)^\alpha \, \mathrm{d}s \right) = \int_0^t \tilde{\mathbb{Q}}_\lambda \left(R(\eta_s) f(s)^\alpha \right) \, \mathrm{d}s.$$

Now since $f(s)^\alpha$ grows much faster than $R(\eta_s)$ it seems reasonable to guess that, to exponential order, the growth rate of $Z^+_\lambda(t)^\alpha$ is the rate of growth of the spine term, that is:

$$\mathbb{Q}_\lambda(\widehat{Z}^+_\lambda(t)^\alpha) \sim \tilde{\mathbb{Q}}^{x,y}_\lambda \left\{ \left(e^{-(\psi^+_\lambda y^2+\lambda x)} e^{\psi^+_\lambda \eta_t^2+\lambda\xi_t-E^+_\lambda t} \right)^\alpha \right\}.$$

We use the following lemma to determine the asymptotic rate of growth of the spine term.

Lemma 16 *Let Y be a random variable on some probability space and suppose that there exists $\alpha_0 > 0$ such that $\mathbb{E}e^{\alpha Y} < \infty$ for all $\alpha < \alpha_0$. Then*

$$\lim_{\alpha\downarrow 0} \alpha^{-1} \log \mathbb{E}(e^{\alpha Y}) = \mathbb{E}Y.$$

Proof. Define $f(\alpha) := \mathbb{E}e^{\alpha Y}$. Then $f'(\alpha) = \mathbb{E}Ye^{\alpha Y}$ and we have $f(0) = 1$, $f'(0) = \mathbb{E}Y$. Now let $g(\alpha) := \log f(\alpha)$, so that

$$g'(\alpha) = \frac{f'(\alpha)}{f(\alpha)} \qquad \text{and} \qquad g'(0) = \mathbb{E}Y.$$

But by definition

$$\mathbb{E}Y = g'(0) = \lim_{\alpha\downarrow 0} \alpha^{-1} g(\alpha),$$

hence result. ■

Applying this result we have

$$\lim_{\alpha\downarrow 0} \alpha^{-1} \tilde{\mathbb{Q}}^{x,y}_\lambda \left\{ \left(e^{-(\psi^+_\lambda y^2+\lambda x)} e^{\psi^+_\lambda \eta_t^2+\lambda\xi_t-E^+_\lambda t} \right)^\alpha \right\}$$
$$= \tilde{\mathbb{Q}}^{x,y}_\lambda \left\{ \psi^+_\lambda(\eta_t^2 - y^2) + \lambda(\xi_t - x) - E^+_\lambda t \right\}.$$

We also know that, under $\tilde{\mathbb{Q}}^{x,y}_\lambda$,

$$\eta_t \sim N\left(ye^{\mu_\lambda t}, \frac{\theta}{2\mu_\lambda}\left(e^{2\mu_\lambda t} - 1\right) \right) \qquad \text{and} \qquad \xi_t \sim N\left(x + a\lambda \int_0^t \eta_s^2 \, \mathrm{d}s, a \int_0^t \eta_s^2 \mathrm{d}s \right),$$

so that

$$\tilde{\mathbb{Q}}^{x,y}_\lambda(\eta_t^2) = \left(\frac{\theta}{2\mu_\lambda} + y^2 \right) e^{2\mu_\lambda t} - \frac{\theta}{2\mu_\lambda}.$$

Moreover we can apply Fubini's theorem to obtain

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\int_0^t \eta_s^2 \,\mathrm{d}s\right) &= \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}(\eta_s^2)\,\mathrm{d}s \\ &= \left(\frac{\theta}{2\mu_\lambda}+y^2\right)\frac{e^{2\mu_\lambda t}}{2\mu_\lambda} - \left(\frac{\theta}{2\mu_\lambda}+y^2\right)\frac{1}{2\mu_\lambda} - \frac{\theta t}{2\mu_\lambda}.\end{aligned}$$

Hence

$$\begin{aligned}&\tilde{\mathbb{Q}}_\lambda^{x,y}\left\{\psi_\lambda^+(\eta_t^2-y^2)+\lambda(\xi_t-x)-E_\lambda^+ t\right\} \\ &\quad= \left(\psi_\lambda^+ + \frac{a\lambda^2}{2\mu_\lambda}\right)\left(\frac{\theta}{2\mu_\lambda}+y^2\right)e^{2\mu_\lambda t} \\ &\qquad - (\psi_\lambda^+ + a\lambda^2 t)\frac{\theta}{2\mu_\lambda} - \left(\frac{\theta}{2\mu_\lambda}+y^2\right)\frac{a\lambda^2}{2\mu_\lambda} - E_\lambda^+ t - \psi_\lambda^+ y^2 \\ &\quad\sim \left(\psi_\lambda^+ + \frac{a\lambda^2}{2\mu_\lambda}\right)\left(\frac{\theta}{2\mu_\lambda}+y^2\right)e^{2\mu_\lambda t}.\end{aligned}$$

These heuristics for the rate of growth of $\tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda^+(t)^\alpha)$ motivate our choice of growth rate $g(\tau)$ in equation (28).

The rest of the argument consists of showing that the probability in equation (28) is at least some $\varepsilon > 0$ and then optimising our choice of parameters to get the correct exponential part of equation (28). Consider the probability in equation (28); we know from large deviation theory that the probability that the spine follows the type path $y(s)$ for $0 \le s \le \tau$ and ends up at the space-type position $(\beta t, \kappa\sqrt{t})$ is roughly

$$\exp\left(-\frac{1}{2\theta}\int_0^\tau \{\dot{y}(s)-\mu_\lambda y(s)\}^2 \,\mathrm{d}s - \frac{\left(\beta t - \lambda a\int_0^\tau y(s)^2\,\mathrm{d}s\right)^2}{2a\int_0^\tau y(s)^2\,\mathrm{d}s}\right),$$

when $y(0)=0$, $y(\tau)=\kappa\sqrt{t}$ and t is very large. This is of order 1 if and only if $y=\widehat{y}$, where $\widehat{y}$ minimises the integral functional

$$\frac{1}{2\theta}\int_0^\tau \{\dot{y}(s)-\mu_\lambda y(s)\}^2\,\mathrm{d}s$$

with the given boundary conditions for $y(s)$, leading to

$$\widehat{y}(s) = \kappa\sqrt{t}\frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau},$$

and then the optimal λ must satisfy

$$\widehat{\lambda} = \frac{\beta t}{a\int_0^\tau \widehat{y}(s)^2\,\mathrm{d}s}.$$

The large deviation heuristics above suggest that when $\eta_\tau \sim \kappa\sqrt{t}$ we must actually have that η_s followed a path 'close' to the optimal path $\widehat{y}(s)$ to reach $\kappa\sqrt{t}$, and then the spatial drift is close enough to the optimal drift to ensure $\xi_\tau \sim \beta t$. We also need to deal with the condition

$$\sup_{s\in[0,\tau]} \widehat{Z}_\lambda^+(s) \le \exp\left\{\left(\psi_\lambda^+ + \frac{\lambda^2 a}{2\mu_\lambda}\right)\left(\frac{\theta}{2\mu_\lambda}+y^2\right)e^{2\mu_\lambda\tau}\right\},$$

but since Z_λ^+ is growing so quickly under $\tilde{\mathbb{Q}}_\lambda$ we might guess that the supremum in $[0,s]$ must be very close to the end point at time s. These ideas will be made rigorous in Section 6.

To get the correct lower bound for the proof of the asymptotic rate of growth we optimise our choice of parameters by taking λ to be the optimal value $\bar{\lambda}$ found in section 3:

$$\lambda = \bar{\lambda} := -\sqrt{\frac{\beta^2\theta(\theta-8r)}{a^2\kappa^4+4a\theta\beta^2}}, \qquad \mu_{\bar{\lambda}} := \frac{\kappa^2\sqrt{\theta(\theta-8r)}}{2\sqrt{\kappa^4+4\theta\beta^2/a}}$$

$$\text{and} \qquad \kappa^2 t := \left(\frac{\theta}{2\mu_{\bar\lambda}} + y^2\right) e^{2\mu_{\bar\lambda}\tau},$$

then with these parameters we have

$$\begin{aligned}\left(\psi^+_{\bar\lambda} + \frac{\bar\lambda^2 a}{2\mu_{\bar\lambda}}\right)\left(\frac{\theta}{2\mu_{\bar\lambda}} + y^2\right) e^{2\mu_{\bar\lambda}\tau} &= (\bar\lambda\beta + \kappa^2\psi^+_{\bar\lambda})t \\ &= \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta-8r)(a^2\kappa^4 + 4a\theta\beta^2)}}{4a\theta} \\ &=: \Theta(\beta,\kappa),\end{aligned}$$

yielding $P^{x,y}(A_\tau) \geq \varepsilon \exp(-\Theta(\beta,\kappa)t)$. It is important to note that choice of parameters fixes the relationship between t and τ. In section 6 we will use this result to recover the proof of the lower bound of Theorem 3.

6 Proof of Theorem 14

In this section we make the heuristics of the previous section rigorous. We first define, for large times t, with $\tau \ll t$,

$$A_\tau := \{\exists k : X_k(\tau) \in ((\beta-\delta)t, \beta t); Y_k(\tau) \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})\},$$

and similarly

$$A^\xi_\tau := \{\xi_\tau \in ((\beta-\delta)t, \beta t); \eta_\tau \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})\}.$$

Proof of Theorem 14. Throughout this section we take $\delta, \varepsilon \ll 1$ as small positive numbers, the actual values of which may change from line to line. With the growth rate heuristics of the previous section in mind we define

$$g_\varepsilon(\tau) := \exp\left\{\left(\psi^+_\lambda + \frac{\lambda^2 a}{2\mu_\lambda} + \varepsilon\right)\left(\frac{\theta}{2\mu_\lambda} + y^2\right) e^{2\mu_\lambda\tau}\right\},$$

and use this as the function $g(\tau)$ in (28) to obtain

$$P^{x,y}(A_\tau) \geq g_\varepsilon(\tau)^{-1} \tilde{\mathbb{Q}}^{x,y}_\lambda \left(A^\xi_\tau; \sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g_\varepsilon(\tau)\right). \tag{31}$$

Our strategy for the rest of this section is to show that the $\tilde{\mathbb{Q}}^{x,y}_\lambda$-probability in (31) is at least some $\varepsilon > 0$ for all t, and then optimise our choice of λ, τ and β so that the decay rate part of (31) matches the desired rate in the statement of the theorem.

Conditioning on the spine's path and birth times, $\tilde{\mathcal{G}}_\infty$, and then making use of some standard properties of conditional expectation we have

$$\begin{aligned}&\tilde{\mathbb{Q}}^{x,y}_\lambda \left(A^\xi_\tau; \sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g_\varepsilon(\tau)\right) \\ &\qquad = \tilde{\mathbb{Q}}^{x,y}_\lambda \left(\tilde{\mathbb{Q}}^{x,y}_\lambda \left(A^\xi_\tau; \sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g_\varepsilon(\tau) \middle| \tilde{\mathcal{G}}_\infty\right)\right) \\ &\qquad = \tilde{\mathbb{Q}}^{x,y}_\lambda \left(\mathbf{1}(A^\xi_\tau)\tilde{\mathbb{Q}}^{x,y}_\lambda \left(\sup_{s\in[0,\tau]} \widehat{Z}^+_\lambda(s) \leq g_\varepsilon(\tau) \middle| \tilde{\mathcal{G}}_\infty\right)\right),\end{aligned}$$

since A^ξ_τ is $\tilde{\mathcal{G}}_\infty$-measurable. We next observe that, conditional on $\tilde{\mathcal{G}}_\infty$, we can write $\widehat{Z}^+_\lambda(s)$ for $s \in [0,\tau]$ as

$$\widehat{Z}^+_\lambda(s) = e^{-(\psi^+_\lambda y^2 + \lambda x)} \left\{\sum_{u<\xi_s} Z^{(u)}_\lambda(s - S_u) + f(s)\right\}, \tag{32}$$

where n_t is the Poisson process of births along the spine, the $Z_\lambda^{(u)}$ are independent copies of Z_λ^+ and $f(s)$ is the contribution to $Z_\lambda^+(s)$ from the spine, which, conditional on $\tilde{\mathcal{G}}_\infty$, is a known function of s. Now if we could show, for $0 < \hat{\varepsilon} < \varepsilon$,

$$\sup_{s\in[0,\tau]} \widehat{f}(s) \le \frac{g_{\hat{\varepsilon}}(\tau)}{2} \quad \text{and} \quad \sup_{s\in[0,\tau]} \left(\widehat{Z}_\lambda^+(s) - \widehat{f}(s)\right) \le \frac{g_\varepsilon(\tau)}{2},$$

where

$$\widehat{f}(s) := e^{-(\psi_\lambda^+ y^2 + \lambda x)} f(s),$$

we would have

$$\sup_{s\in[0,\tau]} \widehat{Z}_\lambda^+(s) \le g_\varepsilon(\tau).$$

Hence

$$\begin{aligned}
&\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\sup_{s\in[0,\tau]} \widehat{Z}_\lambda^+(s) \le g_\varepsilon(\tau)\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right)\\
&\quad\ge \tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\sup_{s\in[0,\tau]} \widehat{f}(s) \le \frac{g_{\hat{\varepsilon}}(\tau)}{2};\ \sup_{s\in[0,\tau]}\left(\widehat{Z}_\lambda^+(s) - \widehat{f}(s)\right) \le \frac{g_\varepsilon(\tau)}{2}\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right)\\
&\quad= \tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\mathbf{1}\left(\sup_{s\in[0,\tau]} \widehat{f}(s) \le \frac{g_{\hat{\varepsilon}}(\tau)}{2}\right)\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\sup_{s\in[0,\tau]}\left(\widehat{Z}_\lambda^+(s) - \widehat{f}(s)\right) \le \frac{g_\varepsilon(\tau)}{2}\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right),
\end{aligned}$$

since, conditional on $\tilde{\mathcal{G}}_\infty$, the supremum of $\widehat{f}$ on $[0,\tau]$ is known.

