Some path large deviation results for a branching diffusion

Robert Hardy and Simon C. Harris*
Department of Mathematical Sciences, University of Bath,
Bath, BA2 7AY, UK.

November 13, 2007

Abstract

We give an intuitive proof of a path large-deviations result for a typed branching diffusion as found in Git, J.Harris and S.C.Harris [4]. Our approach involves an application of a change of measure technique involving a distinguished infinite line of descent, or *spine*, and we follow the spine setup of Hardy and Harris [5, 7, 6]. Our proof combines simple martingale ideas with applications of Varadhan's lemma, and is successful mainly because a 'spine decomposition' effectively reduces otherwise extremely difficult calculations on the whole collection of branching-diffusion particles down to just a single diffusing particle (the spine) whose large-deviations behaviour is well known. A similar approach was first used for branching Brownian motion in Hardy and Harris [6]. Importantly, our techniques should be applicable in a much wider class of branching diffusion large-deviation problems.

AMS subject classification: 60J80.

Keywords: branching diffusions; spatial branching process; path large deviations; spine decomposition; spine change of measure; additive martingales.

1 Overview

Harris and Williams [8] introduced a model of a branching diffusion in which the diffusion and breeding rate of particles is controlled by their type process which moves as an Ornstein-Uhlenbeck process on $\mathbb{R}$, independently of the particle's position, associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{with } \theta > 0 \text{ considered as the } \textit{temperature}. \tag{1}$$

Throughout this article we shall refer to an OU process with generator $\frac{\theta}{2}\frac{\partial^2}{\partial y^2} - \mu y\frac{\partial}{\partial y}$ as an $\mathrm{OU}(\theta,\mu)$.

More precisely, the spatial movement of a particle of type y is a driftless Brownian motion with instantaneous variance

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

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

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

and binary splitting occurs at the fission times. The model has very different behaviour for low temperature values (i.e. low θ), but throughout we consider that $\theta > 8r$ – the high temperature

*Corresponding author: S.C.Harris@bath.ac.uk

regime. Also, the parameter λ must be restricted to an interval $(\lambda_{\min}, 0)$ in order for some of the model's parameters to remain in $\mathbb{R}$.

We can suppose that the probabilities of this are $\{P^{x,y} : x, y \in \mathbb{R}\}$ so that $P^{x,y}$ is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of this branching diffusion process initiated from a single particle positioned at the space-type location (x, y). The configuration of this branching diffusion at time t is to be given by the $\mathbb{R}^2$-valued point process $\mathbb{X}_t := \{(X_u(t), Y_u(t)) : u \in N_t\}$ where N_t is the set of individuals alive at time t, and without loss of generality we can assume that the initial ancestor starts out at the space-type origin – henceforth we use P to mean $P^{0,0}$.

The main aim of this article is to prove upper and lower bounds for the probability of finding at least one of the branching particles very far from the space-type origin at large times, in a suitable large-deviations sense. The question of the lower bound was originally motivated by Git *et al* [4] in order to determine the exponential growth rates and asymptotic shape of the branching diffusion. We briefly discuss this application in [4] in section 2. Actually, our approach here naturally gives rise to a much stronger result where particles not only arrive at a very large space-type location $(\beta t, \kappa\sqrt{t})$ at a fixed time τ, but are also known to have stayed 'near' a specific space-type trajectory throughout the whole time interval $[0, \tau]$. The spine techniques used in this paper involve a change-of-measure that makes a single 'spine' particle closely follow this given trajectory. Whilst the upper bound result is new, more significantly our methods (for both bounds) should also facilitate far more general large-deviation results for branching diffusions, as we hope future work will reveal. Although Git *et al* [4] also used a spine change-of-measure, their original approach was far more model specific and quite different in flavour.

Let

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

and define a space-type trajectory $(x_s, y_s)_{s\in[0,\tau]}$ by

$$\overline{y}_s := \kappa\frac{\sinh\overline{\mu}s}{\sinh\overline{\mu}\tau}, \quad \overline{x}_s := a\overline{\lambda}\int_0^s y_w^2\,\mathrm{d}w, \qquad s \in [0, \tau]. \tag{3}$$

Note that the path endpoints are $y_\tau = \kappa$ and $x_\tau = \beta$. Also define,

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

Theorem 1.1 (The short-climb probability) *Let $\beta < 0$, $\kappa \in \mathbb{R}$ and $\varepsilon > 0$.*

(a) If $\tau > 0$ is sufficiently large, then for all $\delta, \delta' > 0$

$$\liminf_{t\to\infty} t^{-1}\log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|t^{-1}X_u(s)-\overline{x}_s\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(s)-\overline{y}_s\big| < \delta'\Big) \geq -\Theta(\beta,\kappa)-\varepsilon.$$

(b) If $\delta, \delta' > 0$ are sufficiently small, then for all $\tau > 0$

$$\begin{aligned}
&\limsup_{t\to\infty} t^{-1}\log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|t^{-1}X_u(s)-\overline{x}_s\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(s)-\overline{y}_s\big| < \delta'\Big)\\
&\quad\leq \limsup_{t\to\infty} t^{-1}\log P\Big(\exists u \in N_\tau : \big|t^{-1}X_u(\tau)-\beta\big| < \delta, \big|t^{-\frac{1}{2}}Y_u(\tau)-\kappa\big| < \delta'\Big)\\
&\quad\leq -\Theta(\beta,\kappa)+\varepsilon.
\end{aligned}$$

Note that the above new upper bound result shows that the trajectory followed really is 'optimal' in order to achieve the required large position at time τ. We will actually prove a more general result for *any* (as opposed to sufficiently large) fixed time τ with a rate of decay $J(\tau)$, where $J(\tau) \downarrow \theta(\beta, \kappa)$. We state this stronger result as Theorem 3.2. We also note that the above lower bound on the probability of following the optimal paths to reach $(\beta t, \kappa\sqrt{t})$ is for a large *fixed* time τ. In contrast, the method used for the analogous result in [4, Theorem 7] dictated that $\tau = \tau(t) \propto \log t$. Our stronger result would enable a corresponding simplification in the proof of the application in [4].

The principle behind the proof of the lower bound is to design new measures $\mathbb{Q}_t$ for the branching diffusion such that one of the particles (the spine) will closely follow a specific space-type path. Our spine approach, which we briefly lay out in section 4 (and which is fully presented in Hardy and Harris [5]), will allow us to explicitly find the Radon-Nikodym derivatives (martingales) of these new measures with respect to the original measure P. Then, using the spine decomposition together with Doob's submartingale inequality, we shall show that the growth rate of these martingales under $\mathbb{Q}_t$ is exactly the correct rate for the large-deviations lower bound.

In such branching diffusion settings, we comment that a large-deviations upper bound is usually easier to obtain than the lower bound. Generally, we can over estimate the probability that any particle succeeds in performing a certain 'rare' event by the expected number of particles performing that event, which then reduces to a single particle (large deviation) calculation. In the present context, the upper bound of Theorem 1.1 can also be proved directly using some fundamental 'additive' martingales.

The layout of this article is as follows: in the next section we discuss the results of Git *et al* [4] in order to give a context to our work. Section 3 contains a heuristic discussion of the large deviations for the model which motivates the choice of the subsequent martingales. A statement of a stronger path large deviation result (Theorem 3.2) is also found in this section. In section 4, we briefly present the foundations of our spine approach, giving definitions of the underlying space and its filtrations, measures and the fundamental martingales of interest. These strictly-positive martingales, Z_t, are defined in terms of specific paths as suggested by our heuristic arguments. As Radon-Nikodym derivatives, these martingales can define the new measures $\mathbb{Q}_t$ (under which they become *sub*martingales) and we state a key result (Theorem 4.3) on their growth under the measures $\mathbb{Q}_t$; this growth result leads directly to the proof of the large-deviations lower bounds which we present in section 5. (An alternative approach to the large deviation lower bounds is also briefly discussed in section 8.) Section 6 contains proofs for the upper bounds of the two main large deviation results. Section 7 is devoted to proving the martingale growth Theorem 4.3; the proof is not particularly short, but neither is it difficult given the powerful spine technology. It should be noted that this result is the main application of spines in this article, using the so-called *spine decomposition* to reduce the branching-particle martingales Z_t to become a function of just one particle (the spine) whose growth can be determined by a standard application of Varadhan's lemma.

2 The Git *et al.* almost-sure result

Before we move on to prove the above theorem, we summarize the main results from Git *et al* [4] and the earlier Harris and Williams [8] so that the reader might understand how Theorem 1.1 fits into the picture.

Work on large-deviations results for this typed branching-diffusion began in the paper by Harris and Williams [8], where they considered the behaviour in expectation of the counting function

$$N_t(\gamma) = \sum_{u \in N_t} \mathbf{1}_{(X_u(t) \le -\gamma t)}$$

(not to be confused with N_t, the set of individuals alive at time t) for each $\gamma \in \mathbb{R}$. In Git *et al.* [4] it was shown that

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

exists almost surely and is *finite* for all $0 \leq \gamma < \tilde{c}(\theta)$, for some constant $\tilde{c}(\theta)$; in the case that $\gamma \geq \tilde{c}(\theta)$ the limit is $-\infty$ since *no* particles will be near the ray at large times. In other words, this result says that we almost-surely have exponential growth in numbers of particles following close to rays that are *not too steep*.

