Version: Thursday, April 20, 2000 at 10:49

Large deviations and martingales for a typed branching diffusion: II

by

Yoav Git
Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, CB2 1SB, United Kingdom.

and

Simon C. Harris
Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, United Kingdom.

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 [**8**]. A combination of both martingale and large deviation results are used to prove this result where, in particular, we also develop some almost sure path large deviation results which hold for a broad class of one-dimensional branching diffusions.

1. Introduction

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. At various places, we will make use of the work found in Harris and Williams [**8**], Harris [**9**] and Git [**6**],[**7**].

(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[**8**], 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,

$Y_k(t)$ in $\mathbb{R}$ is the 'type' of the k^{th}-born particle,

$\big(N(t); X_1(t), \ldots, X_{N(t)}; Y_1(t), \ldots, Y_{N(t)}\big)$ is the current state of the particle system.

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

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

The *breeding* of a type y particle occurs at a rate

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

and we have one child born at these times (binary splitting). A child inherits its parent's current type and spatial position then moves off *independently* of all others. Particles live forever (once born!).

The model has a very different behaviour for low temperature parameter values 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 $\mathbb{P}^{x,y}$ and $\mathbb{E}^{x,y}$ 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 the paper by Champneys *et al.*[**4**] 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. Amongst many others, we are particularly indebted to an initial paper by McKean[**12**], work of Biggins[**1**,**2**,**3**] and the martingales techniques in Neveu[**14**] concerning standard branching Brownian motions for inspiration in this work. 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[**10**]). The Orstein-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 results for branching diffusions that are presented in section 4.

(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

$$N_t(\gamma) := \sum_{k=1}^{N(t)} \mathrm{I}\{X_k(t) \leq -\gamma t\} \tag{1.3}$$

The limit giving the *expected* rate of growth,

$$\lim_{t\to\infty} t^{-1} \log E(N_t(\gamma)) \tag{1.4}$$

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

$$\Delta(\gamma) = \rho + \frac{\theta}{4} - \frac{1}{4}\sqrt{\theta(\theta - 8r)\,(1 + 4\gamma^2/(\theta a))}. \tag{1.5}$$

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[**8**].

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

$$\begin{aligned} \tilde{c}(\theta) &: = \sup\{\gamma : \Delta(\gamma) > 0\} \\ &= \sqrt{2a\left(r + \rho + \frac{2(2r+\rho)^2}{\theta - 8r}\right)} \end{aligned} \tag{1.6}$$

In this particular situation, this guess that 'expectation' and 'particle' wave-speeds agree was actually *proved* rigorously a martingale change of measure technique in Harris and Williams[**8**]. In this paper, we extend this connection and prove that the 'expected' and 'almost sure' rates of growth of particles actually agree.

(1.7) **Theorem.** *Under each $\mathbb{P}^{x,y}$ 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 [**9**].

We gain the asymptotic speed of the spatially most extreme particle as a corollary to this theorem, a result first proved in Harris and Williams[**8**] by using a martingale change of measure technique:

(1.8) **Corollary.**

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

(1.9) **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

$$N_t(\gamma, \kappa) := \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -\gamma t, Y_k^2(t) \geq \kappa^2 t\right\}. \tag{1.10}$$

It can be shown that the limit

$$\lim_{t\to\infty} t^{-1} \log \mathbb{E}\left(N_t(\gamma, \kappa)\right) \tag{1.11}$$

exists and takes the value

$$\Delta(\gamma, \kappa) = \rho + \frac{(\theta - \kappa^2)}{4} - \frac{1}{4\theta a}\sqrt{\theta(\theta - 8r)(4a\theta\gamma^2 + a^2(\theta + \kappa^2)^2)} \tag{1.12}$$

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.

(1.13) **Theorem.** *Under each $\mathbb{P}^{x,y}$ law, the limit*

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

exists almost surely and is given by

$$D(\gamma, \kappa) = \begin{cases} \Delta(\gamma, \kappa) & \textit{if } \Delta(\gamma, \kappa) > 0, \\ -\infty & \textit{otherwise.} \end{cases} \tag{1.14}$$

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 some large deviation results concerning the probability of any particles in the branching diffusion managing to follow various steeply ascending paths. These large deviation results (see sections 3 and 4) give a very clear insight into typical particle paths that can be observed within the branching diffusion, revealing either the growth rate in the number of particles following close to given a path or the probability that any particle manages to follow the path over a large time period.

After proving theorem 1.13, it becomes straightforward to retrieve the following theorem:

(1.15) **Theorem.** *For any $F \subset \mathbb{R}^2$, define*

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

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 $\mathbb{P}^{x,y}$

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

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

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

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

Although we consider theorem to be the main result of the paper, we emphasize again that the path large deviation results presented in section 4 are applicable in more general branching diffusion setups.

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.

(1.16) **Corollary.** *Let $B \subset \mathbb{R}^2$ be any open set. Under each $\mathbb{P}^{x,y}$ law,*

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

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. In section 3, we will give some large deviation heuristics for the particles within the branching diffusion which give a very good insight into certain particles' behaviour. These large deviation results are made precise in section 4 after we develop some almost sure path results which hold for a broad class of one-dimensional branching diffusions. Martingale results and a discussion of some of their uses are given in section 5, with proofs in section 6. Finally, the proofs of the results stated in this section are found in section 7.

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[**8**] for further details of these techniques, or see Rogers and Williams[**15**],[**16**] for the more general theory.