Next we see from (32) that, *conditional on* $\tilde{\mathcal{G}}_\infty$,

$$\widehat{Z}_\lambda^+(s) - \widehat{f}(s)$$

is a submartingale, because the expectation of each of the martingales in the sum

(33) $$e^{-(\psi_\lambda^+ y^2 + \lambda x)} \sum_{u<\xi_s} Z_\lambda^{(u)}(s - S_u)$$

is constant, so the expectation of the sum cannot decrease, and in fact increases every time there is a birth on the spine. Then by Doob's submartingale inequality we have

$$\begin{aligned}
&\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\sup_{s\in[0,\tau]}\left(\widehat{Z}_\lambda^+(s) - \widehat{f}(s)\right) \le \frac{g_\varepsilon(\tau)}{2}\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\\
&\qquad= 1 - \tilde{\mathbb{Q}}_\lambda^{x,y}\left(\sup_{s\in[0,\tau]}\left(\widehat{Z}_\lambda^+(s) - \widehat{f}(s)\right) \ge \frac{g_\varepsilon(\tau)}{2}\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\\
&\qquad\ge 1 - \frac{2}{g_\varepsilon(\tau)}\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\widehat{Z}_\lambda^+(\tau) - \widehat{f}(\tau)\,\middle|\,\tilde{\mathcal{G}}_\infty\right).
\end{aligned}$$

We must note here that the expectation on the above line is not *a priori* finite. However, the expectation of each term in the sum (33) is bounded by $\sup_{s\in[0,\tau]}\widehat{f}(s)$, which we have control over via an indicator function and so we do not have to worry about this expectation blowing up.

So we need to show that

(34) $$\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\mathbf{1}\left(\sup_{s\in[0,\tau]} \widehat{f}(s) \le \frac{g_{\hat{\varepsilon}}(\tau)}{2}\right)\left(1 - \frac{2}{g_\varepsilon(\tau)}\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\widehat{Z}_\lambda^+(\tau) - \widehat{f}(\tau)\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right)\right) > \varepsilon',$$

for all τ. We claim that

(i) $\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\mathbf{1}\left(\sup_{s\in[0,\tau]} \widehat{f}(s) \le \frac{g_{\hat{\varepsilon}}(\tau)}{2}\right)\right) > \varepsilon', \qquad \forall\tau > 0$, and

(ii) $\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\mathbf{1}(A_\tau^\xi)\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s)\le\frac{g_{\tilde{\varepsilon}}(\tau)}{2}\right)\left(\frac{2}{g_\varepsilon(\tau)}\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\widehat{Z}_\lambda^+(\tau)-\widehat{f}(\tau)\Big|\,\tilde{\mathcal{G}}_\infty\right)\right)\right)$
$\to 0 \qquad \text{as } \tau\to\infty,$

which, when taken together, imply (34).

Proof of (i). We use arguments based on one-particle large deviations to show that (i) holds. Recall that under $\mathbb{Q}_\lambda^{x,y}$ the spine's type process, η_s, has the normal distribution

$$N\left(ye^{\mu_\lambda s},\frac{\theta}{2\mu_\lambda}\left(e^{2\mu_\lambda s}-1\right)\right),$$

so the probability that η_τ is 'close' to $\kappa\sqrt{t}$ in a large-deviations sense is

$$\begin{aligned}
&\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\eta_\tau\in[\kappa\sqrt{t},(\kappa+\varepsilon)\sqrt{t}]\right)\\
&=\mathbb{P}\left(\frac{\kappa\sqrt{t}-ye^{\mu_\lambda\tau}}{\sqrt{\frac{\theta}{2\mu_\lambda}}\sqrt{e^{2\mu_\lambda\tau}-1}}\le Z\le\frac{(\kappa+\varepsilon)\sqrt{t}-ye^{\mu_\lambda\tau}}{\sqrt{\frac{\theta}{2\mu_\lambda}}\sqrt{e^{2\mu_\lambda\tau}-1}}\right)\qquad\text{where } Z\sim N(0,1),\\
&\to\varepsilon'>0,
\end{aligned}$$

since $\sqrt{t}$ grows like $\exp(\mu_\lambda\tau)$. Now that we know $\eta_\tau\sim\kappa\sqrt{t}$ with strictly positive probability, we want to know what path it took to get there. The probability that η follows a path $y(s)$, with $y(0)=0$ and $y(\tau)=\kappa\sqrt{t}$ is approximately given by

$$\exp\left(-\frac{1}{2\theta}\int_0^\tau\{\dot{y}(s)-\mu_\lambda y(s)\}^2\,\mathrm{d}s\right). \tag{35}$$

We only consider paths with $y(0)=0$ here because later on we will use a coupling argument to say that the type position of a particle started at $y>0$ in type space will always be greater than or equal to that of a particle started at 0.

We now maximise (35) over all paths y with the correct start and finish points, which is equivalent to minimising the integral functional

$$\int_0^\tau F(y,\dot{y})\,\mathrm{d}s,\qquad\text{where}\qquad F(y,\dot{y}):=-\frac{1}{2\theta}\{\dot{y}(s)-\mu_\lambda y(s)\}^2.$$

Some Euler-Lagrange optimisation, with the boundary conditions $y(0)=0$ and $y(\tau)=\kappa\sqrt{t}$, gives the optimal path as

$$\widehat{y}(s)=\kappa\sqrt{t}\frac{\sinh\mu_\lambda s}{\sinh\mu_\lambda\tau}.$$

Now we find that

$$\left(\frac{\mathrm{d}\widehat{y}}{\mathrm{d}s}-\mu_\lambda\widehat{y}\right)^2=\frac{\mu_\lambda^2\kappa^2t}{\sinh^2\mu_\lambda\tau}e^{-2\mu_\lambda s},$$

and hence the probability that η_s follows the optimal path on $[0,\tau]$ satisfies the large deviation principle with rate function given by

$$\begin{aligned}
-\frac{1}{2\theta}\int_0^\tau\left\{\frac{\mathrm{d}\widehat{y}}{\mathrm{d}s}-\mu_\lambda\widehat{y}\right\}^2\mathrm{d}s&=-\frac{1}{2\theta}\int_0^\tau\frac{\mu_\lambda^2\kappa^2t}{\sinh^2\mu_\lambda\tau}e^{-2\mu_\lambda s}\,\mathrm{d}s\\
&=\frac{-\mu_\lambda}{4\theta}(1-e^{-2\mu_\lambda\tau})\frac{\kappa^2t}{\sinh^2\mu_\lambda\tau}=:-I(\tau).
\end{aligned}$$

Now we know that $\kappa^2t\sim e^{2\mu_\lambda\tau}$, so $I(\tau)$ is bounded as $\tau\to\infty$.

Now consider the perturbed path $y_\delta(s):=\widehat{y}(s)+\delta\tau$. We find that the probability that η_s follows this path has large deviation rate function

$$-\frac{1}{2\theta}\int_0^\tau\{\dot{y}_\delta(s)-\mu_\lambda\dot{y}_\delta\}^2\,\mathrm{d}s=-I(\tau)-\frac{\mu_\lambda^2\delta^2}{2\theta}\tau^3+O(\tau).$$

This means that, if the spatial position of the spine is near $\kappa\sqrt{t}$ at time τ, the probability that it followed the perturbed path to get there decays exponentially (for all $\delta > 0$). Hence, on any mesh of $\mathbb{R}^+$, there is at least a fixed probability $\varepsilon' > 0$ that $\eta_\tau \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})$ and moreover, by the Borel Cantelli lemma, on this mesh we eventually have that η_s is within $\delta\tau$ of the optimal path, for all $s \in [0, \tau]$.

For the spatial position at time τ, ξ_τ, we recall that the optimal λ satisfies

$$\widehat{\lambda} = \frac{\beta t}{a \int_0^\tau \widehat{y}(s)^2 \,\mathrm{d}s}.$$

Moreover, given that η_s has followed the path $y(s)$ over the time interval $[0, \tau]$, the probability that $\xi_\tau = \beta t$ is approximately

$$\exp\left(- \frac{\left(\beta t - \lambda a \int_0^\tau y(s)^2 \,\mathrm{d}s\right)^2}{2a \int_0^\tau y(s)^2 \,\mathrm{d}s} \right),$$

and this decays exponentially for all spatial positions other than

$$\lambda a \int_0^\tau y(s)^2 \,\mathrm{d}s.$$

Now as long as $\eta_\tau \in [\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t}]$, the large deviations arguments above show that for sufficiently large t, there is some small $\varepsilon' > 0$ such that, with probability 1, η_s follows a path that is at worst

$$y(s) = \kappa\sqrt{t}\frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau} - \delta\tau, \tag{36}$$

and the corresponding total spatial drift up to time τ is at worst

$$\lambda a \int_0^\tau \left\{ \kappa\sqrt{t}\frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau} - \delta\tau \right\}^2 \mathrm{d}s = \beta t + o(t), \tag{37}$$

provided we change measure with $\widehat{\lambda}$, i.e. the optimal choice for λ. Since the probability that ξ_τ is anywhere other than

$$\lambda a \int_0^\tau y(s)^2 \,\mathrm{d}s$$

decays exponentially we have shown that, on any mesh of $\mathbb{R}^+$ and for sufficiently large τ, we have $\xi_\tau \in [(\beta-\delta)t, (\beta+\delta)t]$ with probability 1 (using Borel-Cantelli again). Of course this is conditional on the event $\{\eta_\tau \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})\}$, but we know this event has at last probability $\varepsilon' > 0$ of occurring. The same argument applied to times $s \in [0, \tau]$ shows that in fact ξ_s reaches spatial position βt at time τ by constant spatial drift (with respect to the clock $A(\eta_s)$), modulo some error $\pm\delta t$.

We now know that for sufficiently large t, as long as the type position of the spine at time τ is in the interval $[\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t}]$, it must have followed a path close to the optimal path $\widehat{y}(s)$ and also that the spatial position of the spine is close to βt. Hence

$$\tilde{\mathbb{Q}}_\lambda^{x,y}(A_\tau^\xi) = \tilde{\mathbb{Q}}_\lambda^{x,y}\left(\eta_\tau \in [\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t}]\right) > \varepsilon' > 0,$$

for all sufficiently large $\tau > 0$. Notice that in this proof that $\tilde{\mathbb{Q}}_\lambda^{x,y}(A_\tau^\xi) > \varepsilon'$ we not only see that there is some probability $\varepsilon' > 0$ that the spine is at the desired space-type position at time τ, but that, for sufficiently large τ, (η_s, ξ_s) followed a path very close to the optimal path for $s \in [0, \tau]$. Moreover if we know that $\eta_\tau \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})$, then choosing $0 < \delta < \varepsilon$ such that

$$\sup_{s\in[0,\tau]} |\widehat{y}(s) - \eta(s)| < \delta\tau \qquad \text{for all large enough } \tau,$$

then we guarantee that (eventually) we will have, for some small time ν,

$$\eta_s \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t}) \qquad \text{for all } s \in [\tau-\nu, \tau].$$

Also, since the spatial motion of the spine under $\tilde{\mathbb{Q}}_\lambda^{x,y}$ is a time change of drifting Brownian motion, we also have that ξ_s stays close to βt during the small time interval $[\tau-\nu, \tau]$. So A_τ^ξ is really the event that the spine is arbitrarily close to $(\beta t, \kappa\sqrt{t})$ at time τ *and* is at that location during the small time interval $[\tau-\nu,\tau]$. This fact is important for the last stage of this proof.