For later reference we state that

$$\Delta(\gamma) = \inf_{\lambda\in(\lambda_{\min},0)} \big\{E_\lambda^- + \lambda\gamma\big\} = \rho + \frac{\theta}{4} - \frac{1}{4}\sqrt{\theta(\theta-8r)(1+4\gamma^2/(\theta a))},$$

where $E_\lambda^- \in \mathbb{R}$ is an eigenvalue described in the next section, and

$$\tilde{c}(\theta) := \sup\big\{\gamma : \Delta(\gamma) > 0\big\} = \sqrt{2a\Big(r + \rho + \frac{2(2r+\rho)^2}{\theta - 8r}\Big)}.$$

The work of Git *et al* [4] improves this to obtain the almost-sure rate of growth in numbers of particles at certain *spatial and type* positions at large times. They study the following function that counts how many particles occupy a particular region in the type-space domain:

$$N_t(\gamma,\kappa) := \sum_{u\in N(t)} \mathbf{1}\big\{X_u(t) \leq -\gamma t, Y_u(t)^2 \geq \kappa^2 t\big\}.$$

Theorem 2.1 *Under each $P^{x,y}$ law, the limit*

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

exists almost-surely and is given by

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

Here,

$$\begin{aligned} \Delta(\gamma,\kappa) &= \inf_{\lambda\in(\lambda_{min},0)} \big\{E_\lambda^- + \lambda\gamma - \kappa^2\psi_\lambda^+\big\}, \\ &= \rho + \frac{\theta-\kappa^2}{4} - \frac{1}{4a\theta}\sqrt{\theta(\theta-8r)(4a\theta\gamma^2 + a^2(\theta+\kappa^2)^2)}. \end{aligned} \tag{5}$$

2.1 The almost-sure upper bound

Harris and Williams [8] showed that there are *two* strictly-positive martingales Z_λ^- and Z_λ^+ defined as

$$Z_\lambda^\pm(t) := \sum_{k=1}^{N(t)} v_\lambda^\pm(Y_k(t))e^{\lambda X_k(t) - E_\lambda^\pm t}, \tag{6}$$

where v_λ^- and v_λ^+ are strictly positive eigenfunctions of the self-adjoint operator $\frac{1}{2}\lambda^2 A + R + Q_\theta$, with corresponding eigenvalues $E_\lambda^- < E_\lambda^+$. The explicit form for these eigenfunctions is

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

where $\psi_\lambda^\pm := \frac{1}{4} \pm \frac{\mu_\lambda}{2\theta}$, for a positive parameter μ_λ, and $\psi_\lambda^\pm$ are both positive for all $\lambda \in (\lambda_{\min}, 0)$.

A useful trick to obtain upper bounds is to overestimate indicator functions by an exponential and optimise over parameters. It is often the case that this will bring in one of the martingales of the model: for $\lambda \in (\lambda_{\min}, 0)$,

$$\sum_{k=1}^{N(t)} \mathbf{1}\big\{X_k(t) \le -\gamma t, Y_k(t)^2 \ge \kappa^2 t\big\} \le \sum_{k=1}^{N(t)} \exp\big\{\psi_\lambda^+ (Y_k(t)^2 - \kappa^2 t)\big\} \exp\big\{\lambda(X_k(t) + \gamma t)\big\}$$
$$= e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{(E_\lambda^- + \lambda\gamma - \kappa^2 \psi_\lambda^+)t}. \tag{7}$$

(Importantly for this, the parameter ψ_λ^+ is positive and λ is negative; the functions c_λ^- and c_λ^+ are defined as $c_\lambda^\pm := E_\lambda^\pm/(-\lambda)$.)

The expression for $\Delta(\gamma, \kappa)$ as a Legendre-conjugate – see (5) – explains why $\Delta(\gamma, \kappa)$ relates to (7) above: by choosing λ at the infimum we get

$$N_t(\gamma, \kappa) \le e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{\Delta(\gamma,\kappa)t}. \tag{8}$$

We remember that $N_t(\gamma, \kappa)$ takes only integer values, and a separate theorem by Harris and Git states that

$$\limsup_{t\to\infty} e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) \le 0, \qquad \text{for each } \lambda \in (\lambda_{\min}, 0). \tag{9}$$

Thus if $\Delta(\gamma, \kappa) < 0$ we deduce that almost surely

$$N_t(\gamma, \kappa) = 0, \quad \text{eventually},$$

whence $\lim_{t\to\infty} t^{-1} \log N_t(\gamma, \kappa) = -\infty$, as required.

On the other hand, if $\Delta(\gamma, \kappa) \ge 0$, (8) and (9) immediately imply that

$$\lim_{t\to\infty} t^{-1} \log N_t(\gamma, \kappa) \le \Delta(\gamma, \kappa).$$

□

2.2 A two-phase mechanism for the lower bound

For their proof of the almost-sure lower bound of Theorem 2.1, Git *et al* [4] propose an explicit mechanism by which a sufficient number of particles will obtain a position near $(\gamma T, \kappa\sqrt{T})$ in the type-space domain at large times T. It is made up of two phases:

the long tread: Over a long period $[0, t]$, taking up nearly all of the time, a large number of particles will drift spatially with speed $\gamma\theta/(\theta + \kappa^2)$ – as if their type has had a modified occupation measure, as described by Harris and Williams [8];

the short climb: Following this, over a short period of time $[t, t + \tau]$ with τ a fixed time ($\tau \ll t$), each of the particles from this group will have a small probability of further rushing to the large type position $\kappa\sqrt{t}$ whilst additionally gaining $\big\{\gamma\kappa^2/(\theta + \kappa^2)\big\}t$ in spatial position.

The combination of these two phases will present us with *sufficiently many* particles at the space-type position $\big(\gamma T, \kappa\sqrt{T}\big)$ at the large time $T = t + \tau$, as Git *et al* [4] show in their proof of the lower bound of Theorem 2.1 – we refer the reader to their work for further details. This lower bound requires a substantial amount of technical work, mainly focussed on the short climb in which, in particular, they required $\tau = O(\log t)$. Our Theorem 1.1 includes an alternative *short climb* lower bound for τ fixed and using this would slightly simplify the combination of phases. Our current proof will also provide a cleaner, more intuitive and more generic approach to such path large deviation results in branching diffusions.

3 Large deviations heuristics

We now present some heuristic arguments concerning the large-deviations behaviour of the branching diffusion which will serve as the *intuition* behind our later *rigourous* proofs.

Under a measure $\tilde{P}$ let (ξ_s, η_s) satisfy

$$\mathrm{d}\eta_s = \sqrt{\theta}\mathrm{d}B_s - \frac{\theta}{2}\eta_s \mathrm{d}s, \quad \text{and} \quad \mathrm{d}\xi_s = \sqrt{a}\eta_s \, \mathrm{d}W_s,$$

for two independent $\tilde{P}$-Brownian motions B_s and W_s. We will call (ξ_s, η_s) the *spine* and, under $\tilde{P}$, it moves like a single particle within the branching diffusion. For a large-deviations analysis we observe that for any $t > 0$,

$$\mathrm{d}\Big(\frac{\eta_s}{\sqrt{t}}\Big) = \sqrt{\theta}\Big(\frac{\mathrm{d}B_s}{\sqrt{t}}\Big) - \frac{\theta}{2}\Big(\frac{\eta_s}{\sqrt{t}}\Big)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}\Big(\frac{\xi_s}{t}\Big) = \sqrt{a}\Big(\frac{\eta_s}{\sqrt{t}}\Big)\Big(\frac{\mathrm{d}W_s}{\sqrt{t}}\Big),$$

and therefore it is appropriate to work with the *re-scaled* spine $(\xi_s/t, \eta_s/\sqrt{t})$ since in this way we obtain a variance coefficient of $1/\sqrt{t}$ on the driving Brownian motions.

Definition 3.1 *For each $t > 0$ we define*

$$\xi^t_s := \xi_s/t, \quad \textit{and} \quad \eta^t_s := \eta_s/\sqrt{t},$$

and call (ξ^t_s, η^t_s) the ***re-scaled spine****. We note that under $\tilde{P}$ we have for $s \in [0, \tau]$:*

$$\mathrm{d}\eta^t_s = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}B_s - \frac{\theta}{2}\eta^t_s \mathrm{d}s, \quad \textit{and} \quad \mathrm{d}\xi^t_s = \frac{\sqrt{a}\eta^t_s}{\sqrt{t}}\mathrm{d}W_s,$$

for two independent $\tilde{P}$-Brownian motions B_s and W_s.

Throughout the remainder of this article, and different from the earlier parts, the variable t will not be a time parameter but will bring about this large-deviations scaling; typically we shall use either w or s to denote the time parameter from the time-interval $[0, \tau]$ where $\tau > 0$ is considered as fixed.

Suppose that we are given two paths: a type-path $y : [0, \tau] \to \mathbb{R}$ and a spatial-path $x : [0, \tau] \to \mathbb{R}$. On a *heuristic* level we can say that the probability of the type-diffusion η^t_s closely following y and the space-diffusion ξ^t_s closely following x is roughly

$$\exp\Big(-\frac{t}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2 \,\mathrm{d}s - \frac{t}{2}\int_0^\tau \frac{\dot{x}_s^2}{a y_s^2}\,\mathrm{d}s\Big), \tag{10}$$

for large enough t. See [2], for example, for the large deviation theory of Wentzell-Friedlin.

The reader who is familiar with the large-deviations principle for branching Brownian motion (see Hardy and Harris [6] for a spine proof, or Lee [16] for a classical proof) might make the reasonable guess that the *probability at least one* of the *re-scaled* branching particles $\big(X_u(s)/t, Y_u(s)/\sqrt{t}\big)$ follows the *difficult* type-path y_s and space-path x_s closely over the time interval $[0, \tau]$ is roughly

$$\exp\Big\{-\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2 + \frac{1}{2}\frac{\dot{x}_s^2}{a y_s^2} - r y_s^2 \,\mathrm{d}s\Big)t - \rho w\Big]\Big\},$$

when t is large. This *guess* is obtained by first estimating the *probability* by the *expected number of particles* following the difficult path, and then using the 'one-particle picture' of section 4 and large deviations theory for one particle.