First, we give some key definitions. Let

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}. \tag{2.1}$$

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

$$\lambda_{\min} < \lambda < 0. \tag{2.2}$$

Also, define

$$\mu_\lambda := \frac{1}{2}\sqrt{\theta(\theta - 8r - 4a\lambda^2)}, \qquad \psi_\lambda^{\pm} := \frac{1}{4} \pm \frac{\mu_\lambda}{2\theta} \qquad E_\lambda := \rho + \theta\psi_\lambda^- \tag{2.3}$$

(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

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

where $\xi_t = B\left(\int_0^t A(\eta_s)\,ds\right)$ 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

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

and then

$$\begin{aligned}
\mathbb{E}\sum_{k=1}^{N(t)} f\big(X_k(t), Y_k(t)\big) &= \mathbb{E}_{\frac{\theta}{2},\lambda}\left\{\exp\big(-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t\big) f(\xi_t, \eta_t).M_t^{\mu_\lambda,\frac{\theta}{2}}\right\} \\
&= \mathbb{E}_{\mu_\lambda,\lambda}\left\{\exp\big(-\lambda\xi_t - \psi_\lambda^-\eta_t^2 + E_\lambda t\big) f(\xi_t, \eta_t)\right\}
\end{aligned} \tag{2.4}$$

(2.5) **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 (2.4), we find

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

We can now define

$$\Delta(\gamma) := \inf_{\lambda\in(\lambda_{\min},0)}\{E(\lambda) + \lambda\gamma\} \tag{2.6}$$

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

$$\lambda_\gamma = -\sqrt{\frac{(\theta - 8r)\gamma^2}{\theta a^2 + 4a\gamma^2}} \tag{2.7}$$

and then

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

Of course, choosing this minimizing λ_γ value also means that we have maximized 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, $\xi_t = B\left(\int_0^t A(\eta_s)\,ds\right)$ where B is a BM with drift λ. Almost surely under $\mathbb{P}_{\mu_\lambda,\lambda}$,

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

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

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

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

$$\lim_{t\to\infty} t^{-1}\log E(N_t(\gamma)) = \Delta(\gamma). \tag{2.8}$$

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[**8**] 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).

(2.9) **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 (2.4),

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

Now, continuing with our approximate calculations,

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

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

$$\mathbb{E}\sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t)\sim -\gamma t\,;Y_k(t)^2\geq\kappa^2 t\right\}\sim e^{\left(E_\lambda+\lambda\gamma-\kappa^2\psi_\lambda^+\right)t}\mathbb{P}_{\mu_\lambda,\lambda}\left(\xi_t\sim -\gamma t\,\middle|\eta_t^2\geq\kappa^2 t\right). \tag{2.10}$$

Define

$$\begin{aligned}\Delta(\gamma,\kappa) &:= \inf_{\lambda\in(\lambda_{\min},0)}\left\{E_\lambda+\lambda\gamma-\kappa^2\psi_\lambda^+\right\}\\ &= \rho+(\theta-\kappa^2)/4-\sqrt{\theta(\theta-8r)(4a\theta\gamma^2+a^2(\theta+\kappa^2)^2)}/4\theta a\end{aligned} \tag{2.11}$$

where the infimum is attained with a λ value of

$$\overline{\lambda}(\gamma,\kappa) = -\sqrt{\frac{\gamma^2\theta(\theta-8r)}{a^2(\kappa^2+\theta)^2+4a\gamma^2\theta}}\in(\lambda_{\min},0) \tag{2.12}$$

Then equation (2.10) certainly leads to the inequality

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

It is clear from (2.10) that when minimizing $E_\lambda+\lambda\gamma-\kappa^2\psi_\lambda^+$, we simultaneously maximize 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

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

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

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

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 $\mathbb{P}_{\lambda,\mu_\lambda}$,

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

(ii) *Rapid ascent.* Over a large fixed time τ, but where $t >> \tau$, 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

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

under the $\mathbb{P}_{\lambda,\mu_\lambda}$ law. See, for example, Varadhan[**18**]. Considering the path

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

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

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

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

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

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

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left(\sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -\gamma t; Y_k(t)^2 \geq \kappa^2 t\right\}\right) = \Delta(\gamma,\kappa). \tag{2.13}$$

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 later be used as a basis for our proof of a lower bound for the almost sure growth rate. Martingales will be used to prove the same rate also gives an almost sure upper bound.

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 $Y_k^2(t) > \epsilon t$ for a prolonged period of time. The majority of particles that contribute to $N_\cdot(\gamma,\kappa)$ at time t will fail to contribute to $N_\cdot(\gamma,\kappa)$ 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 traveled 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

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

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 [**8**], 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 $\exp\left(\Delta(\alpha)t\right)$ 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 $s \le t$, 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 descendant 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

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

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

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

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 traveled '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 $\rho_t := \int_0^t (\mu_s - \lambda_s)\,ds$, so here

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

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 $\mathbb{E}(M(\tau)|M(\tau)\geq 1)=W_\tau$.

In our particular case, we have

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

Define the largest effective death rate prior to time τ by

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

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

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

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

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

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 $\exp\left(-L(x,y,\tau)\right)$.