Finally, we need to ensure that the condition on the supremum of $\widehat{f}(s)$ holds. So long as $\eta_\tau \in (\kappa\sqrt{t}, (\kappa+\varepsilon)\sqrt{t})$, with probability 1 we must eventually have have (recall that the spatial drift is negative):

$$\sup_{s\in[0,\tau]} \eta_s \le (\kappa+\varepsilon)\sqrt{t} + \delta\tau \le (\kappa+\varepsilon')\sqrt{t} \qquad \text{and} \qquad \inf_{s\in[0,\tau]} \xi_s \ge (\beta-\delta')t,$$

because we know what paths η and ξ followed (modulo the ε and δ error terms). Also note that we can think of g_ε as a function of t or τ, because the relationship between them is fixed. In terms of t, g_ε becomes

$$g_\varepsilon(t) := \exp\left\{\left(\psi^+_{\bar\lambda} + \frac{\bar\lambda^2 a}{2\mu_{\bar\lambda}} + \varepsilon'\right)\kappa^2 t\right\} = \exp\left((\kappa^2\psi^+_{\bar\lambda} + \bar\lambda\beta + \tilde\varepsilon')t\right),$$

and here we are choosing the optimal value for λ as well as fixing the relationship between t and τ.

So to show

$$\tilde{\mathbb{Q}}^{x,y}_\lambda\left(\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s) \le \frac{g_{\tilde\varepsilon}(\tau)}{2}\right)\right) > 0$$

we want

$$\begin{aligned}&\sup_{s\in[0,\tau]} e^{-(\psi^+_\lambda y^2+\lambda x)}\exp\left(\psi^+_\lambda\eta_s^2 + \lambda\xi_s - E^+_\lambda s\right)\\ &\qquad\le \frac{1}{2}g_{\tilde\varepsilon}(\tau) = \frac{1}{2}\exp\left((\kappa^2\psi^+_\lambda + \lambda\beta + \tilde\varepsilon')t\right).\end{aligned} \tag{38}$$

But we know that, $\tilde{\mathbb{Q}}^{x,y}_\lambda$-almost surely, for sufficiently large τ,

$$\begin{aligned}&\sup_{s\in[0,\tau]} e^{-(\psi^+_\lambda y^2+\lambda x)}\exp\left(\psi^+_\lambda\eta_s^2 + \lambda\xi_s - E^+_\lambda s\right)\\ &\qquad\le e^{-(\psi^+_\lambda y^2+\lambda x)}\exp\left(\psi^+_\lambda(\kappa+\varepsilon')^2 t + \lambda(\beta-\delta')t - E^+_\lambda\tau\right),\end{aligned}$$

so that (38) holds if

$$\exp\left(\psi^+_\lambda(2\varepsilon'\kappa + \varepsilon'^2)t - \lambda\delta' t\right) \le \frac{1}{2}\exp\left(\tilde\varepsilon' t\right),$$

and by taking ε' and δ' small enough we can satisfy this for any $\tilde\varepsilon' > 0$. We have already seen that $\tilde{\mathbb{Q}}^{x,y}_\lambda(A^\xi_\tau) > 0$ for any δ' and ε', so (i) is proved. ■

Proof of (ii). Consider first the expectation of the 'sum term'. We have

$$\begin{aligned}\tilde{\mathbb{Q}}^{x,y}_\lambda\left(\widehat{Z}^+_\lambda(\tau) - \widehat{f}(\tau)\,\middle|\,\mathcal{G}_\infty\right) &= e^{-(\psi^+_\lambda y^2+\lambda x)}\sum_{u<\xi_\tau}\exp\left(\psi^+_\lambda\eta(S_u)^2 + \lambda\xi(S_u) - E^+_\lambda S_u\right)\\ &\le n_\tau \sup_{s\in[0,\tau]}\widehat{f}(s),\end{aligned}$$

where n_τ is the number of births along the spine up to time τ and since each of the terms in the sum is bounded by the supremum of $\widehat{f}$ on $[0,\tau]$. Hence,

$$\begin{aligned}&\tilde{\mathbb{Q}}^{x,y}_\lambda\left(\mathbf{1}(A^\xi_\tau)\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s) \le \frac{g_{\tilde\varepsilon}(\tau)}{2}\right)\left(\frac{2}{g_\varepsilon(\tau)}\tilde{\mathbb{Q}}^{x,y}_\lambda\left(\widehat{Z}^+_\lambda(\tau) - \widehat{f}(\tau)\,\middle|\,\tilde{\mathcal{G}}_\infty\right)\right)\right)\\
&\qquad\le \tilde{\mathbb{Q}}^{x,y}_\lambda\left(\mathbf{1}(A^\xi_\tau)\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s) \le \frac{g_{\tilde\varepsilon}(\tau)}{2}\right)n_\tau\frac{g_{\tilde\varepsilon}(\tau)}{g_\varepsilon(\tau)}\right)\\
&\qquad= \tilde{\mathbb{Q}}^{x,y}_\lambda\left(\mathbf{1}(A^\xi_\tau)\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s) \le \frac{g_{\tilde\varepsilon}(\tau)}{2}\right)n_\tau e^{-(\varepsilon-\tilde\varepsilon)\kappa^2 t}\right)\\
&\qquad= e^{-(\varepsilon-\tilde\varepsilon)\kappa^2 t}\tilde{\mathbb{Q}}^{x,y}_\lambda\left\{\tilde{\mathbb{Q}}^{x,y}_\lambda\left(\mathbf{1}(A^\xi_\tau)\mathbf{1}\left(\sup_{s\in[0,\tau]}\widehat{f}(s) \le \frac{g_{\tilde\varepsilon}(\tau)}{2}\right)n_\tau\,\middle|\,\mathcal{G}_\infty\right)\right\},\end{aligned} \tag{39}$$

where in the last line above we are conditioning on $\mathcal{G}_\infty$, the path of the spine in type space (*not* including the birth times), and using the tower property.

But now, conditional on $\mathcal{G}_\infty$, n_τ is a Poisson random variable with mean given by

$$\int_0^\tau 2\left(r\eta_s^2+\rho\right)\,\mathrm{d}s,$$

and we calculate the outer expectation in (39) using Fubini's theorem:

$$\begin{aligned}
&\tilde{\mathbb{Q}}_\lambda^{x,y}\left(\int_0^\tau 2\left(r\eta_s^2+\rho\right)\,\mathrm{d}s\right)\\
&\quad=\int_0^\tau 2r\tilde{\mathbb{Q}}_\lambda^{x,y}(\eta_s^2)\,\mathrm{d}s+2\rho\tau \qquad \text{(Fubini)}\\
&\quad=\frac{r}{\mu_\lambda}\left(\frac{\theta}{2\mu_\lambda}+y^2\right)e^{2\mu_\lambda\tau}-\left(\frac{\theta}{2\mu_\lambda}+y^2\right)\frac{r}{\mu_\lambda\tau}-\frac{r\theta\tau}{\mu_\lambda}+2\rho\tau\\
&\quad=\frac{r}{\mu_\lambda}\kappa^2t-\left(\frac{\theta}{2\mu_\lambda}+y^2\right)\frac{r}{\mu_\lambda\tau}-\frac{r\theta\tau}{\mu_\lambda}+2\rho\tau=\frac{r}{\mu_\lambda}\kappa^2t+o(\tau).
\end{aligned}$$

So the $\tilde{\mathbb{Q}}_\lambda^{x,y}$-expectation of n_τ only grows linearly in t. Then since $\varepsilon-\tilde{\varepsilon}>0$, the expectation (39) tends to 0 as $t\to\infty$. ■

Hence, taking the limit as $\varepsilon\to 0$ in (31),

$$\begin{aligned}
P^{x,y}(A_\tau)&\geq\varepsilon\exp\left\{-\left(\psi_{\bar\lambda}^{+}+\frac{\bar\lambda^2a}{2\mu_{\bar\lambda}}\right)\left(\frac{\theta}{2\mu_{\bar\lambda}}+y^2\right)e^{2\mu_{\bar\lambda}\tau}\right\}\\
&=\varepsilon\exp\left\{-(\bar\lambda\beta+\kappa^2\psi_{\bar\lambda}^{+})t\right\}=\varepsilon\exp(-\Theta(\beta,\kappa)t),
\end{aligned}$$

where we are changing measure with the optimal martingale $Z_{\bar\lambda}^{+}(t)$. ■

7 Martingale results

We rely heavily on martingales to prove the results in the last section. We state key martingale theorems separately in this section.

We remember from Harris and Williams (1996) that E_λ (also written $E(\lambda)$) and $\Delta(\gamma)$ are Legendre conjugates with

(40) $$\Delta(\gamma)=\inf_{\lambda<0}\{E(\lambda)+\lambda\gamma\},\quad E(\lambda)=\sup_{\gamma>0}\{\Delta(\gamma)-\gamma\lambda\},$$

If, for $\lambda_{\min}<\lambda<0$, we write γ_λ for the γ value which achieves the supremum on the right-hand side of equation (40), then the functions $\lambda\mapsto\gamma_\lambda$ from $(-\lambda_{\min},0)$ to $(0,\infty)$, and $\gamma\mapsto\lambda_\gamma$ from $(0,\infty)$ to $(-\lambda_{\min},0)$ are inverses of each other and, of course, λ_γ is the λ value which achieves the infimum on the left-hand side of equation (40). In addition, we note that

(41) $$\gamma_\lambda=-E'(\lambda)=\sqrt{\frac{\theta a^2\lambda^2}{\theta-8r-4a\lambda^2}},$$

that $E(\lambda)$ and $\Delta(\gamma)$ are convex functions and that

(42) $$\begin{aligned}\tilde{c}(\theta)&=\sup\{\gamma:\Delta(\gamma)>0\}=\inf\{-E(\lambda)/\lambda:\lambda_{\min}<\lambda<0\}\\&=\inf\{c_\lambda^-:\lambda_{\min}<\lambda<0\}=c_{\tilde\lambda(\theta)}^-,\end{aligned}$$

where

(43) $$c_\lambda^-:=-E_\lambda/\lambda,\quad\text{and}\quad\tilde\lambda(\theta):=-\sqrt{\frac{2(\theta-8r)(\theta\rho+2\rho^2+r\theta)}{a(\theta+4\rho)^2}}\in(\lambda_{\min},0).$$

and a formula for $\tilde{c}(\theta)$ is given in equation (4).

7.1 The 'ground-state' martingales

In Harris and Williams (1996), we proved the almost sure speed of the spatially left-most particle by making use of the following martingales:

Lemma 17 *Let $\lambda \in (\lambda_{\min}, 0)$. For $t \geq 0$, define*

$$Z_\lambda^-(t) := \sum_{k=1}^{N(t)} \exp\left(\psi_\lambda^- Y_k(t)^2 + \lambda\left[X_k(t) + c_\lambda^- t\right]\right). \tag{44}$$

This defines a martingale Z_λ^- (under each $P^{x,y}$ measure).

Notice that these martingales *immediately* give the upper bound $D(\gamma) \leq \Delta(\gamma)$. Also, since the martingale is non-negative it must converge. It is easy to check that the function c^- is convex on $(\lambda_{\min}, 0)$, and achieves its minimum at the unique point $\tilde{\lambda}(\theta)$. In Harris and Williams (1996), we used this simple geometric fact and an idea of Neveu (1988) in proving the following:

Theorem 18 *The martingale Z_λ^- is uniformly integrable and has an almost sure strictly positive limit if $\lambda \in \big(\tilde{\lambda}(\theta), 0\big)$.*

7.2 The 'excited-state' martingales and an important convergence theorem

There are further families of martingales related to this problem that prove very useful in various situations. Define

$$h_{n,\lambda}(y) := \sqrt{\frac{\mu_\lambda^{\frac{1}{2}}}{\theta^{\frac{1}{2}} n! 2^n}} H_n\left(\sqrt{\frac{\mu_\lambda}{\theta}} y\right),$$

where H_n is the n^{th} Hermite polynomial so that

$$H_n(z) = (-1)^n e^{z^2} \frac{\mathrm{d}^n}{\mathrm{d}z^n}\left(e^{-z^2}\right)$$

so in particular, $H_0(z) \equiv 1$, $H_1(z) = 2z$, $H_2(z) = 4z^2 - 2$, etc.

See Harris (2000) if you need convincing of the following martingale properties:

Lemma 19 *Let $\lambda \in \big(\lambda_{\min}, 0\big)$. For each $n \in \{0, 1, \dots\}$ and $t \in [0, \infty)$*

$$Z_{n,\lambda}(t) := \sum_{k=1}^{N(t)} e^{n\mu_\lambda t} h_{n,\lambda}(Y_k(t)) e^{\psi_\lambda^- Y_k(t)^2 + \lambda X_k(t) - E_\lambda t}$$

defines a martingale $Z_{n,\lambda}$ for each $P^{x,y}$ starting law.