By standard optimization arguments (Git *et al* [4] give some details of how this can be carried out) this implies that the probability of at least one of the re-scaled branching particles being near the space-type position (β, κ) at a fixed time τ (which is also the event that the non-rescaled particles arrive near $(\beta t, \kappa\sqrt{t})$ of course) should be roughly

$$\exp\Big\{-\inf_{x,y}\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2+\frac{1}{2}\frac{\dot{x}_s^2}{ay_s^2}-ry_s^2\,\mathrm{d}s\Big)t-\rho w\Big]\Big\}, \tag{11}$$

when t is large, and where the infimum is taken over all paths $x, y \in C[0,\tau]$ satisfying

$$y(0)=0, y(\tau)=\kappa, x(0)=0, x(\tau)=\beta. \tag{12}$$

This is typical in a large-deviations setting: although there are many possible trajectories that the (re-scaled) particles could travel along to get to a position (β, κ), the *dominant number* will have followed *optimal* paths.

Although the preceding arguments have been presented as if one is free to choose *any* paths x and y, we note that if $y_s = 0$ when $\dot{x}_s \neq 0$ then a rigourous approach to these arguments may have problems with the term $\frac{\dot{x}_s^2}{ay_s^2}$ in (11) – the heuristics are not really problematic if we interpret this as saying that the probability is $e^{-\infty} = 0$. On an intuitive level this is an expression of the fact that to have $y_s = 0$ equates to turning off the Brownian variance in the spatial diffusion which in turn would imply that no spatial progress is possible and therefore $\dot{x}_s = 0$ would be needed for consistency.

Git *et al* [4] state that for any given type-path y, the optimal space-path x for (11) under the constraint $x(\tau) = \beta$ will always be given by

$$x_s = \lambda \int_0^s ay_w^2\,\mathrm{d}w, \qquad \text{for } s\in[0,\tau], \tag{13}$$

for some value $\lambda \in \mathbb{R}$. Briefly, their arguments rely on the fact that in the definition of our model the spatial diffusion $X_u(s)$ of the branching particles can be seen as a time-changed Brownian motion where the time scaling is determined by its type process $Y_u(s)$:

$$X_u(s) = \hat{B}\Big(\int_0^s aY_u(w)^2\,\mathrm{d}w\Big)$$

for a Brownian motion $\hat{B}(\cdot)$ on $[0,\tau]$. A measure change that introduces a linear drift of λ to this Brownian motion will give

$$X_u(s) = \tilde{B}\Big(\int_0^s aY_u(w)^2\,\mathrm{d}w\Big) + \lambda\int_0^s aY_u(w)^2\,\mathrm{d}w,$$

where $\tilde{B}(\cdot)$ is a Brownian motion under the new measure – this clearly relates to (13). Linear drifts are the optimal path (in a large-deviations sense) for a Brownian motion to be at a given point at a given time, and the constraint $x(\tau) = \beta$ for our problem will determine the value of λ in terms of the type-path y:

$$\lambda = \frac{\beta}{a\int_0^\tau y_s^2\,\mathrm{d}s}. \tag{14}$$

Thus for the event being considered in Theorem 1.1, the optimal spatial-path x is determined *uniquely* by (13) together with (14). Therefore an equivalent but easier statement of our large-deviations result is that the probability of at least one of the re-scaled branching particles being near the space-type position (β, κ) at a fixed time τ is roughly

$$\exp\Big\{-\inf_{y}\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2+\frac{a\lambda^2}{2}y_s^2-ry_s^2\,\mathrm{d}s\Big)t-\rho w\Big]\Big\}, \tag{15}$$

when t is large, and where the infimum is taken over all paths $y \in C[0,\tau]$ and all $\lambda \in (\lambda_{\min}, 0)$ satisfying

$$y(0) = 0, y(\tau) = \kappa, \quad \lambda = \frac{\beta}{a\int_0^\tau y_s^2 \,\mathrm{d}s}. \tag{16}$$

Git *et al* [4] presented alternative heuristic arguments based on birth-death processes to arrive at the expression (15). Using Euler-Lagrange techniques they showed that the specific path

$$y_s = \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0,\tau] \tag{17}$$

is optimal for this expression, where

$$\mu_\lambda = \frac{\sqrt{\theta(\theta - 8r - 4a\lambda^2)}}{2}, \tag{18}$$

and $\lambda \in (\lambda_{\min}, 0)$ is dependent on the choice of τ (which we are anyway considering as fixed throughout) and is chosen to satisfy

$$\frac{\beta}{a\lambda} = \kappa^2\Big(\frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau}\Big). \tag{19}$$

We refer the reader to Git *et al* [4] for details of these relationships between the parameters, but note that particles staying close to this path will arrive near $y(\tau) = \kappa$ at time τ in agreement with the heuristics.

As we mentioned just before the statement of Theorem 1.1, our spine techniques will naturally use the path y_s defined at (17) together with x_s defined at (13), since they are the optimal paths (in a large-deviations sense) for accumulating particles near the point $(\beta t, \kappa\sqrt{t})$ at time τ. In fact our spine proof of Theorem 1.1 will result in a proof of the following stronger result, from which Theorem 1.1(a) would follow as a corollary.

Theorem 3.2 *Let $\tau > 0$ be fixed and suppose $\beta < 0$ and $\kappa \in \mathbb{R}$. Define two paths on $[0,\tau]$ by*

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad x_s := a\lambda \int_0^s y_w^2 \,\mathrm{d}w, \qquad s \in [0,\tau], \tag{20}$$

where $\lambda \in (\lambda_{\min}, 0)$ is chosen so that

$$\beta = a\lambda \int_0^\tau y_w^2 \,\mathrm{d}w = a\lambda\kappa^2\Big(\frac{\coth \mu_\lambda \tau}{2\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau}\Big). \tag{21}$$

Note that the path endpoints are $y_\tau = \kappa$ and $x_\tau = \beta$. Define

$$J(\tau) := \int_0^\tau \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2 - r y_s^2\Big] \,\mathrm{d}s = \lambda\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth \mu_\lambda \tau\Big). \tag{22}$$

(a) For all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|X_u(s) - t x_s\big| < \delta t, \big|Y_u(s) - \sqrt{t} y_s\big| < \delta'\sqrt{t}\Big) \geq -J(\tau),$$

(b) Let $\varepsilon > 0$. For all sufficiently small $\delta, \delta' > 0$,

$$\limsup_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|X_u(s) - t x_s\big| < \delta t, \big|Y_u(s) - \sqrt{t} y_s\big| < \delta'\sqrt{t}\Big) \leq -J(\tau) + \varepsilon.$$

It can be easily checked that $J(\tau) \downarrow \Theta(\beta,\kappa)$, $\lambda \to \overline{\lambda}$ and $(x_s, y_s) \to (\overline{x}_s, \overline{y}_s)$ as $\tau \to \infty$. Theorem 1.1 can now easily be deduced from Theorem 3.2 by choosing τ sufficiently large.

Although some additional work would be required to prove as much, the paths (x_s, y_s) above are chosen as they are the 'best' ones for particles to follow in order to reach position $(\beta t, \kappa\sqrt{t})$ at (fixed) time τ, as found in the large deviation heuristics discussed above and in [4].

Importantly, we emphasise that it should be possible to develop the ideas and techniques used in this article to obtain proofs of large-deviations principles for many other branching-diffusion models, essentially because we can reduce the branching particles down to the spine and in general this gives a technique for deriving large-deviations principles for the branching diffusion from those of the single diffusing particle (the spine) which are already well studied.

4 The spine approach, martingales and measures

In this section we will introduce some of the key concepts used in the proofs of the main results. We shall construct the branching diffusion with a distinguished infinite line of decent, the *spine*, and then perform a change of measure that will make the spine 'closely' follow a given path. Estimates on the martingale associated with this change of measure can then give a lower-bound for large-deviations events in the branching diffusion. The heuristics of the previous section will serve as an important guide. However, although they have already indicated a specific path at (17), it should be noted that in our proofs we use properties of this path only at a few points – elsewhere the techniques can be applied in general to any path. Therefore the reader may suppose that $y : [0,\tau] \to \mathbb{R}$ is any given and fixed path, and we shall be very careful to highlight those points when we use specific properties of the path defined at (17). Also, to keep notational complexity to a reasonable minimum we tend not to make the dependencies of the martingales and action functionals on the underlying chosen paths explicit in the notation.

The spine setup. Recall that the original branching diffusion $\mathbb{X}_t := \big\{(X_u(t), Y_u(t)) : u \in N_t\big\}$ where N_t is the set of individuals alive at time t has associated probability measures $P^{x,y}$ with natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$. We label all particles according to the Ulam-Harris convention. For example, '$\emptyset 213$' represents 'the 3rd child of the 1st child of the 2nd child of the initial ancestor'. For two labels $v, u \in \Omega$ the notation $v < u$ means that v is an *ancestor* of u, $|u|$ is the generation of particle u, and so forth.