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

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

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

$$\inf_{x,y} L(x,y,\tau) = \inf_{x,y} \sup_{s\in[0,\tau]} J(x,y,s) \geq \inf_{x,y} J(x,y,\tau) \tag{3.5}$$

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

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

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

where

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

yielding

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

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

$$a\int_0^\tau y(s)^2\,ds$$

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

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

Then we are left to find

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

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

$$y_\lambda(s) = \kappa\sqrt{t}\left\{\frac{\sinh\mu_\lambda s}{\sinh\mu_\lambda\tau}\right\} \qquad (0 \leq s \leq \tau) \tag{3.6}$$

where

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

and then

$$\sup_\lambda \inf_y \left\{ \int_0^\tau \left(\frac{1}{2\theta}\Big(\dot y(s) + \frac{\theta}{2} y(s)\Big)^2 - r y(s)^2 - \frac{1}{2} a\lambda^2 y(s)^2 \right) ds + \lambda\beta t \right\}$$
$$= \sup_\lambda \left\{ \lambda\beta t + \kappa^2 t \left(\frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \right) \right\}$$

The optimal choice $\hat\lambda$ satisfies

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

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

$$x_\lambda(s) := \lambda a \int_0^s y_\lambda(u)^2\, du = \beta t \left\{ \frac{\sinh 2\mu_\lambda s - 2\mu_\lambda s}{\sinh 2\mu_\lambda \tau - 2\mu_\lambda \tau} \right\} \tag{3.7}$$

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

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

Finally, it is easy to check that $J(\hat x, \hat y, \tau) = L(\hat x, \hat y, \tau)$, whence we have $\inf_{x,y} J(x,y,\tau) \geq \inf_{x,y} L(x,y,\tau)$ and, combining with (3.5), we have found that

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

(3.8) *A note on the optimal paths:* As $\tau \to \infty$, we have

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

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

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

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

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

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.

(3.9) **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

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

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.

Optimizing 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 (2.7) and (2.12).]

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

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

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.

In the next section, we put these large deviation heuristics on a rigorous footing.

4. Large Deviation Results

In this section we first present some path large deviation results for a fairly broad class of one-dimensional branching diffusions. These results are then 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 results can be applied to estimate a probability required in the proof of the asymptotic growth rate of theorem 1.15.

(4.1) **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

$$dZ^T_s = \frac{1}{\sqrt{T}}\sigma(Z^T, s)dB_s + \mu(Z^T, s)ds.$$

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,

$$\sigma(Z_{s\le t}, t) \le l_t(1 + \sup_{s\le t} Z_s),$$
$$\mu(Z_{s\le t}, t) \le l_t(1 + \sup_{s\le t} Z_s),$$

a strong unique solution to the above SDE exists. Freidlin and Wentzell [**20**],[**21**] considered bounded σ and μ and showed that under these conditions Z^T admits a large-deviations principle with a rate function

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

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 Strook [**17**]. In this paper, we use Liptzer [**11**] 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:

(4.2) **Theorem.** *Let Z^T be as above and let the rate function for C_1 functions be*

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

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

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

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

$$J(z,t) = I(z,t) - \int_0^t r(z_s,s)ds.$$

(4.3) **Theorem.** *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*

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

(4.4) *Proof.*

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

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

We apply result 4.2 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

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

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 Lesbegue-measure sense that $\lambda(I_j \setminus U_j) = O(c)$. The constant is proportional to $\min_{j\leq n} \dot{z}(t_j)^{-1}$. We define the tube

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

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 4.3 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 4.2 to answer this question. We get that the

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

The correction term is due to the lack of breeding on $\cup_{j\leq n} I_j \setminus U_j$. 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.

$|z(s) - z(t_z)| \leq \epsilon/4$, and such that result 4.3 holds by time t_- (if $t_z \neq 0$ pick $t_- < t_z$ and pick $t_- = t_z = 0$ otherwise). 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 $N_- \leq N \leq N_+$. As we get finer and finer approximations of r so that $r_\pm \to r$, the lower bound follows. $\square$

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

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

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.

(4.5) **Theorem.** *Almost surely,*

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

Assuming σ, μ and r are constant, this is a dyadic branching Brownian motion. Result 4.5 has been proved in Git [**7**]. Actually a stronger lower-bound statement was proved, namely that for every $\epsilon > 0$,

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

Before proving result 4.5 for general σ, μ and r, we state and prove a related result.

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

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

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

(4.7) *Proof of the upper bound of result 4.6 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 $\epsilon > 0$ there exists $\delta_z > 0$ such that for all $y \in B(z,\epsilon_z)$ we have $J(y,s_z) > J(z,s_z)(1-\epsilon)$. 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

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

The RHS bound decays exponentially like $\inf_z J(z,s_i) \geq L(z_i,t)(1-\epsilon)$. 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 $\min_i L(z_i,t)(1-\epsilon)$. Taking the limit as $\epsilon \downarrow 0$ yields the result. For general closed sets, the upper bound follows by the lower semi-continuity property of the rate function J. □

(4.8) *Proof of the lower bound of result 4.6 for constant 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 $J(z,t_1,t) \leq 0$ 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

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

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

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

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

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

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 4.5 on $[t_{j_i},t_{i+1}]$ to deduce that the induction hypothesis holds. If $j_{i+1}=i+1$ we know (by Chauvin et al. [**5**]) 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 $P(N_{t_{i+1}}^T > 1)$.

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

(ii) Restrict ourselves further to the event that there are at least $e^{-T(J(z,t_{i_j},t_i)+\epsilon)}$ particles in B by time t_i.