It was the study of these 'excited-state' martingales over the parameter range $(\tilde{\lambda}(\theta), 0)$ (where the 'ground-state' martingale is uniformly integrable) in Harris (2000) that facilitated the proof of the following convergence theorem by making use of Hermite martingale expansions.

Theorem 20 *Let $\lambda \in \big(\tilde{\lambda}(\theta), 0\big)$ and $\alpha < 1/4$. For each $P^{x,y}$ starting law and every continuous bounded function $f : \mathbb{R} \mapsto \mathbb{R}$, we have*

$$\sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda\left(X_k(t) + c_\lambda^- t\right)} \overset{a.s.}{\to} f_0\, Z_\lambda^-(\infty).$$

where

$$f_0 := \int_{\mathbb{R}} f(y) e^{\alpha y^2} \Psi_{0,\lambda}(y) \phi(y)\, \mathrm{d}y \tag{45}$$

In this paper, we require a corollary to this theorem which specifies more precisely which particles contribute to the final limit.

Corollary 21 *For $\tilde{\lambda}(\theta) < \lambda < 0$ and $\alpha < 1/4$. For each $P^{x,y}$ starting law and every continuous bounded function $f : \mathbb{R} \mapsto \mathbb{R}$, we have for every $\varepsilon > 0$*

$$\sum_{k=1}^{N(t)} f\left(Y_k(t)\right) e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{|X_k(t)/t + \gamma_\lambda| < \varepsilon\right\} \to f_0\, Z_\lambda^-(\infty)$$

where $\gamma_\lambda = -E'(\lambda)$ and f_0 is given at equation (45).

This last result will enable us to show that the almost sure growth rate is at least as large as the expected growth rate, $D(\gamma) \geq \Delta(\gamma)$. It is easy to see from 21 that when $Z_\lambda^-(\infty) > 0$, there must exist at least one particle near to $\gamma_\lambda t$ in space. Further, because of the decay rate of each term in the sum over particles in 21, it is straightforward to improve this to get the required exponential numbers of particles, $\exp \Delta(\gamma) t$, near $\gamma_\lambda t$ for large times (as long as $Z_\lambda^-(\infty) > 0$).

7.3 Rates of convergence to zero

Define

$$E_\lambda^+ := \rho + \theta \psi_\lambda^+, \qquad c_\lambda^+ := -E_\lambda^+/\lambda.$$

We will prove the following results concerning the rate at which various martingales converge to zero. The result for the Z_λ^+ martingales will be crucial to gain the asymptotic shape bounds for the branching diffusion.

Theorem 22 *For every starting law, $P^{x,y}$:*

(i) if $\lambda \in (\lambda_{\min}, 0)$ then

$$\frac{\log Z_\lambda^\pm(t)}{t} \to \lambda(c_\lambda^\pm - c_\lambda^*) \qquad a.s.$$

where

$$c_\lambda^* := \begin{cases} \tilde{c}(\theta) & \text{if } \lambda_{\min} < \lambda \leq \tilde{\lambda}(\theta), \\ c_\lambda^- & \text{if } \tilde{\lambda}(\theta) \leq \lambda < 0. \end{cases}$$

(ii) if $\lambda \in \left(\lambda_{\min}, \tilde{\lambda}(\theta)\right)$ and $n \in \{1, 2, \ldots\}$ then for all $\varepsilon > 0$

$$e^{-\varepsilon t}\, e^{-\lambda\left(c_\lambda^- - \tilde{c}(\theta)\right)t} Z_{n,\lambda}(t) e^{-n\mu_\lambda t} \to 0 \qquad a.s.$$

The rate of convergence of the Z_λ^+ martingale in part (i) of 22 will immediately give the upper bound on the almost sure growth rate, $D(\gamma, \kappa) \leq \Delta(\gamma, \kappa)$. We also comment that *if* 21 was true for all $\alpha < \psi_\lambda^+$, then we could have gained this upper bound at that point. Although the result 21 is only proved when $\alpha < 1/4$ where we can utilise suitable Hermite expansions, we conjecture that it may well hold beyond this point for all values $\alpha < \psi_\lambda^+$.

Corollary 23 *For every starting law, $P^{x,y}$,*

(i) if $\lambda \in (\lambda_{\min}, 0)$ then $Z_\lambda^+(t) \to 0$ almost surely.

(ii) if $\lambda \in \left(\lambda_{\min}, \tilde{\lambda}(\theta)\right)$ and $n \in \{0, 1, \ldots\}$ then $e^{-n\mu_\lambda t} Z_{n,\lambda}(t) \to 0$ almost surely.

7.4 Martingales at the 'critical' parameter value

Although we will *not* offer a proof in this paper, it should be found that the 'usual' situation occurs where the critical martingale tends to zero,

$$Z^-_{\tilde{\lambda}(\theta)}(t) \to 0 \qquad a.s.$$

but the 'derivative' martingale has a strictly positive limit,

$$\lim_{t\to\infty} \left\{ \frac{\partial Z^-_\lambda}{\partial\lambda}(t)\bigg|_{\lambda=\tilde{\lambda}(\theta)} \right\} > 0 \qquad a.s.$$

See Neveu (1988) for this type of behaviour in the standard branching Brownian motion case.

8 Proof of the martingale results

Proof of corollary 21. Let $\varepsilon > 0$ be small, $\mu := \lambda - \varepsilon$, $\lambda, \mu \in (\tilde{\lambda}(\theta), 0)$, f be a positive, continuous bounded function, $\alpha < 1/4$ and recall $\gamma_\mu > \gamma_\lambda$. Then observe that

$$\begin{aligned}
&\sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{X_k(t) < -\gamma_\mu t\right\} \\
&\qquad \le e^{(E_\mu - E_\lambda - \varepsilon\gamma_\mu)t} \sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \mu X_k(t) - E_\mu t} \mathbf{1}\left\{X_k(t) < -\gamma_\mu t\right\} \\
&\qquad \le \left(\sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \mu X_k(t) - E_\mu t} \right) e^{-(E_\lambda - E_\mu + (\lambda-\mu)\gamma_\mu)t}
\end{aligned}$$

Recall that E_λ is convex with $E''_\lambda \ge 0$ and $E'_\lambda = \gamma_\lambda$, so from the Taylor expansion

$$E_\lambda - E_\mu + (\mu - \lambda)E'_\lambda = (\mu-\lambda)^2/2\, E''_\lambda + o\big((\mu-\lambda)^2\big)$$

Then taking $\varepsilon > 0$ small enough so that $E_\lambda - E_\mu + (\lambda - \mu)\gamma_\mu > 0$ and using theorem 20 we find that for any $\delta > 0$

$$\limsup_{t\to\infty} \sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{X_k(t) < -(\gamma_\lambda + \delta)t\right\} = 0.$$

Similarly, we can show

$$\limsup_{t\to\infty} \sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{X_k(t) > -(\gamma_\lambda - \delta)t\right\} = 0$$

and hence combining with convergence theorem 20 we have

$$\begin{aligned}
f_0\, Z^-_\lambda(\infty) &= \lim_t \sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \\
&= \lim_t \sum_{k=1}^{N(t)} f\left(Y_k(t)\right)\, e^{\alpha Y_k(t)^2 + \lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{|X_k(t)/t + \gamma_\lambda| < \delta\right\}
\end{aligned}$$

since the only contribution to the limit comes from the particles near $-\gamma_\lambda t$ in space. ■

Proof of theorem 22. (a) Remember each $Z^\pm_\lambda$ is a positive martingale. Remembering a useful technique brought to our attention in Neveu (1988), let $p \in (0, 1)$, then by Jensen's inequality, $Z^\pm_\lambda(t)^p$ is a supermartingale and since for $u, v > 0$ we have

$$(u + v)^p \le u^p + v^p$$

then

$$
\begin{aligned}
Z^{\pm}_{\lambda}(t)^p &= \left| \sum_{k=1}^{N(t)} e^{\psi^{\pm}_{\lambda} Y_k(t)^2 + \lambda\left(X_k(t) + c^{\pm}_{\lambda} t\right)} \right|^p \\
&\leq \left(\sum_{k=1}^{N(t)} e^{p\psi^{\pm}_{\lambda} Y_k(t)^2 + p\lambda\left(X_k(t) + c^{-}_{p\lambda} t\right)} \right) e^{p\lambda\left(c^{\pm}_{\lambda} - c^{-}_{p\lambda}\right)t}.
\end{aligned}
$$

For any $\varepsilon > 0$, Doob's supermartingale inequality says

$$
\begin{aligned}
P\left(\sup_{s \leq u \leq s+t} Z^{\pm}_{\lambda}(u)^p > \varepsilon^p \right) &\leq \frac{E Z^{\pm}_{\lambda}(s)^p}{\varepsilon^p} \\
&\leq \varepsilon^{-p} \left\{ E \sum_{k=1}^{N(s)} e^{p\psi^{\pm}_{\lambda} Y_k(s)^2 + p\lambda\left(X_k(s) + c^{-}_{p\lambda} s\right)} \right\} e^{p\lambda(c^{\pm}_{\lambda} - c^{-}_{p\lambda})s}
\end{aligned}
$$

Then

$$
\begin{aligned}
&P\left(\sup_{s \leq u \leq s+t} e^{\delta u} Z^{\pm}_{\lambda}(u) > \varepsilon \right) \\
&\quad \leq P\left(\sup_{s \leq u \leq s+t} Z^{\pm}_{\lambda}(u)^p > e^{-p\delta(s+t)} \varepsilon^p \right) \\
&\quad \leq \varepsilon^{-p} e^{p\delta t} \left\{ E \sum_{k=1}^{N(s)} e^{p\psi^{\pm}_{\lambda} Y_k(s)^2 + p\lambda\left(X_k(s) + c^{-}_{p\lambda} s\right)} \right\} e^{p\left(\lambda(c^{\pm}_{\lambda} - c^{-}_{p\lambda}) + \delta\right)s}
\end{aligned}
$$

Now, if we can choose $p \in (0,1)$ such that $\lambda(c^{\pm}_{\lambda} - c^{-}_{p\lambda}) + \delta < 0$ and $p\psi^{\pm}_{\lambda} < \psi^{+}_{p\lambda}$, we must have $e^{\delta u} Z^{\pm}_{\lambda}(u) \to 0$ almost surely by using a familiar Borel-Cantelli argument. (The condition $p\psi^{\pm}_{\lambda} < \psi^{+}_{p\lambda}$ guarantees that the expectation in the last line above tends to a finite limiting value, hence stays bounded over all times s, as can be checked by using formula (12), for example.)

For all $0 \leq p < 1$ we find $p\psi^{\pm}_{\lambda} < \psi^{+}_{p\lambda}$. Considering the graph of $c^{\pm}_{\lambda}$ we quickly see that for $\lambda \in [\tilde{\lambda}(\theta), 0)$ taking p as close to 1 as we like gives the best rate. For $\lambda \in [\lambda_{\min}, \tilde{\lambda}(\theta))$ we can choose p so that $p\lambda = \tilde{\lambda}(\theta)$ to yield the best rate.

Recall, we already know $Z^{-}_{\lambda}(\infty) > 0$ when $\lambda \in (\tilde{\lambda}(\theta), 0)$ from theorem 18. Then, so far, we have proved the following:

Lemma 24 *For every starting law, $P^{x,y}$, and for all $\varepsilon > 0$, if $\lambda \in (\lambda_{\min}, 0)$ then*

$$
e^{-\varepsilon t}\, e^{-\lambda(c^{\pm}_{\lambda} - c^{*}_{\lambda})t} Z^{\pm}_{\lambda}(t) \to 0 \qquad a.s.
$$

where

$$
c^{*}_{\lambda} := \begin{cases} \tilde{c}(\theta) & \text{if } \lambda_{\min} < \lambda \leq \tilde{\lambda}(\theta), \\ c^{-}_{\lambda} & \text{if } \tilde{\lambda}(\theta) \leq \lambda < 0. \end{cases}
$$

It is clear that this gives the required upper bound of

$$
\limsup_{t \to \infty} \frac{\log Z^{\pm}_{\lambda}(t)}{t} \leq \lambda(c^{\pm}_{\lambda} - c^{*}_{\lambda}).
$$

Now, for any $\varepsilon > 0$, if $\lambda \in \left(\lambda_{\min}, \tilde{\lambda}(\theta)\right]$ then

$$
e^{\varepsilon t} \sum_{k=1}^{N(t)} e^{\psi^{\pm}_{\lambda} Y_k(t)^2 + \lambda(X_k(t) + \tilde{c}(\theta)t)} \geq e^{\lambda(L_t + \tilde{c}(\theta)t) + \varepsilon t} \to \infty \qquad a.s.
$$

since we know that $L_t := \inf_{k\le N(t)} X_k(t)$ satisfies $L(t)/t \to -\tilde{c}(\theta)$ a.s. Otherwise, with $\lambda \in (\tilde{\lambda}(\theta), 0)$,

$$e^{\varepsilon t} \sum_{k=1}^{N(t)} e^{\psi_\lambda^+ Y_k(t)^2 + \lambda(X_k(t)+c_\lambda^- t)} \geq e^{\varepsilon t} Z_\lambda^-(t) \to \infty \qquad a.s.$$

since here $Z_\lambda^-(\infty) > 0$ a.s. Thus, in all cases,

$$\liminf_{t\to\infty} \frac{\log Z_\lambda^\pm(t)}{t} \geq \lambda(c_\lambda^\pm - c_\lambda^*).$$

and we have completed the proof of the first part of theorem 22.