A *spine* ξ is a *distinguished* infinite line of descent starting with the initial ancestor, where $\xi = \{\xi_0, \xi_1, \xi_2, ...\}$ with $\xi_0 = \emptyset$, ξ_n the label of the spine at the n-th generation and $u \in \xi$ means that $u = \xi_i$ for some $i \geq 0$. Let $\{(\xi(t), \eta(t))\}_{t\geq 0}$ represent the space-type path of spine, that is, the time position of the spine at time t is $(\xi(t), \eta(t)) := (X_u(t), Y_u(t))$ for $u \in N(t) \cap \xi$. Define $n = \{n_t : t \geq 0\}$ to be the counting process for the number of fissions that have occurred along the path of the spine by time t, with the actual fission times along the spine denoted by $\{S_i\}_{\geq 1}$. Note that ξ_{n_t} is the label of the spine at time t.

We will make important use of a variety of filtrations for the process with a distinguished spine. Let the *enriched* filtration for the branching diffusion with distinguished spine be $\widetilde{\mathcal{F}}_t = \sigma\left(\mathcal{F}_t, \{\xi_{n_s}\}_{s\leq t}\right)$. Then $\tilde{\mathcal{F}}_t$ knows everything up to time t, all particle paths, genealogy and identification of spine, whereas $\mathcal{F}_t$ knows about the paths and genealogy of all particles up to the time t, but does not know the identity of the spine. In addition, let $\tilde{\mathcal{G}}_t := \sigma\left(\{(\xi(s), \eta(s), n_s, \xi_{n_s})\}_{s\leq t}\right)$ and $\mathcal{G}_t := \sigma\left(\{(\xi(s), \eta(s))\}_{s\leq t}\right)$. Then $\tilde{\mathcal{G}}_t$ knows everything *along* the spine up to time t - the *spine's* motion, the *spine's* genealogy and the *spine's* fission times. It doesn't know about any information 'off' the spine. On the other hand, $\tilde{\mathcal{G}}_t$ only knows about the *spine's* motion but not about the births along the spine, the spine's genealogical information, nor any information 'off' the spine.

The spine construction. Under a measure $\tilde{P}^{x,y}$, the branching diffusion $(\mathbb{X}_s)_{s\geq 0}$ with distinguished spine ξ is constructed as follows:

- the spine process (ξ_s, η_s) starts at (x, y) and diffuses as a solution to

$$\mathrm{d}\eta_s = \sqrt{\theta}\mathrm{d}B_s - \frac{\theta}{2}\eta_s \mathrm{d}s, \quad \text{and} \quad \mathrm{d}\xi_s = \sqrt{a}\eta_s \,\mathrm{d}W_s, \tag{23}$$

where B_s and W_s are standard Brownian motions.

- at rate $R(\eta_s)$ the spine undergoes fission producing two particles;
- with equal probability, one of these two particles is selected to continue the spine ξ;
- the other particle initiates, from its birth space-type position, an independent copy of the original P branching diffusion with branching rate $R(\cdot)$.

We note that $\tilde{P}^{x,y}$ is an extension of the original measure $P^{x,y}$, with $P = \tilde{P}|_{\mathcal{F}_\infty}$. In fact, it is easy to see that an alternative way to construct the process under $\tilde{P}$ is to first construct the entire tree of the branching diffusion according to P and, secondly, choose the spine by starting from the initial ancestor then following the spine path forward in time with independent uniform choices made from the particles produced at each fission. In particular, for $u \in N(t)$ we note that $\tilde{P}(u \in \xi|\mathcal{F}_t) = \prod_{v<u} 2^{-1} = 2^{-|u|}$.

We also have the very useful and intuitive *'one particle picture' (OPP)*. For example, for any measurable function f of single particle paths on $[0, t]$,

$$\tilde{P}^{x,y}\left(\sum_{u\in N(t)} f(X_u(s), Y_u(s); s \le t)\right) = \tilde{P}^{x,y}\left(e^{\int_0^t \beta(\eta(s))\,ds} f(\xi(s), \eta(s); s \le t)\right)$$

For further details of this spine set-up and various key results, see Hardy and Harris[5]. Also see Lyons *et al* [17, 13, 18] and other recent work based on these (examples are Kyprianou [14], Kyprianou and Sani [15], Athreya [1], Olofsson [19] amongst others) for similar spine based approaches in branching processes.

Changes of measure. We can now able to perform our 'spine change of measure'. For any $t > 0$ and any given $y : [0, \tau] \to \mathbb{R}$ that is square-integrable along with its derivative

$$\exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_s - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big) \times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big),$$

is a strictly-positive $\tilde{P}$-martingale over the time period $w \in [0, \tau]$ (see Øksendal [12] for example). As one part of the change of measure defined below, this martingale will introduce drift terms into the diffusions η_s and ξ_s such that $\eta_s^t \sim y_s$ and $\xi_s^t \sim a\lambda y_s^2$ when t is large, and we note a comparison between this martingale and the expression (15) above.

The process n_w which counts the number of fission times on the spine up to time w is a Cox process of rate $R(\eta_s)$ and therefore for $w \in [0, \tau]$,

$$w \mapsto e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w}$$

is also a $\tilde{P}$-martingale. We can use the product of these two martingales to define a new measure:

Theorem 4.1 (Spine change of measure.) *Let $\tau > 0$ be fixed. For $t > 0$, we define a measure $\tilde{\mathbb{Q}}_t$ on $\tilde{\mathcal{F}}_\tau$ where*

$$\begin{aligned}\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_t}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_w} := \tilde{\zeta}_t(w) := \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_s - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big) \\ \times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big) \times e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w},\end{aligned} \tag{24}$$

for $w \in [0,\tau]$. Under the measure $\tilde{\mathbb{Q}}_t^{x,y}$ we can give a pathwise construction of the branching-diffusion $(\mathbb{X}_s)_{s\in[0,\tau]}$ with distinguished spine ξ:

- *the spine process (ξ_s,η_s) starts at (x,y) and diffuses as a solution to*

$$\mathrm{d}(\eta_s - \sqrt{t}y_s) = \sqrt{\theta}\mathrm{d}\tilde{B}_s - \frac{\theta}{2}(\eta_s - \sqrt{t}y_s)\mathrm{d}s \tag{25}$$

  and

$$\mathrm{d}\xi_s = \sqrt{a}\eta_s\,\mathrm{d}\tilde{W}_s + a\lambda\sqrt{t}y_s\eta_s\,\mathrm{d}s, \tag{26}$$

  where $\tilde{B}$ and $\tilde{W}$ are standard Brownian motions under $\tilde{\mathbb{Q}}_t$, with

$$\mathrm{d}\tilde{B}_s = \mathrm{d}B_s - \frac{\sqrt{t}}{\sqrt{\theta}}\Big(\dot{y}_s + \frac{\theta}{2}y_s\Big)\mathrm{d}s, \qquad \mathrm{d}\tilde{W}_s = \mathrm{d}W_s - \sqrt{at}\lambda y_s\,\mathrm{d}s;$$

- *at the accelerated rate $2R(\eta_s)$ the spine undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected to continue the spine;*
- *the other particle initiates, from its birth space-type position, an independent copy of the original P branching diffusion with normal branching rate $R(\cdot)$.*

Due to our formulation of the underlying spine foundations in terms of filtrations and sub-filtrations, we can project this new measure $\tilde{\mathbb{Q}}_t$ down onto the branching-diffusion particles and define a measure $\mathbb{Q}_t$ on $\mathcal{F}_\tau$ by $\mathbb{Q}_t := \tilde{\mathbb{Q}}_t|_{\mathcal{F}_\tau}$.

Theorem 4.2 *Let $\tau > 0$ be fixed. For each fixed $t > 0$, define the $\big((\mathcal{F}_w)_{0\le w\le\tau}, P\big)$-martingale $Z_t(w)$ for $w \in [0,\tau]$ by*

$$Z_t(w) := \frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}P}\bigg|_{\mathcal{F}_w} = \tilde{P}\big(\tilde{\zeta}_t(w)|\mathcal{F}_w\big).$$

Then

$$Z_t(w) = \sum_{u\in N(w)} f_{t,w}(u)$$

where

$$f_{t,w}(u) := e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}\exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w(\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_u(s) - \frac{t}{2\theta}\int_0^w(\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big)$$
$$\times\exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_u(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big) \tag{27}$$

and $B_u(s)$ and $W_u(s)$ to denote the P-Brownian motions driving the type and spatial processes of particle u in the branching diffusion.

See Hardy and Harris [5], or Engländer and Kyprianou [3], for more details and how to prove such results.

The spine decomposition. Consider those particles alive at time τ and group them together according to the time that they first branched off the spine's path. Since, under $\tilde{\mathbb{Q}}_t$ particles 'off' the spine behave as if under P and since Z_t is a P-martingale, it is easy to see the following 'spine decomposition':

$$\tilde{\mathbb{Q}}_t(Z_t(\tau)|\tilde{\mathcal{G}}_\infty) = f_{t,\tau}(\xi) + \sum_{i=1}^{n_t} f_{t,S_i}(\xi)$$

where

$$f_{t,w}(\xi) := e^{-\int_0^w R(\eta(s))\,\mathrm{d}s} \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B(s) - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big)$$
$$\times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big) \quad (28)$$

The spine decomposition is a very powerful tool and, in particular, it will prove essential for our key martingale growth estimates.

The growth of martingale Z_t. For our proof of Theorem 1.1 (and its stronger version of Theorem 3.2) it is important to know how quickly $Z_t(\tau)$ grows under the measure $\mathbb{Q}_t$. The following key result is the main application of spines in this article:

Theorem 4.3 *For the specific path y defined at (17), and for any $\alpha \in [0,1]$ we have*

$$\limsup_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) \le \alpha J(\tau) + \alpha^2 M(\tau),$$

where we define

$$J(w) := \int_0^w \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2 - ry_s^2\Big]\,\mathrm{d}s, \quad (29)$$

and

$$M(w) := \int_0^w \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2\Big]\,\mathrm{d}s. \quad (30)$$

We emphasize that without the technology of spines the proof of this result would be exceptionally difficult – witness the proofs of $\mathcal{L}^p$-convergence in Harris [9] for the simpler martingale Z_λ, where this classical approach succeeded mainly thanks to a martingale inequality taken from Biggins [10] and the fact that explicit $\mathcal{L}^2$-theory was available. It can notoriously difficult to deal with operations like $Z_t(\tau)^\alpha$ since these martingales Z_t are defined via *sums*, and classical inequalities aren't always good enough.