(iii) Perform $N_{t_i}^T(B)$ 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 $\epsilon\downarrow 0$ in the three steps). The induction hypothesis holds. □

(4.9) *Proof of the upper bound of result 4.5 for general σ,μ and r.*

Fix n. Let A_n^T be the event that $T^{-1}\log N_t^T(C) > -K(z,t)+1/n$. From the upper bound of result 4.3 we must have that $P(A_n^T)$ 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. □

(4.10) *Proof of results 4.5 and 4.6 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 [**7**] 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) 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 4.5 and 4.6 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 4.5. 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 $||\hat{\sigma} - \sigma||_\infty \leq \epsilon$ and $||\hat{\mu} - \mu||_\infty \leq \epsilon$ 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

$$dY_t^T = \frac{1}{\sqrt{T}}\hat{\sigma}(Y_t^T)dB_t + \hat{\mu}(Y_t^T)dt.$$

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

$$dX_t^T = \frac{1}{\sqrt{T}}a(X_t^T, Z_t^T)dB_t + b(X_t^T, Z_t^T)dt.$$

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 4.2 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.

$$\limsup_{T\to\infty} T^{-1}\log P_\epsilon^T \leq rt - \inf\left\{ \tfrac{1}{2}\int_0^t \{\dot{x}_s - b(x_s,z_s)\}^2/a(x_s,z_s)^2 ds : x(t) = ct \right\}$$

Since $||a||, ||b|| \leq \epsilon$, for all x and z we have

$$\begin{aligned}
\frac{1}{2}\int_0^t \{\dot{x}_s - b(x_s,z_s)\}^2/a(x_s,z_s)^2 ds &\geq \frac{t}{2\epsilon^2}\frac{1}{t}\int_0^t (\dot{x} - b(x,z))^2 \\
&\geq \frac{t}{2\epsilon^2}\left(\frac{1}{t}\int_0^t \dot{x} - b(x,z)ds\right)^2 \\
&\geq \frac{t}{2\epsilon^2}(c-\epsilon)^2.
\end{aligned}$$

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 a 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

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

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 $(c - \epsilon)^2/2\epsilon^2$ 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

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

By letting $\epsilon \to 0$ 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. □

(4.11) **Large deviations for higher dimensions.**

It is possible to extend results 4.3, 4.5 and 4.6 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

$$\begin{aligned} dZ_t^T &= \frac{1}{\sqrt{T}} \sigma(Z_t^T) dB_t^T + \mu(Z_t^T) dt \\ dW_t^T &= \frac{1}{\sqrt{T}} \hat{\sigma}(Z_t^T, Y_t^T) d\hat{B}_t^T + \hat{\mu}(Z_t^T, Y_t^T) dt \end{aligned}$$

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 4.3, 4.5 and 4.6 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 4.2 with rate function $\hat{I}(z,w,t)$. We get the following result which is equivalent to result 4.3.

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

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

(4.13) *Proof.*

We utilise the independence of W given Z and N_t^T, just as we did in the proof of result 4.3 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 $\pi_z B$.

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

The large-deviations behaviour of $E[\pi_z N_t^T(B)]$ is given in result 4.3. 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.

(4.14) **Theorem.** *Almost surely*

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

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 4.5 and 4.6.

(4.15) *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 4.6 applied to Z^T. We then estimate the probability that this particle is in B using result 4.2 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 $-\dot{J}(z,t)(1-\epsilon)$. This allows us to apply results 4.5 and 4.6 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. □

(4.16) **Application of the large deviation results.**

We now sketch how results 4.12 and 4.14 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

$$Z_s^t = Z_0 + t^{-\frac{1}{2}}B_s - \tfrac{1}{2}\int_0^s Z_u^t du.$$

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

$$X_{t+s} - X_t = \int_0^s \sqrt{at}Z_u^t dB_u.$$

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

$$dZ_s^t = t^{-\frac{1}{2}}dB_s - \tfrac{1}{2}Z_s^t ds,$$
$$dW_s^t = t^{-\frac{1}{2}}\sqrt{a}Z_s^t d\hat{B}_s.$$

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 3.8 and equations (3.6), (3.7)). 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 4.14. 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 4.2 does not hold initially for the spatial diffusion. Let $\epsilon > 0$. The type-diffusion satisfies all the conditions of result 4.5 and 4.6 on $[0, \epsilon]$. Let $P^t_{[0,\epsilon]}$ be the probability of finding a particle near $\overline{y}(\epsilon)$. Using result 4.6 we know that

$$\lim_{t\to\infty} t^{-1} \log P^t_{[0,\epsilon]} \geq -J(\overline{y}, \epsilon)$$

Using the symmetry of the spatial diffusion, the probability that the spatial position of a particle changes by a positive amount during $[0, \epsilon]$ is exactly $\frac{1}{2}$. Let $P^t_{[\epsilon,\tau]}$ be the probability for a particle which starts at $(\overline{x}(\epsilon), \overline{y}(\epsilon))$ to be near $(\overline{x}, \overline{y})$ on $[\epsilon, \tau]$. We now apply the lower bound of result 4.14 to get that

$$\lim_{t\to\infty} t^{-1} \log P^t_{[\epsilon,\tau]} \geq -J(\overline{x}, \overline{y}, \epsilon, \tau)$$

We observe that $J(\overline{x}, \overline{y}, \epsilon, \tau) + J(\overline{y}, \epsilon) \leq J(\overline{x}, \overline{y}, \epsilon, \tau) + J(\overline{x}, \overline{y}, \epsilon) = J(\overline{x}, \overline{y}, \tau)$. Lets imagine a particle:
(i) Its type is near $\overline{y}$ on $[0, \epsilon]$.
(ii) Its position at time ϵ is greater than its position at time 0.
(iii) Its type is near $\overline{y}$ on $[\epsilon, \tau]$.
(iv) Its position is at least $\overline{x}(\tau) - \overline{x}(\epsilon)$.
The probability of finding such an ideal particle decays by at most $J(\overline{x}, \overline{y}, \tau)$. Since $\lim_{\epsilon\downarrow 0} \overline{x}(\epsilon) = 0$, we are done.