(b) For the 'excited-state' martingales we can make a comparison to the 'ground-state' martingales as follows. Fix $n \in \{1, 2, \dots\}$ and let $\lambda \in \big[\lambda_{\min}, \tilde{\lambda}(\theta)\big)$. Now $\Psi_n(y) = h_n(y)\, e^{\psi_\lambda^- y^2}$ where h_n is a polynomial of degree n. Hence we can find a constant such that $|h_n(y)| \leq C\,|y|^n$ for all $y \in \mathbb{R}$. Define $R_t := \sup_{k\le N(t)} Y_k(t)$, then

$$\begin{aligned}
|Z_{n,\lambda}(t)| &= \left| \sum_{k=1}^{N(t)} \Psi_n\big(Y_k(t)\big)\, e^{\lambda\left(X_k(t)+c_\lambda^- t\right)+n\mu_\lambda t} \right| \\
&\leq e^{n\mu_\lambda t} \sum_{k=1}^{N(t)} \big|h_n\big(Y_k(t)\big)\big|\, e^{\psi_\lambda^- Y_k(t)^2 + \lambda\left(X_k(t)+c_\lambda^- t\right)} \\
&\leq C\ |R_t|^n\ e^{n\mu_\lambda t} Z_\lambda^-(t).
\end{aligned}$$

We now derive a simple bound on the growth of R_t. Since Z_0^- is a convergent martingale, we require

$$\limsup_t \left\{ \psi_0^- R_t^2 - E_0 t \right\} < +\infty$$

hence

$$\limsup_t \frac{R_t^2}{t} \leq \frac{E_0}{\psi_0^-}.$$

Combining the above with 22(a) now gives the required result. ■

9 Proof of the main results

Proof of theorem 1. Let $\lambda \in (\lambda_{\min}, 0)$, then

$$\begin{aligned}
\sum_{k=1}^{N(t)} \mathbf{1}\{X_k(t) \leq -\gamma t\} &\leq \sum_{k=1}^{N(t)} e^{\lambda(X_k(t)+\gamma t)} \\
&\leq \sum_{k=1}^{N(t)} e^{\psi_\lambda^- Y_k(t)^2 + \lambda X_k(t) + \lambda\gamma t} = e^{(E_\lambda + \lambda\gamma)t} Z_\lambda^-(t).
\end{aligned}$$

Taking the infimum over λ,

$$\Delta(\gamma) = \inf_{\lambda\in(\lambda_{\min},0)} \{E_\lambda + \lambda\gamma\},$$

and recalling that the infimum is achieved with parameter $\lambda_\gamma \in (\lambda_{\min}, 0)$ then

$$\sum_{k=1}^{N(t)} \mathbf{1}\{X_k(t) \leq -\gamma t\} \leq e^{\Delta(\gamma)t} Z_{\lambda_\gamma}^-(t).$$

Now, if $\gamma > \tilde{c}(\theta)$, corresponding to $\lambda_\gamma \in \big(\lambda_{\min}, \tilde{\lambda}(\theta)\big)$ and having $\Delta(\gamma) < 0$, we know from theorem 18 that $Z_{\lambda_\gamma}(\infty) = 0$ almost surely. [*Note:* Although not proved here, we should find that $Z^-_{\tilde{\lambda}(\theta)}(\infty) = 0$ a.s. as well.] Then,

$$\gamma > \tilde{c}(\theta) \Rightarrow \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -\gamma t\right\} = 0 \quad \text{eventually, a.s.}$$

Otherwise, if $\gamma \in \big(0, \tilde{c}(\theta)\big)$, corresponding to $\lambda_\gamma \in (\tilde{\lambda}(\theta), 0)$ and having $\Delta(\gamma) > 0$, theorem 18 tells us that $Z^-_{\lambda_\gamma}(\infty) > 0$ almost surely, hence

$$\limsup_{t\to\infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -\gamma t\right\} \le \Delta(\gamma).$$

Now, let $\varepsilon > 0$ be small, $\tilde{\lambda}(\theta) < \lambda < 0$ and $\mu = \lambda - \varepsilon$. We recall now that E_λ is convex so $\frac{\partial^2 E_\lambda}{\partial \lambda^2} \ge 0$ and $\gamma_\mu > \gamma_\lambda$.

$$\begin{aligned}
&\sum_{k=1}^{N(t)} e^{\lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{-(\gamma_\lambda + \varepsilon)t \le X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\} \\
&\qquad \le \sum_{k=1}^{N(t)} e^{\lambda(-(\gamma_\lambda+\varepsilon)t) - E_\lambda t} \mathbf{1}\left\{-(\gamma_\lambda + \varepsilon)t \le X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\} \\
&\qquad = e^{(-\lambda\gamma_\lambda - E_\lambda - \lambda\varepsilon)t} \sum_{k=1}^{N(t)} \mathbf{1}\left\{-(\gamma_\lambda + \varepsilon)t \le X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\} \\
&\qquad \le e^{(-\lambda\gamma_\lambda - E_\lambda - \lambda\varepsilon)t} \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\}
\end{aligned}$$

Then

$$\begin{aligned}
&t^{-1} \log \sum_{k=1}^{N(t)} e^{\lambda X_k(t) - E_\lambda t} \mathbf{1}\left\{|X_k(t)/t + \gamma_\lambda| < \varepsilon\right\} \\
&\qquad \le -\lambda\gamma_\lambda - E_\lambda - \lambda\varepsilon + t^{-1} \log \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\}
\end{aligned}$$

Letting $t \to \infty$, using corollary 21 and remembering that for $\tilde{\lambda}(\theta) < \lambda \le 0$ we have $Z^-_\lambda(\infty) > 0$ a.s., we find

$$0 \le -\lambda\gamma_\lambda - E_\lambda - \lambda\varepsilon + \liminf_{t\to\infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -(\gamma_\lambda - \varepsilon)t\right\}$$

and as $\varepsilon > 0$ can be arbitrarily small we have

$$\liminf_{t\to\infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -\gamma_\lambda t\right\} \ge E_\lambda + \lambda\gamma_\lambda.$$

Equivalently,

$$\liminf_{t\to\infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathbf{1}\left\{X_k(t) \le -\gamma t\right\} \ge E_{\lambda_\gamma} + \lambda_\gamma\gamma = \Delta(\gamma)$$

and hence the limsup and liminf agree as required. ■

We note that these proofs will easily adapt to cover a multi-type branching Brownian motion where the types evolve as a finite state Markov chain, such as found in Champneys et al. (1995), where it will also be possible to prove the analogous convergence theorem required when we have a finite type space by adapting the proof of theorem 20 found in Harris (2000).

In the standard branching Brownian motion case things are even simpler to adapt (where, of course, there is no need for any convergence result akin to theorem 20). All the information necessary is contained in the martingales $\sum \exp\big(\lambda X_k(t) - (\lambda^2/2 + r)t\big)$ studied by Neveu (1988) and, as first came to our attention during discussions with J. Warren, the martingale with parameter λ can *only* be capable of 'counting' particles near $\gamma_\lambda t$ in space at large times t, so when this martingale is uniformly integrable particles *must* perpetually be found with the corresponding speed. Of course, in this case more precise results, in the spirit of Watanabe (1967), also exist.

Proof of theorem 3: the upper bound. Simply observe that for $\lambda \in (\lambda_{\min}, 0)$,

$$
\begin{aligned}
&\sum_{k=1}^{N(t)} \mathbf{1}\left\{Y_k(t)^2 \geq \kappa^2 t; X_k(t) \leq -\gamma t\right\} \\
&\qquad\leq \sum_{k=1}^{N(t)} \mathbf{1}\left\{Y_k(t)^2 \geq \kappa^2 t; X_k(t) \leq -\gamma t\right\} e^{\psi_\lambda^+ (Y_k(t)^2 - \kappa^2 t) + \lambda (X_k(t) + \gamma t)} \\
&\qquad\leq e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t)\, e^{(E_\lambda + \lambda\gamma - \kappa^2 \psi_\lambda^+)t} \\
&\qquad\leq e^{-\lambda(c_\lambda^+ - c_\lambda^*)t} Z_\lambda^+(t)\, e^{(E_\lambda + \lambda\gamma - \kappa^2 \psi_\lambda^+)t}.
\end{aligned}
$$

Recall that

$$
\Delta(\gamma, \kappa) := \inf_{\lambda \in (\tilde{\lambda}(\theta), 0)} \left\{E_\lambda + \lambda\gamma - \kappa^2 \psi_\lambda^+\right\}
$$

where it is easy to check that the infimum is achieved at

$$
\bar{\lambda}(\gamma, \kappa) = -\sqrt{\frac{\gamma^2 \theta(\theta - 8r)}{a^2(\kappa^2 + \theta)^2 + 4a\gamma^2\theta}}.
$$

In cases where $\Delta(\gamma, \kappa) < 0$, we can use the optimal value for λ and theorem 22 to see that

$$
\sum_{k=1}^{N(t)} \mathbf{1}\left\{Y_k(t)^2 \geq \kappa^2 t; X_k(t) \leq -\gamma t\right\} \to 0
$$

almost surely. Hence,

$$
D(\gamma, \kappa) = -\infty \quad a.s. \quad \text{if} \quad \Delta(\gamma, \kappa) < 0.
$$

Otherwise, we have $\Delta(\gamma, \kappa) \geq 0$ (which in fact guarantees that $\gamma \in \big(0, \tilde{c}(\theta)\big]$ and hence $\bar{\lambda}(\gamma, \kappa) \in [\tilde{\lambda}(\theta), 0)$), and since

$$
D(\gamma, \kappa) \leq \limsup_{t \to \infty} \left\{t^{-1} \log e^{-\lambda(c_\lambda^+ - c_\lambda^*)t} Z_\lambda^+(t)\right\} + (E_\lambda + \lambda\gamma - \kappa^2 \psi_\lambda^+)
$$

we can again make use of theorem 22 and the minimising λ value, $\bar{\lambda}(\gamma, \kappa)$, to get the bound

$$
D(\gamma, \kappa) \leq \Delta(\gamma, \kappa) \qquad a.s.
$$

as desired. ■

Proof of theorem 3: the lower bound. Recall that the proof of the lower bound of Theorem 3 requires us to show that

$$
\lim_{t \to \infty} t^{-1} \ln N_t(\gamma, \kappa) \geq D(\gamma, \kappa).
$$

Now we know from Theorem 1 that, for large t, the number of particles near αt in spatial position is almost surely of order $\exp(\{\Delta(\alpha)\}t)$, where $0 < \alpha < \tilde{c}(\theta)$. So with probability 1 we have that there are at least $\exp(\{\Delta(\alpha) - \delta_1\}t)$ particles near αt in space for any $\delta_1 > 0$, for sufficiently large t. Suppose we have a particle with type position greater than 0. By a coupling argument, its type position will always be greater than or equal to a particle at type position exactly equal to 0 at time t. Hence, for sufficiently large t, the number of particles at position $((\alpha+\beta)t, \kappa\sqrt{t})$ at time $t+\tau$ is at least as big as the binomial random variable, X, below:

$$N_{t+\tau}((\alpha+\beta)t, \kappa\sqrt{t}) \geq \mathrm{Bin}\left(e^{(\Delta(\alpha)-\delta_1)t}, \varepsilon e^{-\Theta(\beta,\kappa)t}\right) =: X,$$

because we know from Theorem 14 that

$$P^{\alpha t,0}(A_\tau) \geq \varepsilon e^{-\Theta(\beta,\kappa)t},$$

and here we are fixing the relationship between t and τ to be $\kappa^2 t = (\theta/2\mu_{\bar{\lambda}})e^{2\mu_{\bar{\lambda}}}$ because we have put all particles back to the type origin.