In contrast to this, the spine decomposition gives us a proper methodology for reducing the additive structure of these martingales to essentially a single-particle problem, and since it does this through a conditional-expectation operation rather than with an inequality, it is *exact* and therefore can lead to tight estimates that are useful. Due to its length, we dedicate the whole of section 7 to the spine proof of this above theorem, and now proceed to show how this result can be used to obtain the upper-bound on $Z_t(\tau)$ that we require for Theorem 1.1.

It is not difficult to verify that for any $\alpha \in [0,1]$, $Z_t(w)^\alpha$ is a submartingale with respect to the measure $\mathbb{Q}_t$. Given Theorem 4.3, we can therefore use Doob's submartingale inequality to prove the following:

Theorem 4.4 *Let $\tau > 0$ be fixed. Then for all $\varepsilon > 0$,*

$$\lim_{t\to\infty}\mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \le e^{(J(\tau)+\varepsilon)t}\Big) \to 1.$$

Proof: For a given $\varepsilon > 0$ and for any $\alpha \in [0,1]$, Doob's inequality gives

$$\mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t}\Big) = \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s)^\alpha > e^{\alpha(J(\tau)+\varepsilon)t}\Big) \le \frac{\mathbb{Q}_t\big(Z_t(\tau)^\alpha\big)}{e^{\alpha(J(\tau)+\varepsilon)t}}.$$

From Theorem 4.3 we know that for each $\alpha \in [0,1]$ and for all large t we have

$$\mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t}\Big) \le e^{(\alpha M(\tau)-\varepsilon)\alpha t}.$$

If we also have $\alpha \in (0, \varepsilon/M(\tau))$ then clearly this above is a decaying exponential and so it follows that

$$\lim_{t\to\infty} \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t} \Big) \to 0.$$

$\square$

For the specific y defined at (17), it can be shown that

$$J(\tau) = \int_0^\tau \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2} y_s)^2 + \frac{a\lambda^2}{2} y_s^2 - r y_s^2 \Big]\, \mathrm{d}s = \lambda\beta + \kappa^2 \Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \Big),$$

where we recall that this $\lambda \in (\lambda_{\min}, 0)$ was specifically determined by (19). In fact, Git *et al* [4] explain that this choice of λ was optimal in that

$$\lambda\beta + \kappa^2 \Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta} \coth \mu_\lambda \tau \Big) = \sup_\gamma \Big\{ \gamma\beta + \kappa^2 \Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta} \coth \mu_\gamma \tau \Big) \Big\}. \tag{31}$$

On the other hand we can find a similar representation for the parameter $\Theta(\beta,\kappa)$: if we define

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

then

$$\Theta(\beta,\kappa) = \bar{\lambda}\beta + \kappa^2 \psi^+_{\bar{\lambda}} = \sup_\gamma \Big\{ \gamma\beta + \kappa^2 \Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta} \Big) \Big\} = \lim_{\tau\to\infty} \sup_\gamma \Big\{ \gamma\beta + \kappa^2 \Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta} \coth \mu_\gamma \tau \Big) \Big\},$$

where we recall that

$$\psi^+_{\bar{\lambda}} := \frac{1}{4} + \frac{\mu_{\bar{\lambda}}}{2\theta}.$$

In this way it can be deduced from (31) that $J(\tau) > \Theta(\beta,\kappa)$ with

$$J(\tau) \downarrow \Theta(\beta,\kappa), \quad \text{as } \tau \to \infty.$$

It is now easy to deduce the following corollary to Theorem 4.4:

Corollary 4.5 *Given $\varepsilon > 0$, for $\tau > 0$ chosen sufficiently large,*

$$\lim_{t\to\infty} \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \le e^{(\Theta(\beta,\kappa)+\varepsilon)t} \Big) \to 1.$$

5 Proving the large-deviations lower bound

Barring the proof of Theorem 4.3 which we cover fully in section 7, we now have all the ingredients required to prove the large-deviations lower-bound for the short-climb event of Theorem 3.2.

Throughout this proof we are focussing on the specific path

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0,\tau]$$

where $\lambda \in (\lambda_{\min}, 0)$ satisfies

$$\frac{\beta}{a\lambda} = \kappa^2 \Big(\frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau} \Big).$$

as discussed at (19). We define the event that the space-type location $\big(X_u(s), Y_u(s)\big)$ of a particular particle $u \in N_\tau$ remains near $(a\lambda t \int_0^s y_w^2 \, \mathrm{d}w, \sqrt{t} y_s)$ throughout the interval $s \in [0, \tau]$:

$$A_t(u) := \left\{ \forall s \in [0,\tau], \big|X_u(s) - a\lambda t \int_0^s y_w^2 \, \mathrm{d}w\big| < \delta t, \big|Y_u(s) - \sqrt{t} y_s\big| < \delta' \sqrt{t} \right\},$$

where $\delta, \delta' > 0$ are given and fixed. In addition, we define the event that any of the particles performs this event (whilst emphasising the parameter dependence) by

$$A_{t,\tau}^{\delta,\delta'} := \bigcup_{u \in N(\tau)} A_t(u)$$

Noting that this event is $\mathcal{F}_\tau$-measurable since it depends only on the branching particles and does not refer to the spine, it follows that on this event the change of measure is carried out by Z_t, as noted in Theorem 4.2. The upper bound that we have derived for Z_t at Corollary 4.5 will serve as a lower bound for $1/Z_t(\tau)$ in this change of measure, and will combine with the fact that under the measure $\tilde{\mathbb{Q}}_t$ (for large t) we know that the spine will carry out the large-deviations behaviour that we want.

Then for any $\varepsilon > 0$,

$$\begin{aligned}
P\Big(A_{t,\tau}^{\delta,\delta'}\Big) &= \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u)\Big) \\
&\geq \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big) \\
&\geq e^{-(J(\tau)+\varepsilon)t}\, \mathbb{Q}_t\Big(\exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{J(\tau)+\varepsilon)t}\Big) \\
&\geq e^{-(J(\tau)+\varepsilon)t}\, \tilde{\mathbb{Q}}_t\Big(A_t(\xi); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big).
\end{aligned} \tag{32}$$

Given (25) and (26), standard theory says that under the measure $\tilde{\mathbb{Q}}_t$ (with t large) the re-scaled spine (ξ_s^t, η_s^t) will tend to stay close to the space-type paths $(a\lambda \int_0^s y_w^2 \, \mathrm{d}w, y_s)$ over the whole time interval $[0, \tau]$:

$$\xi_s^t \sim a\lambda \int_0^s y_w^2 \, \mathrm{d}w, \quad \text{and} \quad \eta_s^t \sim y_s,$$

by which we mean that for a fixed $\tau > 0$ and any $\delta, \delta' > 0$,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi_s^t - a\lambda \int_0^s y_w^2 \, \mathrm{d}w\big| < \delta, \big|\eta_s^t - y_s\big| < \delta', \text{for all } s \in [0,\tau]\Big) \to 1,$$

which can equally be written as:

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi_s - a\lambda t \int_0^s y_w^2 \, \mathrm{d}w\big| < \delta t, \big|\eta_s - y_s\sqrt{t}\big| < \delta'\sqrt{t}, \text{for all } s \in [0,\tau]\Big) \to 1,$$

equivalently,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\big(A_t(\xi)\big) = 1.$$

At the same time, Theorem 4.4 says,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big) = 1,$$

and since $\varepsilon > 0$ was arbitrary, it follows from (32) that for all fixed $\tau > 0$ and $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(A_{t,\tau}^{\delta,\delta'}\Big) \geq -J(\tau),$$

which gives the proof of the lower bound of Theorem 3.2. $\square$

6 Proving the large-deviations upper bounds

We first give a quick and direct martingale proof of the upper bound of Theorem7.2 that identifies the optimal path to reach $(\beta t, \kappa\sqrt{t})$ at large time τ. This follows along similar lines to the almost sure upper bound of section 2.1.

Let $\gamma, \kappa > 0$. Then for $\theta \in (0, -\lambda_{min})$ and $\phi \in (0, \psi_\theta^+]$ to ensure that expectations remain finite for all time, and recalling the martingale Z_θ^+ at equation (6),

$$\begin{aligned}
&P\Big(\exists u \in N_\tau : X_u(\tau) \le -\gamma t, Y_u(\tau)^2 \ge \kappa^2 t\Big)\\
&\qquad \le P\Big(\sum_{u\in N_\tau} \mathbf{1}\{X_u(\tau)+\gamma t \le 0, Y_u(\tau)^2 - \kappa^2 t \ge 0\}\Big)\\
&\qquad \le P\Big(\sum_{u\in N_\tau} e^{-\theta(X_u(\tau)+\gamma t)+\phi(Y_u(\tau)^2-\kappa^2 t)}\Big)\\
&\qquad = e^{-\theta\gamma t - \kappa^2\phi t + E_\theta^+\tau} P\Big(\sum_{u\in N_\tau} e^{-\theta X_u(\tau)+\phi Y_u(\tau)^2 - E_\theta^+\tau}\Big)\\
&\qquad \le e^{-\theta\gamma t - \kappa^2\phi t + E_\theta^+\tau} P\Big(Z_\theta^+(\tau)\Big) = e^{-(\theta\gamma+\kappa^2\phi)t + E_\theta^+\tau}.
\end{aligned}$$

Hence, taking t limits and then optimising over θ we find

$$\limsup_{t\to\infty} t^{-1}\log P\Big(\exists u \in N_\tau : X_u(\tau) \le -\gamma t, Y_u(\tau)^2 \ge \kappa^2 t\Big) \le -\sup_{\theta>0}\big\{\theta\gamma + \kappa^2\phi\big\} = \Theta(\gamma,\kappa) \quad (33)$$

which is then sufficient to give part (b) of Theorem 1.1.