Note: The reader may wonder why we do not apply results 4.12 and 4.14 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 $\int_0^t Y_s^2 ds$ 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 [**6**] 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'.

5. 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[**8**] that E_λ (also written $E(\lambda)$) and $\Delta(\gamma)$ are Legendre conjugates with

$$\Delta(\gamma) = \inf_{\lambda<0}\{E(\lambda)+\lambda\gamma\}, \quad E(\lambda) = \sup_{\gamma>0}\{\Delta(\gamma)-\gamma\lambda\}, \tag{5.1}$$

If, for $\lambda_{\min} < \lambda < 0$, we write γ_λ for the γ value which achieves the supremum on the right-hand side of (5.1), then the functions ı $\lambda \mapsto \gamma_\lambda$ from $(-\lambda_{\min}, 0)$ to $(0,\infty)$, and ı $\gamma \mapsto \lambda_\gamma$ 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 (5.1). In addition, we note that

$$\gamma_\lambda = -E'(\lambda) = \sqrt{\frac{\theta a^2\lambda^2}{\theta - 8r - 4a\lambda^2}}, \tag{5.2}$$

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

$$\begin{aligned} \tilde{c}(\theta) &= \sup\{\gamma : \Delta(\gamma) > 0\} = \inf\{-E(\lambda)/\lambda : \lambda_{\min} < \lambda < 0\} \\ &= \inf\{c_\lambda^- : \lambda_{\min} < \lambda < 0\} = c^-_{\tilde{\lambda}(\theta)}, \end{aligned} \tag{5.3}$$

where

$$c_\lambda^- := -E_\lambda/\lambda, \quad \text{and} \quad \tilde{\lambda}(\theta) := -\sqrt{\frac{2(\theta-8r)(\theta\rho+2\rho^2+r\theta)}{a(\theta+4\rho)^2}} \in (\lambda_{\min}, 0). \tag{5.4}$$

and a formula for $\tilde{c}(\theta)$ is given in equation (1.6).

(5.5)**The 'ground-state' martingales.** In Harris and Williams[**8**], we proved the almost sure speed of the spatially left-most particle by making use of the following martingales:

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

$$Z_\lambda^-(t) := \sum_{k=1}^{N(t)} \exp\left(\psi_\lambda^- Y_k(t)^2 + \lambda\left[X_k(t) + c_\lambda^- t\right]\right). \tag{5.7}$$

This defines a martingale Z_λ^- (under each $\mathbb{P}^{x,y}$ 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[**8**], we used this simple geometric fact and an idea of Neveu[**14**] in proving the following:

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

(5.9) **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

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

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

See Harris [**9**] if you need convincing of the following martingale properties:

(5.10) **Lemma.** *Let* $\lambda \in \big(\lambda_{\min}, 0\big)$. *For each* $n \in \{0, 1, \ldots\}$ *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* $\mathbb{P}^{x,y}$ *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[**9**] that facilitated the proof of the following convergence theorem by making use of Hermite martingale expansions.

(5.11) **Theorem.** *Let* $\lambda \in \big(\tilde\lambda(\theta), 0\big)$ *and* $\alpha < 1/4$. *For each* $\mathbb{P}^{x,y}$ *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

$$f_0 := \int_{\mathbb{R}} f(y) e^{\alpha y^2} \Psi_{0,\lambda}(y) \phi(y)\, dy \tag{5.12}$$

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

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

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

where $\gamma_\lambda = -E'(\lambda)$ *and* f_0 *is given at equation (5.12).*

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 5.13 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 5.13, 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 $Z_\lambda^-(\infty) > 0$).

(5.14) **Other useful martingales and rates of convergence to zero.** Defining

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

we recall the following family of martingales:

(5.15) **Lemma.** *Let* $\lambda \in \big(\lambda_{\min}, 0\big)$. *For* $t \in [0, \infty)$

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

defines a martingale Z_λ^+ *for each* $\mathbb{P}^{x,y}$ *starting law.*

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.

(5.16) **Theorem.** *For every starting law,* $\mathbb{P}^{x,y}$*:*
(a) 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

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

(b) if $\lambda \in \big(\lambda_{\min}, \tilde{\lambda}(\theta)\big)$ *and* $n \in \{1, 2, \cdots\}$ *then for all* $\epsilon > 0$

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

The rate of convergence of the Z_λ^+ martingale in part (a) of 5.16 will immediately give the upper bound on the almost sure growth rate, $D(\gamma, \kappa) \leq \Delta(\gamma, \kappa)$. We also comment that *if* 5.13 was true for all $\alpha < \psi_\lambda^+$, then we could have gained this upper bound at that point. Although the result 5.13 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 $\alpha < \psi_\lambda^+$.