Optimising over $\alpha+\beta=\gamma$, we find that with

$$\bar{\alpha} = \gamma\theta/(\theta+\kappa^2) \quad \text{and} \quad \bar{\beta} = \gamma\kappa^2/(\theta+\kappa^2),$$

we have $\Delta(\alpha) - \Theta(\beta,\kappa) \leq \Delta(\bar{\alpha}) - \Theta(\bar{\beta},\kappa) = \Delta(\gamma,\kappa)$. Chebyshev's inequality gives us that, for $\delta_2 > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\left|X - \varepsilon e^{(\Delta(\gamma,\kappa)-\delta_1)t}\right| > e^{(\Delta(\gamma,\kappa)-\delta_2)t}\right)\\
&\quad\leq e^{-2(\Delta(\gamma,\kappa)-\delta_2)t}\left(\varepsilon e^{(\Delta(\gamma,\kappa)-\delta_1)t}(1-\varepsilon e^{-\Theta(\beta,\kappa)t})\right)\\
&\quad\leq \varepsilon e^{-\Delta(\gamma,\kappa)t} e^{(2\delta_2-\delta_1)t}.
\end{aligned}$$

This means that, if we choose $0 < \delta_1 < \delta_2 = \Delta(\gamma,\kappa)/3$, say, the above probability decays exponentially; then by the Borel-Cantelli Lemma we have, on any mesh of $\mathbb{R}^+$ and for t sufficiently large,

$$\mathbb{P}\left(X > \varepsilon e^{(\Delta(\gamma,\kappa)-\delta_1)t} - e^{(\Delta(\gamma,\kappa)-\delta_2)t}\right) = 1.$$

Since this holds for any $\delta_1 > 0$, we see that

$$\lim_{t\to\infty} t^{-1} \ln N_t(\gamma,\kappa) \geq \Delta(\gamma,\kappa) \tag{46}$$

on any mesh of $\mathbb{R}^+$. Now recall that in the proof of Theorem 14 we actually showed that for sufficiently large τ, A_τ is the event that a single particle makes the 'short climb' *and* stays at the wavefront for a short fixed time. So by taking a sufficiently fine mesh we see that in fact (46) holds on the whole of $\mathbb{R}^+$. ∎

10 An alternative large-deviations approach

In this section we first present some path large deviation conjectures for a fairly broad class of one-dimensional branching diffusions. We then suggest how these can be extended to a class of two-dimensional 'typed' branching diffusions, which will include the main model considered in this paper. The section is concluded with a discussion of how one of these large deviation ideas can be applied to provide a lower bound for the 'short climb probability' required in the proof of the asymptotic growth rate of Theorem 4. The intuitive picture in section 3 motivated these ideas, and while alternative spine techniques have already provided the rigorous proofs required for this papers main results, the authors included the heuristics in section 3 for general interest as they provide good background to the problem.

We refer to the results in this section as conjectures because, although the proofs provided sketch the main ideas that provide heuristic justification for the results, rigour is lacking from the arguments in places. However, more recent work in Hardy and Harris (2004*c*) provides an alternative spine approach where the spine particle is made to stay 'close' to a given trajectory and it is hoped that in future research we will be able to provide the rigour and elegance currently lacking in this sections conjectures.

10.1 Large deviations for a one-dimensional branching diffusion process

Let $T \in \mathbb{N}$ be an indexing parameter of a family of diffusion processes $Z^T_{s\le t}$ which satisfy

$$\mathrm{d}Z^T_s = \frac{1}{\sqrt{T}}\sigma(Z^T, s)\,\mathrm{d}B_s + \mu(Z^T, s)\,\mathrm{d}s.$$

The results below come under the scope of the Freidlin-Wentzell theory. We let σ and μ depend on the entire path of the process. As long as for some increasing function in time, ℓ_t, we have that

$$\sigma(Z_{s\le t}, t) \le \ell_t(1 + \sup_{s\le t} Z_s) \qquad \text{and}$$
$$\mu(Z_{s\le t}, t) \le \ell_t(1 + \sup_{s\le t} Z_s),$$

a strong unique solution to the above SDE exists. Freidlin and Wentzell (1984) and Wentzell (1990) considered bounded σ and μ and showed that under these conditions Z^T admits a large-deviations principle with a rate function

$$I(z,t) = \frac{1}{2}\int_0^t \{\dot{z}_s - \mu(z,s)\}^2/\sigma(z,s)^2\,\mathrm{d}s.$$

We will need σ and μ to be unbounded. For large-deviations results for jumpless diffusion processes with unbounded coefficients, we refer the reader to Stroock (1984). In this paper, we use Liptser and Pukhalskii (1992) as our source for large-deviations limit theorems for semi-martingales, so we state the conditions we require for the next theorem to hold.

Let z be a path in $C[0,t]$ with $I(z,t) < \infty$. Define the cumulant function

$$G(\lambda, z, t) = \lambda\mu(Z,t) + \frac{1}{2}\lambda^2\sigma^2(Z,t),$$

with its Legendre-transform

$$H(y, Z, t) = \sup_\lambda\{\lambda y - G(\lambda, Z, t)\}.$$

We need the following non-degeneracy conditions:

(i) The function $\lambda_y(Z,t)$ exists in some open neighbourhood of z and

$$y = \frac{\partial}{\partial\lambda}G(\lambda_y, Z, t).$$

(ii) For all N there exist δ and r such that if $|y| < N$ and $Z \in B(z,\delta)$ then $|\lambda_y(s)| \le r$ for all $s \le t$.

(iii) If $Z \in B(z,\delta)$ then $H(y, Z, s) = \lambda_y y - G(\lambda_y, Z, s)$ for all $s \le t$.

All these conditions will be *assumed* to hold throughout this section, or can be verified to hold in the application in mind discussed at the end of this section. We observe that when σ and μ are just a function of the current position of Z, then $H(y,Z,t) = (y-\mu)^2/2\sigma^2$ and $\lambda_y = (y-\mu)/\sigma^2$. We then get the following result:

Theorem 25 *Let Z^T be as above and let the rate function for C_1 functions be*

$$I(z,t) = \frac{1}{2}\int_0^t \{\dot{z}_s - \mu(z_s,s)\}^2/\sigma(z_s,s)^2\,\mathrm{d}s.$$

Then for every open set B and closed set C in $C[0,t]$ equipped with the supremum topology,

$$\liminf_{T\to\infty} T^{-1}\log P(Z^T_s \in B) \ge -\inf_{z\in B} I(z,t),$$
$$\limsup_{T\to\infty} T^{-1}\log P(Z^T_s \in C) \le -\inf_{z\in C} I(z,t).$$

In fact, from now on, let us assume that σ and μ are time-independent piece-wise differentiable and increase at most linearly (as this is all we will need later on). Let us include branching in our calculations. We let N_t^T denote the set of particles alive at time t. Each particle $i \in N_s^T$ at position $Z_i^T(s)$ will split into two particles at an exponential rate $T \cdot r(Z_i^T(s), s)$. We define the expectation rate function

$$J(z,t) = I(z,t) - \int_0^t r(z_s, s)\,\mathrm{d}s.$$

Theorem 26 *For every set $B \subset C[0,t]$, let $N_t^T(B)$ be the set of particles whose paths satisfy $Z_i^T(s \le t) \in B$. Then*

$$\liminf_{T\to\infty} T^{-1} \log E(N_t^T(B)) \ge -\inf_{z\in B} J(z,t),$$
$$\limsup_{T\to\infty} T^{-1} \log E(N_t^T(C)) \le -\inf_{z\in C} J(z,t).$$

Proof. We first observe that all particles share the same law. If r is constant, we can separate the diffusion and breeding and get

$$\begin{aligned} E(N_t^T(B)) &= E\{\sum_{i\in N_t^T} \mathbf{1}(Z_i^T(s \le t) \in B)\},\\ &= E\{E(\sum_{i\in N_t^T} \mathbf{1}(Z_i^T(s \le t) \in B)|N_t^T)\},\\ &= E(N_t^T)P_t^T(B),\\ &= e^{rtT}P_t^T(B). \end{aligned}$$

We apply result 25 to $P_t^T(B)$ and we are done. Suppose now that r is a piece-wise constant function (e.g. $r(x) = \lfloor x^2 \rfloor$). We prove the lower bound using a gluing argument, and omit the details of the upper bound proof. Suppose we have a differentiable path z. Our aim is to show that for any open neighbourhood $B(z,c)$ of z we have

$$\liminf_{T\to\infty} T^{-1} \log E(N_t^T(B_c)) \ge -J(z,t).$$

We do not even need to prove this for all $z \in C^1[0,t]$ as any dense subset in $C[0,t]$ will do. Thus, without loss of generality, we may assume that z is a piece-wise linear function with a gradient which is never 0. Divide $[0,t]$ into $0 = t_0 < \ldots < t_{n+1} = t$ such that on each interval $s \in I_j = (t_j, t_{j+1})$ we have $r(z(s)) = r_j$ is constant. For all $c > 0$ there exists a set $U_j \subset I_j$ s.t. if $s \in U_j$ we have $r([z(s)-c, z(s)+c]) = r_j$. As $c \downarrow 0$, we have $U_j \uparrow I_j$ in the Lebesgue-measure sense that $\lambda(I_j \setminus U_j) = O(c)$. The constant is proportional to $\min_{j\le n} \dot{z}(t_j)^{-1}$. We define the tube

$$V_j = \{(x,y) : x \in U_j, |y - z(x)| \le c\}.$$

For a particle starting anywhere on the left of tube V_j, we can ask how many descendants do we expect it to have which will stay inside V_j. Since inside the tube the breeding rate is constant, we can apply result 26 for constant rates. For a descendent which is at the right of tube V_j, we can ask what is the probability that it itself (excluding any breeding) will manage to hit the left of tube V_{j+1} while staying in $B(z,c)$. We can apply result 25 to answer this question. We get that

$$\liminf_{T\to\infty} T^{-1} \log E(N_t^T(B_c)) \ge -J(z,t) - O(c)$$

The correction term is due to the lack of breeding on $\cup_{j\le n} I_j \setminus U_j$. We let $c \downarrow 0$ to complete the argument. We left some details out, but we hope that the reader is now comfortable enough with the gluing concept to be able to fill in the missing details.

Finally, if r is a continuous function, bound it above and below by two piece-wise constant functions r_- and r_+. We construct the three diffusion processes on the same probability space such that $N_- \le N \le N_+$. As we get finer and finer approximations of r so that $r_\pm \to r$, the lower bound follows. ■

Let $t_z = \inf\{t : J(z,t) > 0\}$. Define the rate function $K(z,t)$ as

$$K(z,t) := \begin{cases} J(z,t) & \text{if } t \le t_z, \\ \infty & \text{otherwise.} \end{cases}$$

We will now show that when $K(z,t)$ is negative, it measures the rate of growth of particles following near the path z up to time t.

Theorem 27 *Almost surely,*

$$\liminf_{T\to\infty} T^{-1} \log N_t^T(B) \ge -\inf_{z\in B} K(z,t),$$
$$\limsup_{T\to\infty} T^{-1} \log N_t^T(C) \le -\inf_{z\in C} K(z,t).$$

Assuming σ, μ and r are constant, this is a dyadic branching Brownian motion. Result 27 has been proved in Git (1998*a*). Actually a stronger lower-bound statement was proved, namely that for every $\varepsilon > 0$,

$$P\left(T^{-1}\log N_t^T(B) \le -\inf_{z\in B} K(z,t) - \varepsilon\right) \to 0 \qquad \text{exponentially in } T.$$

Before proving result 27 for general σ, μ and r, we state and prove a related result.

Theorem 28 *Let $L(z,t) = \sup_{s\le t} J(z,s)$.*

$$\liminf_{T\to\infty} T^{-1} \log P(N_t^T(B) \ge 1) \ge -\inf_{z\in B} L(z,t),$$
$$\limsup_{T\to\infty} T^{-1} \log P(N_t^T(C) \ge 1) \le -\inf_{z\in C} L(z,t).$$

We comment that $L(z,t) \ge J(z,0) = 0$ and we will be interested in L when it is a strictly positive function.