Whilst this upper bound is good enough for sufficiently large τ, for fixed τ the expectation upper bound can be sharpened with a bit more work to yield part (b) of Theorem 3.2. Firstly, recalling the one particle picture and that $Z_t(w) = \sum_{u\in N(w)} f_{t,w}(u)$ we have,

$$\begin{aligned}
P\left(\exists u \in N_\tau \text{ such that } A_t(u)\right) &\le P\left(\sum_{u\in N(\tau)} \mathbf{1}\{A_t(u)\}\right)\\
&= \tilde{P}\left(e^{\int_0^\tau R(\eta_s)\,ds}; A_t(\xi)\right) = \tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)}; A_t(\xi)\right) \qquad (34)
\end{aligned}$$

Then if a good upper bound for $f_{t,\tau}(\xi)$ can be found on the event $A_t(\xi)$, a suitable lower bound for the required probability will follow.

Consider y and τ fixed and let us restrict attention to the event $A_t(\xi)$. It is now relatively straightforward making use of Itõ calculus and the constraints from the event, to consider terms in the expression for $f_{t,\tau}(\xi)$ one at at time to derive the following bounds (we omit the proof):

Proposition 6.1 *(a) There exists $K > 0$ depending only on y and τ such that, for all $w \in [0, \tau]$,*

$$\left|\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_u(s) - \frac{t}{\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\right| < \frac{t}{\theta}K\delta'$$

for almost every path in the event $A_t(\xi)$.

(b) There exists a $K' > 0$ depending only on y and τ, such that for all $w \in [0, \tau]$,

$$\left|\int_0^w \eta_s^2\,ds - t\int_0^w y_s^2 ds\right| \le t\delta'(2K' + \tau)$$

for almost every path in the event $A_t(\xi)$.

(c)

$$\tilde{\mathbb{Q}}_t\left(e^{-\sqrt{at}\lambda\int_0^\tau y_s\,dW_s}; A_t(\xi)\right) \le e^{-a\lambda^2 t\int_0^\tau y_s^2\,ds + t(\frac{1}{2}a\lambda^2\delta'^2\tau + \delta\lambda)}$$

Then,

$$\tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)};A_t(\xi)\right) =\tilde{\mathbb{Q}}_t\left(e^{\int_0^\tau R(\xi_s)\,\mathrm{d}s}\times e^{-\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s}\times e^{\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}W_s};A_t(\xi)\right)$$
$$\times e^{\frac{t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times e^{t\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s}$$
$$\leq e^{t\int_0^\tau ry_s^2\,\mathrm{d}s+\rho\tau}\times e^{t\delta'(2K'+\tau)}\times e^{-\frac{t}{\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times e^{t\frac{K}{\theta}\delta'}$$
$$\times e^{-ta\lambda^2\int_0^\tau y_s^2\,\mathrm{d}s}\times e^{t(\frac{1}{2}a\lambda^2\delta'\tau+\delta\lambda)}\times e^{-\frac{t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s+t\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s}$$

Then, given $\epsilon>0$, we can choose $\delta,\delta'>0$ sufficiently small such that

$$\tilde{\mathbb{Q}}_t\left(\frac{1}{f_{t,\tau}(\xi)};A_t(\xi)\right)\leq e^{t\left\{\int_0^\tau ry_s^2\,\mathrm{d}s+\rho\tau-\frac{1}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s+\epsilon\right\}}=e^{-tJ(\tau)+\rho\tau+\epsilon\tau}$$

□

7 A spine proof of the martingale upper-bound

In this section we use the spine decomposition of the martingale Z_t to prove Theorem 4.3. It is Jensen's inequality that immediately allows us to concentrate on the spine decomposition since

$$\mathbb{Q}_t\big(Z_t(\tau)^\alpha\big)\leq\tilde{\mathbb{Q}}_t\Big(\tilde{\mathbb{Q}}_t\big(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\big)^\alpha\Big),\qquad\text{for }\alpha\in[0,1].$$

The spine decomposition of $Z_t(\tau)$ is

$$\tilde{\mathbb{Q}}_t\big(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\big)=e^{-r\int_0^\tau\eta_s^2\,\mathrm{d}s-\rho\tau}e^{[\sqrt{a}\lambda\int_0^\tau y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]}$$
$$+\sum_{k=1}^{n_\tau}e^{-r\int_0^{S_k}\eta_s^2\,\mathrm{d}s-\rho S_k}e^{[\sqrt{a}\lambda\int_0^{S_k}y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^{S_k}y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]}.$$

We consider the two parts of this spine decomposition separately – the **spine term** and then the **sum term** – and aim to show that they both have exponential growth of the same order.

Definition 7.1 *We define*

$$\textbf{spine term}:=e^{-r\int_0^\tau\eta_s^2\,\mathrm{d}s-\rho\tau}e^{[\sqrt{a}\lambda\int_0^\tau y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]},$$

and

$$\textbf{sum term}:=$$
$$\sum_{k=1}^{n_\tau}e^{-r\int_0^{S_k}\eta_s^2\,\mathrm{d}s-\rho S_k}e^{[\sqrt{a}\lambda\int_0^{S_k}y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^{S_k}y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]}.$$

In each case we first use some martingale techniques to factor out exponential terms that give us the correct growth rate (and here we are guided by the heuristics), and then use Varadhan's lemma to show that the remaining terms do not contribute any further exponential growth. The spine term is simpler to deal with and is considered first.

7.0.1 Factoring out the spine term

Girsanov's theorem (see Øksendal [12]) states that under the new measure $\tilde{\mathbb{Q}}_t$ we have

$$\mathrm{d}B_s = \mathrm{d}\tilde{B}_s + \frac{\sqrt{t}}{\sqrt{\theta}}\Big(\dot{y}_s + \frac{\theta}{2}y_s\Big)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}W_s = \mathrm{d}\tilde{W}_s + \sqrt{at}\lambda y_s\,\mathrm{d}s, \tag{35}$$

where $\tilde{B}$ and $\tilde{W}$ are BMs under $\tilde{\mathbb{Q}}_t$, and these representations can be substituted into the spine term to give,

$$\begin{aligned}\textbf{spine term} &= e^{t\int_0^\tau \frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s-\rho\tau} \times e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\\ &= e^{tJ(\tau)-\rho\tau}\times e^{rt\int_0^\tau[(\eta_s^t)^2-y_s^2]\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.\end{aligned} \tag{36}$$

Using the standard martingale

$$e^{\alpha\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s-\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s},$$

we can factor out one of the terms of the expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big)\\ &= e^{\alpha tJ(\tau)-\alpha\rho\tau}e^{\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha r\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big).\end{aligned}$$

This final expectation can be dealt with by another change of measure:

$$\begin{aligned}&\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big)\\ &= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\Big),\\ &= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\end{aligned}$$

where we have used the martingale

$$e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau \dot{y}_s+\frac{\theta}{2}y_s\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}$$

to change the measure from $\tilde{\mathbb{Q}}_t$ to $\tilde{\mathbb{Q}}_t^\alpha$. Another application of the Girsanov theorem implies that under the measure $\tilde{\mathbb{Q}}_t^\alpha$, the re-scaled process η_s^t satisfies

$$\mathrm{d}\big(\eta_s^t-(1+\alpha)y\big) = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\bar{B}_s - \frac{\theta}{2}\big(\eta_s^t-(1+\alpha)y_s\big)\mathrm{d}s \tag{37}$$

where $\bar{B}_s$ is a Brownian motion, which is to say that η^t is an OU($\frac{\theta}{t}$, $\frac{\theta}{2}$) along the *perturbed* path $(1+\alpha)y$.

Putting this all together we are left with a neat factorization expressed in terms of the re-scaled type process η_s^t:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}e^{\alpha^2 tM(\tau)}\times\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\\ &\le e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\times\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\end{aligned} \tag{38}$$

where we remember that $M(\tau) := \int_0^\tau\Big[\frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\Big]\mathrm{d}s$. The term $\alpha\rho\tau$ becomes insignificant in the large deviations limit (for which $t\to\infty$), and therefore it is convenient to have removed it here.

The martingale techniques have now played their part, and we move on to use Varadhan's lemma to show that the term $\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big)$ decays exponentially as $t\to\infty$.