(5.17) **Corollary.** *For every starting law,* $\mathbb{P}^{x,y}$,
(a) if $\lambda \in (\lambda_{\min}, 0)$ *then* $Z_\lambda^+(t) \to 0$ *almost surely.*
(b) if $\lambda \in \big(\lambda_{\min}, \tilde{\lambda}(\theta)\big)$ *and* $n \in \{0, 1, \ldots\}$ *then* $e^{-n\mu_\lambda t} Z_{n,\lambda}(t) \to 0$ *almost surely.*

(5.18) **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[**14**] for this type of behaviour in the standard branching Brownian motion case.

6. Proof of the martingale results.

(6.1) **Proof of corollary 5.13.**

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

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

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 $\epsilon > 0$ small enough so that $E_\lambda - E_\mu + (\lambda - \mu)\gamma_\mu > 0$ and using theorem 5.11 we find that for any $\delta > 0$

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

Similarly, we can show

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

and hence combining with convergence theorem 5.11 we have

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

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

(6.2) **Proof of theorem 5.16**

(a) Remember each $Z_\lambda^\pm$ is a positive martingale. Remembering a useful technique brought to our attention in Neveu[**14**], let $p\in(0,1)$, then by Jensen's inequality, $Z_\lambda^\pm(t)^p$ is a supermartingale and since for $u,v>0$ we have

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

then

$$Z_\lambda^\pm(t)^p = \left|\sum_{k=1}^{N(t)} e^{\psi_\lambda^\pm Y_k(t)^2+\lambda\left(X_k(t)+c_\lambda^\pm t\right)}\right|^p \le \left(\sum_{k=1}^{N(t)} e^{p\psi_\lambda^\pm Y_k(t)^2+p\lambda\left(X_k(t)+c_{p\lambda}^- t\right)}\right) e^{p\lambda\left(c_\lambda^\pm - c_{p\lambda}^-\right)t}.$$

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

$$\mathbb{P}\left(\sup_{s\le u\le s+t} Z_\lambda^\pm(u)^p > \epsilon^p\right) \le \frac{\mathbb{E}Z_\lambda^\pm(s)^p}{\epsilon^p}$$
$$\le \epsilon^{-p}\left\{\mathbb{E}\sum_{k=1}^{N(s)} e^{p\psi_\lambda^\pm Y_k(s)^2+p\lambda\left(X_k(s)+c_{p\lambda}^- s\right)}\right\} e^{p\lambda(c_\lambda^\pm-c_{p\lambda}^-)s}$$

Then

$$\mathbb{P}\left(\sup_{s\le u\le s+t} e^{\delta u} Z_\lambda^\pm(u) > \epsilon\right) \le \mathbb{P}\left(\sup_{s\le u\le s+t} Z_\lambda^\pm(u)^p > e^{-p\delta(s+t)}\epsilon^p\right)$$
$$\le \epsilon^{-p}e^{p\delta t}\left\{\mathbb{E}\sum_{k=1}^{N(s)} e^{p\psi_\lambda^\pm Y_k(s)^2+p\lambda\left(X_k(s)+c_{p\lambda}^- s\right)}\right\} e^{p\left(\lambda(c_\lambda^\pm-c_{p\lambda}^-)+\delta\right)s}$$

Now, if we can choose $p\in(0,1)$ such that $\lambda(c_\lambda^\pm - c_{p\lambda}^-)+\delta<0$ and $p\psi_\lambda^\pm < \psi_{p\lambda}^+$, we must have $e^{\delta u}Z_\lambda^\pm(u)\to 0$ almost surely by using a familiar Borel-Cantelli argument. (The condition $p\psi_\lambda^\pm < \psi_{p\lambda}^+$ 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 (2.4), for example.)

For all $0 \le p < 1$ we find $p\psi_\lambda^\pm < \psi_{p\lambda}^+$. Considering the graph of $c_\lambda^\pm$ 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 5.8. Then, so far we have proved the following:

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

$$e^{-\epsilon t}\, e^{-\lambda(c_\lambda^\pm - c_\lambda^*)t} Z_\lambda^\pm(t) \to 0 \qquad a.s.$$

where

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

It is clear that this gives the required upper bound of

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

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

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

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 $\lambda \in (\tilde\lambda(\theta), 0)$,

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

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

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

and we have completed the proof of the first part of theorem 5.16.

(b) For the 'excited-state' martingales we can make a comparison to the 'ground-state' martingales as follows. Fix $n \in \{1, 2, \ldots\}$ 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 $|h_n(y)| \le C\,|y|^n$ for all $y \in \mathbb{R}$. Define $R_t := \sup_{k \le N(t)} Y_k(t)$, then

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

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

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

Combining the above with 5.16(a) now gives the required result. □

7. Proof of the main results.

(7.1) **Proof of theorem 1.7.**

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

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

Taking the infimum over λ,

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

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

$$\sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \le -\gamma t\right\} \le e^{\Delta(\gamma)t} Z_{\lambda_\gamma}^-(t).$$

Now, if $\gamma > \tilde{c}(\theta)$, corresponding to $\lambda_\gamma \in \left(\lambda_{\min}, \tilde{\lambda}(\theta)\right)$ and having $\Delta(\gamma) < 0$, we know from theorem 5.8 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,

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

Otherwise, if $\gamma \in \left(0, \tilde{c}(\theta)\right)$, corresponding to $\lambda_\gamma \in (\tilde{\lambda}(\theta), 0)$ and having $\Delta(\gamma) > 0$, theorem 5.8 tells us that $Z_{\lambda_\gamma}^-(\infty) > 0$ almost surely, hence