Proof of the upper bound of result 28 for general σ, μ and r. We prove the upper bound for a compact set C. If there exists $z \in C$ such that $L(z,t) = 0$ then there is nothing to prove. Hence for all $z \in C$ there exists a time $s_z \in (0,t]$ which satisfies $L(z,t) = J(z,s_z) > 0$. For every $\varepsilon > 0$ there exists $\delta_z > 0$ such that for all $y \in B(z,\varepsilon_z)$ we have $J(y,s_z) > J(z,s_z)(1-\varepsilon)$. These balls form an open cover of C and hence a finite sub-cover exists whose closure we denote as $C_1,\ldots,C_n$. For each of these closed balls we have at time s_z

$$P(N_{s_i}^T(C_i) \ge 1) \le E(N_{s_i}^T(C)).$$

The RHS bound decays exponentially like $\inf_z J(z,s_i) \ge L(z_i,t)(1-\varepsilon)$. On the other hand, if there are no particles inside C_i at time s_i, there are certainly no particles in C_i at time t. Hence in the limit, the probability of finding a particle in $C \subset \cup_i C_i$ decays with the smallest rate $\min_i L(z_i,t)(1-\varepsilon)$. Taking the limit as $\varepsilon \downarrow 0$ yields the result. For general closed sets, the upper bound follows by the lower semi-continuity property of the rate function J. ■

Proof of the lower bound of result 28 for general σ, μ and r. We prove the lower bound for an arbitrary open neighbourhood B of any piece-wise linear function z. Decompose $[0,t]$ into several intervals $0 = t_0 < \ldots < t_n = t$ such that $\dot z$ is constant on each $[t_j, t_{j+1}]$. Looking at $\dot J(z,s) = (\dot z - \mu)^2/2\sigma^2 - r$ which is also constant on $[t_j, t_{j+1}]$ we realise that J attains its supremum at some t_i. We will prove the lower bound using induction on i but first we need two definitions. Let $J(z,t_1,t_2) = J(z,t_2) - J(z,t_1)$. If for all $t \in [t_1,t_2]$ we have $J(z,t_1,t) \le 0$ then we define $K(z,t_1,t_2)$ to be the same as $J(z,t_1,t_2)$. For each i, let j_i be such that

$$\max_{j\le i} J(z,t_j) = J(z,t_{j_i}) = L(z,t_i).$$

Our induction hypothesis is that for any B, an open neighbourhood of z we have

$$\liminf_{T\to\infty} T^{-1} \log P(N_{t_i}^T(B) \ge 1) \ge -L(z,t_i).$$

Further, we state that conditional on $N^T_{t_{j_i}}(B) \geq 1$,

$$P(T^{-1} \log N^T_{t_i}(B) < -J(z, t_{i_j}, t_i) - \varepsilon | N^T_{t_{j_i}}(B) \geq 1) \to 0 \qquad \text{exponentially in } T.$$

When $i = 0$ the induction hypothesis holds trivially. Suppose it holds until time t_i. If $j_i = j_{i+1}$, then $J(z, t_{j_i}, t)$ is non-positive for all $t \leq t_{i+1}$. Conditional on there being a particle in B at time t_{j_i} we can apply result 27 on $[t_{j_i}, t_{i+1}]$ to deduce that the induction hypothesis holds. If $j_{i+1} = i + 1$ we know (by Chauvin and Rouault (1988)) that given there is a particle at t_i in B, the probability of there being a descendent of the particle in B by time t_{i+1} decays with a rate which is at most $J(z, t_i, t_{i+1})$. We take three steps to calculate $P(N^T_{t_{i+1}} > 1)$.

(i) Condition on there being a particle in B at time t_{j_i}, which happens with probability of at least $e^{-T(J(z,t_{j_i})+\varepsilon)}$.

(ii) Restrict ourselves further to the event that there are at least

$$e^{-T(J(z,t_{i_j},t_i)+\varepsilon)}$$

particles in B by time t_i.

(iii) Perform $N^T_{t_i}(B)$ independent experiments with each particle in B at time t_i trying to produce an offspring in B at time t_{i+1}.

Everything works out. The probability of observing a particle in B by time t_{i+1} behaves like $e^{-TJ(z,t_{i+1})}$ (we let $\varepsilon \downarrow 0$ in the three steps). The induction hypothesis holds. ■

Proof of the upper bound of result 27 for general σ, μ and r. Fix n. Let A^T_n be the event that $T^{-1} \log N^T_t(C) > -K(z, t) + 1/n$. From the upper bound of result 26 we must have that $P(A^T_n)$ decays exponentially in T. Using Borel-Cantelli Lemma in T we deduce that if

$$A_n = \left\{ \omega : \limsup_{T\to\infty} T^{-1} \log N^T_t(C) > - \inf_C K(z, t) + 1/n \right\}.$$

then $P(A_n) = 0$. Taking the union over n we deduce that

$$A = \left\{ \omega : \limsup_{T\to\infty} T^{-1} \log N^T_t(C) > - \inf_C K(z, t) \right\},$$

is a set of probability zero, which was what we wanted. ■

Proof of results 27 and 28 for general σ, μ and r.

(i) If the functions σ and μ are piece-wise constants (so for example, $\mu(x) = \lfloor x \rfloor$), apply the results in Git (1998*a*) repeatedly over regions where σ and μ and are constants and then "glue" the pieces together.

(ii) The difficult part of the proof is to extend these results to continuous σ and μ, which we prove later.

(iii) (iii) If r is piece-wise constant while σ and μ are continuous. Glue together regions over which r is constant to get the result.

(iv) Finally, if r is continuous, bound it above and below by two piece-wise constant functions. Run two other diffusion processes on the same probability space with the breeding rates bounding the true r. The number of particles is monotonic in r and as the two approximations converge to r, so does the number of particles found in any region. Since this particle-wise convergence is so powerful we may, and do, apply it to r which varies continuously in both space and time.

The two proofs extending the lower bound of results 27 and 28 from piece-wise constant μ and σ to continuous functions are nearly identical. Hence we demonstrate the most difficult part which is the lower bound extension of result 27. We aim to show that for every z such that $K(z, t) < 0$, and for every ball $B(z, 2c)$ around z we have

$$\liminf_{T\to\infty} N^T_t(B(z, 2c)) \geq -K(z, t).$$

Let $||\hat{\sigma} - \sigma||_\infty \le \varepsilon$ and $||\hat{\mu} - \mu||_\infty \le \varepsilon$ be piece-wise constant functions approximating σ and μ. Since we can still "glue" regions together, WLOG we assume that $\hat{\sigma}$ and $\hat{\mu}$ are constant. We construct an approximate diffusion process on the same probability space Y^T satisfying

$$\mathrm{d}Y_t^T = \frac{1}{\sqrt{T}} \hat{\sigma}(Y_t^T)\,\mathrm{d}B_t + \hat{\mu}(Y_t^T)\,\mathrm{d}t.$$

We ask ourselves, how big is $D_t^T = \sup_{i \in N_t^T} |Y_i^T(t) - Z_i^T(t)|$. We construct both Y^T and Z^T such that both are driven by the same diffusion process. The difference process $X^T = Z^T - Y^T$ will satisfy

$$\mathrm{d}X_t^T = \frac{1}{\sqrt{T}} a(X_t^T, Z_t^T)\,\mathrm{d}B_t + b(X_t^T, Z_t^T)\,\mathrm{d}t.$$

Regardless of what is the path of Z^T (indeed, not only do we not need it to obey a large-deviations principle, it can be even deterministic) the upper bound of result 25 means we get a large deviations function for X^T (possibly conditioned on the path of Z^T). We thus estimate P_ε^T, the probability that the right-most particle of X^T is above level ct.

$$\limsup_{T\to\infty} T^{-1} \log P_\varepsilon^T \le rt - \inf\left\{ \frac{1}{2}\int_0^t \{\dot{x}_s - b(x_s, z_s)\}^2 / a(x_s, z_s)^2\,\mathrm{d}s : x(t) = ct \right\}$$

Since $||a||, ||b|| \le \varepsilon$, for all x and z we have

$$\begin{aligned}
\frac{1}{2}\int_0^t \{\dot{x}_s - b(x_s, z_s)\}^2 / a(x_s, z_s)^2\,\mathrm{d}s &\ge \frac{t}{2\varepsilon^2}\frac{1}{t}\int_0^t (\dot{x} - b(x,z))^2\,\mathrm{d}s \\
&\ge \frac{t}{2\varepsilon^2}\left(\frac{1}{t}\int_0^t \dot{x} - b(x,z)\,\mathrm{d}s\right)^2 \\
&\ge \frac{t}{2\varepsilon^2}(c - \varepsilon)^2.
\end{aligned}$$

Suppose we wanted to find particles in a tube $B(z, 2c)$ of radius $2c$ around the function z. We would pick a small enough ε and good piece-wise constant approximations $\hat{\sigma}$ and $\hat{\mu}$. Depending on z, we then dissect the line $[0,t]$ into intervals $[t_j, t_{j+1}]$ so that $\hat{\sigma}, \hat{\mu} : B(z, 2c)|_{[t_j, t_{j+1}]}$ are constant. We run the approximate process Y^T to find $\hat{N}_t^T(B(z,c))$ particles of the Y^T diffusion inside the narrower tube which satisfy

$$\liminf_{T\to\infty} T^{-1} \log \hat{N}_t^T(B(z,c)) \ge -\hat{K}(z,t).$$

We say that things have gone wrong if at any time t_j we discovered that

$$\sup_{i \in \hat{N}_{t_j}^T} |X_i^T(t_j)| > (c/t)t_j.$$

The joint distribution of Z^T and X^T are not independent. We will want to know that even when we condition on an unlikely event happening to Z^T, the probability of things going wrong decays exponentially. This is true since we can pick an ε small enough so that $(c-\varepsilon)^2/2\varepsilon^2$ is much larger than any fixed decay rate associated with Z^T. By the Borel-Cantelli Lemma, almost surely, things will not go wrong eventually. We have that the particles in $\hat{N}_t^T(B(z,c))$ satisfy $|Z_i^T(t_j) - z(t_j)| < 2c$. We now wish to show that these particle are in $B(z,2c)$ for all t. We can force the dissection of $[0,t]$ to be as fine as we wish. Define $A \subset C[0,t]$ as follows:

$$A := \{y \in C[0,1] \setminus B(z,2c) : |y(t_j) - z(t_j)| < 2c \text{ for all } t_j\}$$

Let $\bar{A}$ be the closure of A. We pick a dissection of $[0,t]$ which is fine enough such that $J(\bar{A}, t) > K(z,t) + 2\delta$ and get using the upper bound that almost surely

$$\limsup T^{-1} \log N_t^T(\bar{A}) \le -K(z,t) - \delta.$$

By letting $\varepsilon \to 0$ and observing that $\hat{K} \to K$, we conclude that once $|\hat{K}(z,t) - K(z,t)| < \delta$, most of $\hat{N}_t^T(B(z,c))$ not only stayed inside $B(z,2c)$ at times t_j but also for all t. We are done. ∎

10.2 Large deviations for higher dimensions

It is possible to extend results 26, 27 and 28 to higher dimensions but for the application we have in hand, we do not need really need to work in complete generality. This is because we can think of the type of the particles as a one-dimensional diffusion process while the position of each particle is simply "marking time" so to speak. We have a branching diffusion process in two dimensions (W^T, Z^T) which satisfies

$$\begin{aligned} \mathrm{d}Z_t^T &= \frac{1}{\sqrt{T}}\sigma(Z_t^T)\,\mathrm{d}B_t^T + \mu(Z_t^T)\,\mathrm{d}t \\ \mathrm{d}W_t^T &= \frac{1}{\sqrt{T}}\hat{\sigma}(Z_t^T, Y_t^T)\,\mathrm{d}\hat{B}_t^T + \hat{\mu}(Z_t^T, Y_t^T)\,\mathrm{d}t \end{aligned}$$

Further, we have that the breeding rate is given by $T \cdot r(Z_t^T)$ and that $\hat{B}_t$ and B_t are independent for all particles. It follows that the branching diffusion family Z^T satisfies the conditions of results 26, 27 and 28 with rate function $J(z,t)$. Conditional on the path of Z^T, the W^T position of each particle follows a diffusion process which satisfies the conditions of result 25 with rate function $\hat{I}(z,w,t)$. We get the following result which is equivalent to result 26.

Theorem 29 *Let B and C be open (closed) sets in $C[0,t]^2$.*

$$\begin{aligned} \liminf_{T\to\infty} T^{-1}\log E(N_t^T(B)) &\geq -\inf_B(J(z,t)+\hat{I}(w,z,t)) \\ \limsup_{T\to\infty} T^{-1}\log E(N_t^T(C)) &\leq -\inf_C(J(z,t)+\hat{I}(w,z,t)) \end{aligned}$$

Proof. We utilise the independence of W given Z and N_t^T, just as we did in the proof of result 26 for constant breeding rates. So for example, to estimate $E(N_t^T(B))$, we write $\pi_z N_t^T(B)$ to denote the number of particles whose projection onto the z coordinates is in $\pi_z B$.

$$\begin{aligned} E(N_t^T(B)) &= E\left[\sum_{i\in\pi_z N_t^T(B)} \mathbf{1}\{(W,Z)_i^T(s\leq t)\in B\}\right] \\ &= E\left[E\left(\sum_{i\in\pi_z N_t^T(B)} \mathbf{1}\{(W,Z)_i^T(s\leq t)\in B\}|N_{s\leq t}^T, Z_i^T\right)\right] \\ &= E[\pi_z N_t^T(B)]\cdot P[(W,Z)_i^T(s\leq t)\in B|i\in\pi_z N_t^T(B)] \end{aligned}$$

The large-deviations behaviour of $E[\pi_z N_t^T(B)]$ is given in result 26. Unfortunately, there may be variations in the paths inside B which Z^T took, which will affect the large-deviations behaviour of $P[(W,Z)_i^T(s\leq t)\in B|i\in\pi_z N_t^T(B)]$. If we wanted to prove the lower bound, it is sufficient to prove it for an arbitrary small open balls of all differentiable z. Then, since μ and σ are continuous functions in Z, as B becomes smaller, we can bound the large deviations behaviour of $P[(W,Z)_i^T(s\leq t)\in B|i\in\pi_z N_t^T(B)]$ from below. Similarly, if we wanted to prove the upper bound, we would take a finite open cover of a compact set C. We omit the details. ■

We define $J(w,z,t) = J(z,t)+\hat{I}(w,z,t)$, $L(w,z,t) = \sup_{s\leq t} J(w,z,s) \geq 0$ and the rate function $K(w,z,t)$ by

$$K(w,z,t) := \begin{cases} J(w,z,t) & \text{if } L(w,z,t)=0, \\ \infty & \text{otherwise.} \end{cases}$$

We state the following result, which is not trivial to prove, but all the machinery to prove it has been introduced already.