7.0.2 A first application of Varadhan's lemma

Under the measure $\tilde{\mathbb{Q}}_t^\alpha$ the process η^t is an $\mathrm{OU}(\frac{\theta}{t}, \frac{\theta}{2})$ along the perturbed path $(1+\alpha)y$ (or equivalently we can say that $[\eta^t_s - (1+\alpha)y_s]$ is an $\mathrm{OU}(\frac{\theta}{t}, \frac{\theta}{2})$), and therefore it satisfies a large-deviations principle:

Theorem 7.2 *If we use the notation η^t to refer to the element (path) in $C[0,\tau]$ defined by*

$$\eta^t(s) := \eta^t_s, \quad \text{for } s \in [0,\tau]$$

then there is a large-deviations principle for η^t with respect to the measure $\tilde{\mathbb{Q}}_t^\alpha$:

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

$$\limsup_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha(\eta^t_s \in C) \le -\inf_{g\in C} I(g,\tau),$$

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

$$\liminf_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha(\eta^t_s \in V) \ge -\inf_{g\in V} I(g,\tau),$$

where

$$I(g,w) := \int_0^w \frac{1}{2\theta}\Big[\dot{g}_s + \frac{\theta}{2} g_s - (1+\alpha)\big(\dot{y}_s + \frac{\theta}{2} y_s\big)\Big]^2 \,\mathrm{d}s.$$

if $g \in C[0,\tau]$ with $g(0)=0$ is square-integrable along with its derivative; otherwise we define $I(g) = \infty$.

Given the upper bound (38) we now want to understand the behaviour of the expectation term $\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta^t_s)^2]\,\mathrm{d}s}\Big)$ for large t. Varadhan's lemma is a common way to deal with expectations of this form, and we quote the following from Dembo and Zeitouni [2].

Theorem 7.3 (Varadhan) *Let $(X^t)_{t\ge 0}$ be a family of random variables taking values in the space $\mathcal{X}$, and let μ_t denote the probability measures associated with $(X^t)_{t\ge0}$.*

Suppose that the measures μ_t satisfy the LDP with a good rate function $I : \mathcal{X} \to [0,\infty]$, and let $\phi : \mathcal{X} \to \mathbb{R}$ be any continuous function. Assume further that the following moment condition holds for some $\gamma > 1$,

$$\limsup_{t\to\infty} t^{-1} \log \mathbb{E}\Big[e^{\gamma t \phi(X^t)}\Big] < \infty. \tag{39}$$

Then

$$\lim_{t\to\infty} t^{-1} \log \mathbb{E}\Big[e^{t\phi(X^t)}\Big] = \sup_{x\in\mathcal{X}} \big[\phi(x) - I(x)\big].$$

This powerful result will confirm our hopes that the expectation decays as $t \to \infty$.

Theorem 7.4 *For each $\alpha > 0$ the expectation decays exponentially to 0:*

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta^t_s)^2]\,\mathrm{d}s}\big) < 0. \tag{40}$$

For small α we can give more precise expression of the exponential decay:

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta^t_s)^2]\,\mathrm{d}s}\big) = \\ -\alpha^2 \Big\{ k_1 \Big[\int_0^\tau r y_s^2 \,\mathrm{d}s\Big] + k_2 \Big[\frac{1}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2} y_s\big)^2 \,\mathrm{d}s\Big]\Big\} + o(\alpha^2), \quad \text{as } \alpha \to 0,$$

where k_1, k_2 are strictly positive.

Proof: Given the large-deviations principle stated in Theorem 7.2, we shall be equating $\mathcal{X} = C[0,\tau]$, $X^t = \eta^t$ and $\mu_t = \tilde{\mathbb{Q}}_t^\alpha$ and have $\phi(\eta^t) = \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s$; the moment condition (39) is satisfied because

$$\tilde{\mathbb{Q}}_t^\alpha\big(e^{2\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big) < e^{2\alpha r t\int_0^\tau y_s^2\,\mathrm{d}s}.$$

Varadhan's lemma implies that

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big) = \sup_{z\in C_0[0,\tau]}\Big\{\Big(\int_0^\tau \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\}. \tag{41}$$

Standard Euler-Lagrange techniques for maximizing the right-hand integral lead to the following differential equation for z:

$$\ddot{z}_s - \Big(\frac{\theta^2}{4} + 2\theta\alpha r\Big)z_s = (1+\alpha)\ddot{y}_s - \frac{\theta^2}{4}(1+\alpha)y_s, \tag{42}$$

which in general will give the optimal path as a solution in terms of the given path y.

With the *specific* path (17) that resulted from the Harris and Git optimizations of the large-deviations heuristics, it is relatively simple to solve (42) and find that the optimal path z is just a constant multiple of the path y:

$$z_s = K_\alpha y_s, \qquad \text{where } K_\alpha := \frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r}(1+\alpha). \tag{43}$$

Substituting for z into (41) we find that

$$\begin{aligned}&\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big)\\ &= \alpha(1-K_\alpha^2)\Big[\int_0^\tau r y_s^2\,\mathrm{d}s\Big] - (K_\alpha - (1+\alpha))^2\Big[\frac{1}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2\,\mathrm{d}s\Big],\end{aligned} \tag{44}$$

and the following simple bound on K_α implies that this is a negative quantity

Lemma 7.5 *For all $\alpha > 0$,*

$$1 < K_\alpha < 1+\alpha. \tag{45}$$

This small lemma can be proved with simple algebra from the definition of μ_λ given at (18): we can use this to show that $\mu_\lambda^2 - \theta^2/4 = -2\theta r - a\theta\lambda^2 < 0$, from which it follows that

$$\frac{1}{1+\alpha} < \frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r} < 1.$$

If we make a Taylor expansion about $\alpha = 0$:

$$\frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r} = \frac{1}{1-k\alpha} = 1 + k\alpha + k^2\alpha^2 + o(\alpha^2) + \cdots$$

where $k := \frac{2\theta r}{\mu_\lambda^2 - \theta^2/4}$, it follows that for strictly positive constants k_1 and k_2,

$$\alpha(1-K_\alpha^2) = -k_1\alpha^2 + o(\alpha^2), \quad \text{and} \quad (K_\alpha - (1+\alpha))^2 = k_2\alpha^2 + o(\alpha^2) \qquad \text{as } \alpha\to 0,$$

completing the proof □

7.0.3 Dealing with the sum term

Focusing on the sum term, we can again substitute for $\mathrm{d}W_s$ and $\mathrm{d}B_s$ with (35) and immediately factor out the term $J(S_k)$ by over-estimating:

$$\begin{aligned}
\textbf{sum term} &= \sum_{k=1}^{n_\tau} e^{tJ(S_k)-\rho S_k} e^{r\int_0^{S_k}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{[\sqrt{at}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]} \\
&\le e^{t\left(\sup_{0\le w\le\tau} J(w)\right)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k}\eta_s^2-y_s^2\,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.
\end{aligned}$$

For the particular path y that we chose at (17), it was shown by Git *et al* [4] that

$$\sup_{0\le w\le\tau} J(w) = J(\tau)$$

and therefore we have

$$\textbf{sum term} \le e^{tJ(\tau)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k}\eta_s^2-y_s^2\,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.$$

The following small result is very useful for dealing with the sum term:

Proposition 7.6 *If $\alpha\in(0,1]$ and $u,v>0$ then $(u+v)^\alpha \le u^\alpha+v^\alpha$,*

This proposition implies that for $0\le\alpha\le1$,

$$\tilde{\mathbb{Q}}_t\big(\textbf{sum term}^\alpha\big) \le e^{\alpha tJ(\tau)}\,\tilde{\mathbb{Q}}_t\Big(\sum_{k=1}^{n_\tau} e^{\alpha rt\int_0^{S_k}[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{at}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big),$$

and we can transform the sum into an integral by standard techniques (see Kallenberg [11] for example), since the fission times on the spine form a Cox process of rate $2(r\eta_w+\rho)$, as explained in Theorem 4.1:

$$= 2e^{\alpha tJ(\tau)}\,\tilde{\mathbb{Q}}_t\bigg(\int_0^\tau e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\big[r\eta_w^2+\rho\big]\,\mathrm{d}w\bigg);$$

Fubini's theorem can be applied to this, and the transformations that worked on the spine term to give (38) can here too be applied to arrive at

$$\begin{aligned}
&= 2\,e^{\alpha tJ(\tau)}\int_0^\tau e^{\alpha^2\int_0^w\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s} e^{\frac{\alpha^2}{2\theta}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s} \times \tilde{\mathbb{Q}}_t^\alpha\Big(\big[rt(\eta_w^t)^2+\rho\big]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w, \\
&\le 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)} \times \int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\Big(\big[rt(\eta_w^t)^2+\rho\big]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w.
\end{aligned}$$

We want to take advantage of the fact that the terms in the integral look similar to those already dealt with for the spine term. A first step in this direction is to replace the random factor $rt(\eta_w^t)^2$ at the front of the expectation with the deterministic rty_w^2, and since the value of α will eventually be chosen and fixed the following estimate is sufficient for our purposes.