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

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

$$\sum_{k=1}^{N(t)} e^{\lambda X_k(t) - E_\lambda t} \mathrm{I}\left\{-(\gamma_\lambda + \epsilon)t \leq X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

$$\leq \sum_{k=1}^{N(t)} e^{\lambda(-(\gamma_\lambda + \epsilon)t) - E_\lambda t} \mathrm{I}\left\{-(\gamma_\lambda + \epsilon)t \leq X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

$$= e^{(-\lambda\gamma_\lambda - E_\lambda - \lambda\epsilon)t} \sum_{k=1}^{N(t)} \mathrm{I}\left\{-(\gamma_\lambda + \epsilon)t \leq X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

$$\leq e^{(-\lambda\gamma_\lambda - E_\lambda - \lambda\epsilon)t} \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

Then

$$t^{-1} \log \sum_{k=1}^{N(t)} e^{\lambda X_k(t) - E_\lambda t} \mathrm{I}\left\{|X_k(t)/t + \gamma_\lambda| < \epsilon\right\} \leq -\lambda\gamma_\lambda - E_\lambda - \lambda\epsilon$$

$$+ t^{-1} \log \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

Letting $t \to \infty$, using corollary 5.13 and remembering that for $\tilde{\lambda}(\theta) < \lambda \leq 0$ we have $Z_\lambda^-(\infty) > 0$ a.s., we find

$$0 \leq -\lambda\gamma_\lambda - E_\lambda - \lambda\epsilon + \liminf_{t \to \infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -(\gamma_\lambda - \epsilon)t\right\}$$

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

$$\liminf_{t \to \infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -\gamma_\lambda t\right\} \geq E_\lambda + \lambda\gamma_\lambda.$$

Equivalently,

$$\liminf_{t \to \infty} t^{-1} \log \sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \leq -\gamma t\right\} \geq E_{\lambda_\gamma} + \lambda_\gamma \gamma = \Delta(\gamma)$$

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.*[**4**], 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 5.11 found in Harris[**9**].

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 5.11). 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[**14**] 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[**19**], also exist.

(7.2) **Proof of theorem 1.13: the upper bound.**

Simply observe that for $\lambda \in (\lambda_{\min}, 0)$,

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

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

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

In cases where $\Delta(\gamma, \kappa) < 0$, we can use the optimal value for λ and theorem 5.16 to see that

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

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 $\overline{\lambda}(\gamma, \kappa) \in [\tilde\lambda(\theta), 0)$), 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 5.16 and the minimizing λ value, $\overline{\lambda}(\gamma, \kappa)$, to get the bound

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

as desired. □

(7.3) **Proof of theorem 1.13: the lower bound**

Consider $\alpha, \beta > 0$ and $y_0 > 0$ as fixed. Define

$$\Gamma(t) := \{i \le N(t) \ : \ X_i(t) \ge \alpha t, \ Y_i(t) \in [0, y_0]\}$$

Then earlier results tell us that

$$\frac{\log |\Gamma(t)|}{t} \to \Delta(\alpha) \qquad a.s.$$

as $t \to \infty$.

Let

$$Z(t) := \sum_{i \in \Gamma_\alpha(t)} \mathrm{I}\{A_i(t)\}$$

where, for any given constants $\eta, \delta > 0$ and $\tau > \eta$, $A_i(t)$ is the event that there exists a particle k alive at time $t + \tau$ which has descended from the particle i alive at time t and has

$$X_k(t+s) - X_k(t) \ge \beta t(1-\delta), \qquad Y_k(t+s)^2 \ge \kappa^2 t(1-\delta)$$

for all $s \in [\tau - \eta, \tau]$.

(7.4)**Lemma.** *Given $\epsilon, \delta, y_0 > 0$, $\exists \eta > 0, \tau > \eta, T > 0$ such that with*

$$p_t(y) = \mathbb{P}^{0,y}\Big(\exists k \le N(\tau) \ s.t. \ X_k(\tau - s) \ge \beta t(1-\delta), \ Y_k(\tau - s)^2 \ge \kappa^2 t(1-\delta), \ \forall s \in [0, \eta]\Big)$$

we have

$$\exp\left(-t\{\Theta(\beta,\kappa) + \epsilon\}\right) \le p_t(y) \le \left(e^{\psi_\lambda^+ y^2 + E_\lambda^+ \tau}\right) \exp\left(-t \ \Theta(\beta,\kappa) \ \right)$$

for $\forall y \in [0, y_0]$ and $\forall t > T$.

Proof. The upper bound is by making a simple over estimate of the probability by using the Z_λ^+ martingales. The lower bound is a simple corollary to the large deviation results from section 4.16 using the path $(\overline{x}, \overline{y})$ and its properties mentioned in sub-section 3.8. □

Now, conditionally on $\mathcal{F}_t$, the events $A_i(t)$ are all independent with

$$P_i(t) := \mathbb{P}\big(A_i(t) \big| \mathcal{F}_t\big) = p_t\big(Y_i(t)\big)$$

and then

$$\mathbb{E}(Z_t | \mathcal{F}_t) = \sum_{i \in \Gamma_\alpha(t)} P_i(t) =: Q(t),$$

$$\mathrm{var}\,(Z_t | \mathcal{F}_t) = \sum_{i \in \Gamma_\alpha(t)} P_i(t)\Big(1 - P_i(t)\Big) \le Q(t)$$

Using Chebychev's inequality, for all $\epsilon' > 0$,

$$\mathbb{P}\left(\left|\frac{Z_t}{Q_t} - 1\right| > \epsilon' \middle| \mathcal{F}_t\right) \leq \frac{\operatorname{var}(Z_t|\mathcal{F}_t)}{\epsilon'^2 Q(t)^2} \leq \frac{1}{\epsilon'^2 Q(t)}$$