Theorem 30 *Almost surely*

$$\begin{aligned} \liminf_{T\to\infty} T^{-1}\log N_t^T(B) &\geq -\inf_B K(w,z,t) \\ \limsup_{T\to\infty} T^{-1}\log N_t^T(C) &\leq -\inf_C K(w,z,t) \end{aligned}$$

Further,

$$\liminf_{T\to\infty} T^{-1}\log P(N_t^T(B)\geq 1) \geq -\inf_B L(w,z,t)$$
$$\limsup_{T\to\infty} T^{-1}\log P(N_t^T(C)\geq 1) \leq -\inf_C L(w,z,t)$$

To prove the upper bounds, we refer the reader to the upper bounds proofs of results 27 and 28.

Proof of the lower bound. We wish to prove the lower bounds for arbitrary open neighbourhood B of (w,z).

(i) Suppose $J_z(z,s)$ is increasing on $[0,t]$. We estimate $P(\pi_z N_t^T(B)\geq 1)$ using result 28 applied to Z^T. We then estimate the probability that this particle is in B using result 25 applied to W^T.

(ii) Suppose $J(z,s)$ is decreasing on $[0,t]$. We bound below $\pi_z N_t^T(B)$ by a breeding process of rate $-\dot{J}(z,t)(1-\varepsilon)$. This allows us to apply results 27 and 28 to W^T, while thinking of $\pi_z N_t^T(B)$ as a time-dependent breeding process.

(iii) Suppose $\dot{J}$ changes signs on $[0,t]$? We just get our hands messy with glue.

■

10.3 Application of the large deviation results

We now sketch how results 29 and 30 can help us make the heuristics presented in section 3 precise. Suppose we start with a particle whose type at time t is greater than 0. By a coupling argument, its type position will always be greater or equal to a particle whose type position at time t is exactly 0. We run the particle from t until $t+\tau$. Define $Z_s^t = Y_{t+s}/\sqrt{t}$. We get

$$Z_s^t = Z_0 + t^{-\frac{1}{2}}B_s - \frac{1}{2}\int_0^s Z_u^t\,\mathrm{d}u.$$

Similarly, looking at the Brownian displacement of each particle since time t, we see that for each particle, for all $s\in[t,t+\tau]$

$$X_{t+s}-X_t = \int_0^s \sqrt{at}Z_u^t\,\mathrm{d}B_u.$$

Setting $W_s^t = \frac{1}{t}(X_{t+s}-X_t)$ we finally get

$$\mathrm{d}Z_s^t = t^{-\frac{1}{2}}\,\mathrm{d}B_s - \frac{1}{2}Z_s^t\,\mathrm{d}s,$$
$$\mathrm{d}W_s^t = t^{-\frac{1}{2}}\sqrt{a}Z_s^t\,\mathrm{d}\hat{B}_s.$$

The instantaneous breeding rate is $\rho + rt\{Z^t\}^2$ so it grows linearly in t. The effective death rate in section 3.3 is always positive for the optimal functions $(\overline{x},\overline{y})$ along the whole path (see section 3.5 and equations (21), (22)). This means that $J(\overline{x},\overline{y},s)$ is an increasing function in s so that $L(\overline{x},\overline{y},t)=J(\overline{x},\overline{y},t)$ and we wish estimate the probability of the particle which started at t being in an open neighbourhood of $(\overline{x},\overline{y})$ by time $t+\tau$ using result 30. However, the spatial diffusion has a small problem, namely, that we have

$$\hat{\sigma}^2(\overline{x}(0),\overline{y}(0)) = a\overline{y}(0)^2 = 0.$$

This means that the lower bound of result 25 does not hold initially for the spatial diffusion. Let $\varepsilon>0$. The type-diffusion satisfies all the conditions of result 27 and 28 on $[0,\varepsilon]$. Let $P_{[0,\varepsilon]}^t$ be the probability of finding a particle near $\overline{y}(\varepsilon)$. Using result 28 we know that

$$\lim_{t\to\infty} t^{-1}\log P_{[0,\varepsilon]}^t \geq -J(\overline{y},\varepsilon).$$

Using the symmetry of the spatial diffusion, the probability that the spatial position of a particle changes by a positive amount during $[0,\varepsilon]$ is exactly $\frac{1}{2}$. Let $P_{[\varepsilon,\tau]}^t$ be the probability for a particle which starts at $(\overline{x}(\varepsilon),\overline{y}(\varepsilon))$ to be near $(\overline{x},\overline{y})$ on $[\varepsilon,\tau]$. We now apply the lower bound of result 30 to get that

$$\lim_{t\to\infty} t^{-1}\log P_{[\varepsilon,\tau]}^t \geq -J(\overline{x},\overline{y},\varepsilon,\tau).$$

We observe that $J(\overline{x},\overline{y},\varepsilon,\tau)+J(\overline{y},\varepsilon)\leq J(\overline{x},\overline{y},\varepsilon,\tau)+J(\overline{x},\overline{y},\varepsilon)=J(\overline{x},\overline{y},\tau)$. Lets imagine a particle:

(i) Its type is near $\overline{y}$ on $[0, \varepsilon]$.

(ii) Its position at time ε is greater than its position at time 0.

(iii) Its type is near $\overline{y}$ on $[\varepsilon, \tau]$.

(iv) Its position is at least $\overline{x}(\tau) - \overline{x}(\varepsilon)$.

The probability of finding such an ideal particle decays by at most $J(\overline{x}, \overline{y}, \tau)$. Since $\lim_{\varepsilon \downarrow 0} \overline{x}(\varepsilon) = 0$, we are done.

Note: The reader may wonder why we do not apply results 29 and 30 for the whole path without resorting to the martingale techniques. The answer is that it *is* possible, but requires even more machinery. This is because what we are looking for are paths with non-typical $\int_0^t Y_s^2 \, \mathrm{d}s$ rather than non-typical Y_t position. This is equivalent to analysing the large deviations of the branching integral-diffusion process $\int Y^2$. We refer the reader to the last chapter of Git (1998*b*) for such a large-deviations analysis. The martingale methods we use provide a neat and precise alternative to a more direct large deviation analysis although, of course, it is exactly the same 'deviant' particles that the martingale limits end up 'picking out'.

Acknowledgement 31 *This work was conducted whilst J.W.Harris was in receipt of an EPSRC research studentship.*

References

Athreya, K. B. (2000), 'Change of measures for Markov chains and the $L \log L$ theorem for branching processes', *Bernoulli* **6**(2), 323–338.

Biggins, J. D. (1977), 'Martingale convergence in the branching random walk', *J. Appl. Probability* **14**(1), 25–37.

Biggins, J. D. (1979), 'Growth rates in the branching random walk', *Z. Wahrsch. Verw. Gebiete* **48**(1), 17–34.

Biggins, J. D. (1992), 'Uniform convergence of martingales in the branching random walk', *Ann. Probab.* **20**(1), 137–151.

Champneys, A., Harris, S. C., Toland, J., Warren, J. and Williams, D. (1995), 'Algebra, analysis and probability for a coupled system of reaction-diffusion equations', *Phil. Trans. R. Soc. Lond.* **350**, 69–112.

Chauvin, B. and Rouault, A. (1988), 'KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees', *Prob. Th. Rel. Fields* **80**, 299–314.

Engländer, J. and Kyprianou, A. E. (2002), Local extinction versus local exponential growth for spatial branching processes. To appear in *Ann. Prob.*

Freidlin, M. I. and Wentzell, A. D. (1984), *Random perturbations of dynamical systems*, Vol. 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, New York. Translated from the Russian by Joseph Szücs.

Git, Y. (1998*a*), Almost sure path properties of branching diffusion processes, *in* 'Séminaire de Probabilités, XXXII', Vol. 1686 of *Lecture Notes in Math.*, Springer, Berlin, pp. 108–127.

Git, Y. (1998*b*), Some aspects of large deviations theory, PhD thesis, University of Bath.

Hardy, R. (2004), Branching diffusions, PhD thesis, University of Bath.

Hardy, R. and Harris, S. C. (2004*a*), 'A new formulation of the spine approach to branching diffusions', (0404). Mathematics Preprint, University of Bath.

Hardy, R. and Harris, S. C. (2004*b*), 'A spine proof of a large-deviations principle for branching Brownian motion', (0407). Mathematics Preprint, University of Bath.

Hardy, R. and Harris, S. C. (2004*c*), 'A spine proof of a lower-bound for a typed branching diffusion', (0408). Mathematics Preprint, University of Bath.

Hardy, R. and Harris, S. C. (2004*d*), 'Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales', (0405). Mathematics Preprint, University of Bath.

Harris, S. C. (2000), Convergence of a 'Gibbs-Boltzmann' random measure for a typed branching diffusion, *in* 'Séminaire de Probabilités, XXXIV', Vol. 1729 of *Lecture Notes in Math.*, Springer, Berlin, pp. 239–256.

Harris, S. C. and Williams, D. (1996), 'Large deviations and martingales for a typed branching diffusion. I', *Astérisque* (236), 133–154. Hommage à P. A. Meyer et J. Neveu.

Itô, K. and McKean, H. P. (1965), *Diffusion processes and their sample paths*, Die Grundlehren der Mathematischen Wissenschaften, Band 125, Academic Press Inc., Publishers, New York.

Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y. (1997), A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes, *in* K. B. Athreya and P. Jagers, eds, 'Classical and modern branching processes (Minneapolis, MN, 1994)', Vol. 84 of *IMA Vol. Math. Appl.*, Springer, New York, pp. 181–185.

Kyprianou, A. E. (2003), Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris' probabilistic analysis. To appear in *Ann. Inst. H. Poincaré.*

Liptser, R. S. and Pukhalskii, A. A. (1992), 'Limit theorems on large deviations for semimartingales', *Stochastics Stochastics Rep.* **38**(4), 201–249.

Lyons, R. (1997), A simple path to Biggins' martingale convergence for branching random walk, *in* K. B. Athreya and P. Jagers, eds, 'Classical and modern branching processes (Minneapolis, MN, 1994)', Vol. 84 of *IMA Vol. Math. Appl.*, Springer, New York, pp. 217–221.

Lyons, R., Pemantle, R. and Peres, Y. (1995), 'Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes', *Ann. Probab.* **23**(3), 1125–1138.

McKean, H. P. (1975), 'Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskounov', *Comm. Pure Appl. Math.* **28**, 323–331.

Neveu, J. (1988), Multiplicative martingales for spatial branching processes, *in* 'Seminar on Stochastic Processes, 1987 (Princeton, NJ, 1987)', Vol. 15 of *Progr. Probab. Statist.*, Birkhäuser Boston, Boston, MA, pp. 223–242.

Rogers, L. C. G. and Williams, D. (2000*a*), *Diffusions, Markov processes, and martingales. Vol. 1: Foundations*, Cambridge Mathematical Library, Cambridge University Press, Cambridge. Reprint of the second (1994) edition.

Rogers, L. C. G. and Williams, D. (2000*b*), *Diffusions, Markov processes, and martingales. Vol. 2: Itô calculus*, Cambridge Mathematical Library, Cambridge University Press, Cambridge. Reprint of the second (1994) edition.

Stroock, D. W. (1984), *An introduction to the theory of large deviations*, Universitext, Springer-Verlag, New York.

Varadhan, S. R. S. (1984), *Large deviations and applications*, Vol. 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Watanabe, S. (1967), Limit theorems for a class of branching processes, *in* J. Chover, ed., 'Markov processes and potential theory', Vol. 1967 of *Proceedings of a Symposium conducted by the Mathematics Research Center and the United States Army at the University of Wisconsin, Madison, Wis., May 1–3*, John Wiley & Sons Inc., New York, pp. 205–232.

Wentzell, A. D. (1990), *Limit theorems on large deviations for Markov stochastic processes*, Vol. 38 of *Mathematics and its Applications (Soviet Series)*, Kluwer Academic Publishers Group, Dordrecht. Translated from the Russian.