Lemma 7.7 *For all $\alpha>0$, and for all large enough t,*

$$\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\Big(\big[rt(\eta_w^t)^2+\rho\big]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w \le \frac{1}{\alpha} + \int_0^\tau\big[rty_w^2+\rho\big]\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w.$$

Proof: Noting that the expectation looks something like $\partial/\partial w\, \tilde{\mathbb{Q}}_t^\alpha\big(e^{-\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big)$, we shall use integration by parts. From

$$\frac{\partial}{\partial w}\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) = \tilde{\mathbb{Q}}_t^\alpha\Big(\alpha rt[y_w^2-(\eta_w^t)^2]e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big), \tag{46}$$

it follows that

$$\begin{aligned}\tilde{\mathbb{Q}}_t^\alpha\Big(\big[rt(\eta_w^t)^2+\rho\big]e^{-\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) = \big[rty_w^2+\rho\big]\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)&\\ -\frac{1}{\alpha}\frac{\partial}{\partial w}\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)&.\end{aligned}$$

Integration by parts now proves

$$\begin{aligned}\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\Big(\big[rt(\eta_w^t)^2+\rho\big]e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w = \int_0^\tau \big[rty_w^2+\rho\big]\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w&\\ +\frac{1}{\alpha}\Big[1-\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\Big]&.\end{aligned}$$

The exponential decay proved in Theorem 7.4 implies $\lim_{t\to\infty}\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)=0$, and this completes the proof □

It follows therefore that for all large enough t,

$$\tilde{\mathbb{Q}}_t\big(\mathbf{sum\ term}^\alpha\big) \le 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\times\Big(\frac{1}{\alpha}+\int_0^\tau \big[rty_w^2+\rho\big]\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w\Big).$$

We now make some simple over-estimates of the integral. Firstly, it is immediate that

$$\int_0^\tau \big[rty_w^2+\rho\big]\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w \le \big[rt\kappa^2+\rho\big]\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w$$

since $\big(\sup_{0\le w\le\tau} y_w^2\big)=\kappa^2$. Then, for each $w\in[0,\tau]$, it is true by definition that

$$e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} \le e^{\alpha rt(\sup_{0\le w\le\tau}\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s)},$$

and therefore

$$\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) \le \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\big(\sup_w\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big).$$

Since this holds for all $w\in[0,\tau]$ we can deduce

$$\sup_{0\le w\le\tau}\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) \le \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\big(\sup_w\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big),$$

which we can use to get:

$$\begin{aligned}\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)\,\mathrm{d}w &\le \tau\times\sup_{0\le w\le\tau}\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\\ &\le \tau\times\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\big(\sup_w\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big).\end{aligned}$$

Thus we arrive at a simple upper bound for the sum term: for all $\alpha\in[0,1]$ and all large t,

$$\tilde{\mathbb{Q}}_t\big(\mathbf{sum\ term}^\alpha\big) \le 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\Big\{\frac{1}{\alpha}+\big[rt\kappa^2+\rho\big]\tau\,\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\big(\sup_w\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big)\Big\}. \tag{47}$$

7.0.4 A second application of Varadhan's lemma

We already applied Varadhan's lemma to the term $\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big)$, and now we show how it can in fact deal with the more complex term $\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \sup_w \int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big)$ without much more effort.

Once again the observation

$$\tilde{\mathbb{Q}}_t^\alpha\Big(e^{2\alpha r t\big(\sup_w \int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big) < \tilde{\mathbb{Q}}_t^\alpha\Big(e^{2\alpha r t \tau(\sup_w y_w^2)}\Big)$$

shows that the moment condition (39) is satisfied and therefore from Varadhan's lemma, Theorem 7.3, it follows that:

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\big(\sup_w \int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big) = \sup_{z\in C_0[0,\tau]}\bigg\{\Big(\sup_{0\le w\le\tau}\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\bigg\}.$$

For any path z, the action functional $I(z,w)$ is non-decreasing in w and therefore

$$\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau) \le \Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w),$$

and taking the supremum over $w\in[0,\tau]$ of both sides we deduce:

$$\Big(\sup_w\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau) \le \sup_w\bigg\{\Big(\int_0^w \alpha r\big[y_s^2 z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\bigg\},$$

We now take the supremum of both sides over the set of paths $z\in C_0[0,\tau]$, and interchange the order to obtain:

$$\begin{aligned}\sup_z\bigg\{\Big(\sup_w\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\bigg\} &\le \sup_z\sup_w\bigg\{\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\bigg\} \\ &= \sup_{0\le w\le\tau}\sup_z\bigg\{\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\bigg\}.\end{aligned} \tag{48}$$

If we compare the term

$$\sup_z\bigg\{\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\bigg\}$$

with (41) from our first application of Varadhan's lemma, it is clear that Euler-Lagrange optimization techniques will result in exactly the same optimal path for this integral, namely $z_s = K_\alpha y_s$ as at (43). Furthermore, evaluating the left-hand side of (48) shows that we actually have the equality:

$$\begin{aligned}\sup_z\bigg\{\Big(\int_0^\tau \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\bigg\} &= \sup_z\bigg\{\Big(\sup_w\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\bigg\}, \\ &= \alpha(1-K_\alpha^2)\bigg[\int_0^\tau r y_s^2\,\mathrm{d}s\bigg] - (K_\alpha-(1+\alpha))^2\bigg[\frac{1}{2\theta}\int_0^\tau \big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2\,\mathrm{d}s\bigg], \\ &< 0 \quad \big(\text{and } = O(\alpha^2) \text{ as } \alpha\to 0\big).\end{aligned}$$

Consequently we see that there is no difference in the growth rate between the remaining terms of the spine term and the sum term:

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\big(\sup_w \int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big) = \lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) < 0. \tag{49}$$

7.0.5 Concluding the upper-bound for $Z_t(\tau)$

We have shown that

$$\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) \le e^{\alpha t J(\tau)} e^{\alpha^2 t M(\tau)} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big),$$

and since we clearly have $\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big) \le \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s\right)}\Big)$, it follows that:

$$\begin{aligned}\tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) &\le \tilde{\mathbb{Q}}_t\big(\textbf{spine term}^\alpha\big) + \tilde{\mathbb{Q}}_t\big(\textbf{sum term}^\alpha\big)\\ &\le e^{\alpha t J(\tau)} e^{\alpha^2 t M(\tau)} \bigg\{ \Big(1 + 2\big[rt\kappa^2 + \rho\big]\tau\Big) \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s\right)}\Big) + \frac{2}{\alpha}\bigg\}. \quad (50)\end{aligned}$$

Thus

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) \le \alpha J(\tau) + \alpha^2 M(\tau),$$

and the proof of Theorem 4.3 is completed. □

8 An alternative approach to the lower bound

Finally, we comment on an alternative approach to gaining the lower bound of Theorem 3.2 to complement our approach to the upper bound in section 6. It would not require calculation of the $\tilde{\mathbb{Q}}_t(Z_t(\tau)^\alpha)$ in order to control the size of the martingale $Z_t(s)$, but would nethertheless need some estimates requiring comparable levels of work. The key tool is nevertheless still the spine decomposition, of course!

Making use of some simple estimation, the tower property, recalling that $\mathcal{G}_\infty$ contains information only about the spine's spatial trajectory and using the conditional form of Jensen's inequality, we have

$$\begin{aligned} P\Big(\exists u \in N_\tau \text{ such that } A_t(u)\Big) &= \tilde{\mathbb{Q}}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u)\Big)\\ &\ge \tilde{\mathbb{Q}}_t\Big(\frac{1}{Z_t(\tau)}; A_t(\xi)\Big) = \tilde{\mathbb{Q}}_t\left\{\tilde{\mathbb{Q}}_t\left(\frac{1}{Z_t(\tau)}\middle|\,\mathcal{G}_\infty\right); A_t(\xi)\right\}\\ &\ge \tilde{\mathbb{Q}}_t\left\{\frac{1}{\tilde{\mathbb{Q}}_t\left(Z_t(\tau)\middle|\,\mathcal{G}_\infty\right)}; A_t(\xi)\right\}\\ &= \tilde{\mathbb{Q}}_t\left\{\frac{1}{f_{t,\tau}(\xi) + \int_0^\tau 2R(\xi_s) f_{t,s}(\xi)\,\mathrm{d}s}; A_t(\xi)\right\}. \end{aligned}$$

If a good upper bound for $f_{t,\tau}(\xi)$ and $\int_0^\tau 2R(\xi_s) f_{s,\tau}(\xi)\,\mathrm{d}s$ can be found on the event $A_t(\xi)$, a suitable lower bound for the required probability will follow. This is a similar idea as used in the upper bound approach (see equation (34)), except more work would be required here to control the integral over time (although it will still be dominated in exponential order by the final value of $f_{t,\tau}(\xi)$, matching the upper bound). This approach should work far more generally, as we hope to investigate in further work.

References

[1] Krishna B. Athreya, *Change of measures for Markov chains and the $L\log L$ theorem for branching processes*, Bernoulli **6** (2000), no. 2, 323–338. MR 2001g:60202

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

[3] János Engländer and Andreas E. Kyprianou, *Local extinction versus local exponential growth for spatial branching processes*, Ann. Probab. **32** (2004), no. 1A, 78–99. MR MR2040776

[4] Y. Git, J. W. Harris, and S. C. Harris, *Exponential growth rates in a typed branching diffusion*, Ann. App. Probab., **17** (2007), no. 2, 609-653. `doi:10.1214/105051606000000853`

[5] Hardy, R., and Harris, S. C. *A new formulation of the spine approach for branching diffusions*, `arXiv:math.PR/0611054`, (2006).

[6] Hardy, R., and Harris, S. C. *A conceptual approach to a path result for branching Brownian motion*, Stochastic Processes and their Applications, **116** (2006), no. 12, 1992–2013. `doi:10.1016/j.spa.2006.05.010`

[7] Hardy, R., and Harris, S. C. *Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales*, `arXiv:math.PR/0611056`, (2006).

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

[9] Simon C. Harris, *Convergence of a 'Gibbs-Boltzmann' random measure for a typed branching diffusion*, Séminaire de Probabilités XXXIV, Lecture Notes in mathematics, vol. 1729, Springer, 2000, pp. 133–154.

[10] J.D.Biggins, *Uniform convergence in the branching random walk*, Annals of Probability **20** (1992), 137–151.

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

[12] Bernt Øksendal, *Stochastic differential equations*, fifth ed., Springer, 2000.

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

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

[15] Andreas E. Kyprianou and A. Rahimzadeh Sani, *Martingale convergence and the functional equation in the multi-type branching random walk*, Bernoulli **7** (2001), no. 4, 593–604. MR 2002g:60069

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

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

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

[19] Peter Olofsson, *The $x \log x$ condition for general branching processes*, J. Appl. Probab. **35** (1998), no. 3, 537–544. MR 99k:60214