Lemma 7.4 gives $p_t(y) \approx \exp(-\Theta(\beta,\kappa))$ and then we find that for all $\epsilon > 0$ and $\delta > 0$, we can choose $\eta > 0$, $\tau > \eta$ and find a $T < \infty$ such that

$$\frac{\log Q(t)}{t} \in [\Delta(\alpha) - \Theta(\beta,\kappa) - \epsilon, \Delta(\alpha) - \Theta(\beta,\kappa) + \epsilon] \qquad \forall t > T$$

Using the optimal parameters $\alpha = \overline{\alpha}$ and $\beta = \overline{\beta}$ where $\overline{\alpha} + \overline{\beta} = \gamma$ and if $\Delta(\gamma,\kappa) > 0$, we have $\Delta(\overline{\alpha}) - \Theta(\overline{\beta},\kappa) = \Delta(\gamma,\kappa) > 0$. Letting $t_n := n\eta$ for each $n \in \mathbb{N}$, the Borel-Cantelli lemma now reveals that only finitely many of the events

$$\left|\frac{Z_{t_n}}{Q(t_n)} - 1\right| > \epsilon'$$

occur almost surely. Then, almost surely there exists $T > 0$ such that for all $t > T$, if we choose n so $t \in [\tau + t_n, \tau + t_{n+1}]$ then

$$\sum_{k=1}^{N(t)} \mathrm{I}\left\{X_k(t) \geq \gamma t_{n+1}(1-\delta), Y_k(t)^2 \geq \kappa^2 t_{n+1}(1-\delta)\right\} \geq Z_{t_{n+1}} \geq (1-\epsilon')Q(t_{n+1})$$
$$\geq (1-\epsilon')\exp\left(t_{n+1}\{\Delta(\gamma,\kappa) - \epsilon\}\right)$$

Since $t_{n+1}/t \to 1$ as $t \to \infty$ and $\epsilon, \epsilon' > 0$ and $\delta > 0$ can all be made arbitrarily small, this gives us the required lower bound on the growth rate of particles near $(\gamma t, \kappa\sqrt{t})$ as $\Delta(\gamma,\kappa)$.
□

References

[**1**] Biggins, J. D. (1977) Martingale convergence in the branching random walk. *J. appl. Probab.* **14**, 25–37.

[**2**] Biggins, J. D. (1979) Growth rates in the branching random walk. *Z. Wahr.* **48**, 17–34.

[**3**] Biggins, J. (1992) Uniform convergence of martingales in the branching random walk, *Ann. Probab.*, **20**, 137–151

[**4**] Champneys, A., Harris, S. C., Toland, J. F., Warren, J. & Williams, D. (1995) Analysis, algebra and probability for a coupled system of reaction-diffusion equations, *Phil. Trans. Roy. Soc. London (A)*, **350**, 69–112.

[**5**] Chauvin, B. and Rouault, A. (1988) KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees., *Probability Theory and Related Fields*, **80**, 299–314.

[**6**] Git, Y. (1997) *Some aspects of large-deviations theory.* Ph.D. thesis, University of Bath.

[**7**] Git, Y. (1998) Almost sure path properties of branching diffusion processes, *Séminaire de Probabilités*, **XXXII**, Springer.

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

[**9**] Harris, S.C. (1999) Convergence of a 'Gibbs-Boltzmann' random measure for a typed branching diffusion, *Séminaire de Probabilités*, **XXXIV**, Springer.

[**10**] Itô, K. & McKean, H.P. (1965) *Diffusion Processes and their Sample Paths.* Springer, Berlin.

[**11**] Liptser, R.S.H. & Pukhalskii, A.A. (1992) Limit Theorems on Large Deviations for Semi-martingales, *Stochastics and Stochastics Reports*, **38**, 201–249.

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

[**13**] McKean, H. P. (1976) Correction to the above. *Comm. Pure Appl. Math.* **29**, 553–554.

[**14**] Neveu, J. (1987) Multiplicative martingales for spatial branching processes. *Seminar on Stochastic Processes* (ed. E. Çinlar, K. L Chung and R. K. Getoor), Progress in Probability and Statistics **15**. pp 223–241. Birkhäuser, Boston.

[**15**] Rogers, L.C.G. & Williams, D. (1994) *Diffusions, Markov processes and martingales. Volume 1: Foundations.* (Second Edition.) Wiley, Chichester and New York.

[**16**] Rogers, L.C.G. & Williams, D. (1987) *Diffusions, Markov processes and martingales. Volume 2: Itô Calculus.* Wiley, Chichester and New York.

[**17**] Strook, D.W. (1984) *An introduction to the theory of large deviations.* Springer-Verlag, Berlin.

[**18**] Varadhan, S. R. S. (1984) *Large deviations and applications.* CBMS-NSF regional conference series in applied mathematics, **46**. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

[**19**] Watanabe, S. (1967) Limit theorems for a class of branching processes. In *Markov processes and potential theory* (ed. J. Chover), 205-232. Wiley, New York.

[**20**] Wentzell, A.D. & Freidlin, M.I. (1984) *Random Perturbations of Dynamical Systems.* Springer-Verlag, Berlin.

[**21**] Wentzell, A.D. (1989) *Limit theorems on large deviations for Markov random processes.* Reidel, Dordrecht.