A spine proof of a lower bound for a typed branching diffusion

Robert Hardy and Simon C. Harris

March 8, 2004

Abstract

We follow the spine approach as found in Hardy and Harris [6, 8, 7] to define new martingales for use in spine change of measure techniques to give a simpler, more intuitive proof of an important and difficult path large-deviations lower bound for a typed branching diffusion as found in Git, J.Harris and S.C.Harris [5]. Our proof combines simple martingale ideas with applications of Varadhan's lemma, and is successful mainly because the 'spine decomposition' effectively reduces 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 [7] and, importantly, we hope to show in the future that these techniques will enable generalization to a far wider class of branching diffusion large-deviation problems.

1 Overview

Harris and Williams [9] 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 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 a lower bound for the probability of finding at least one of the branching particles far from the space-type origin, in a large-deviations sense. The precise form of the large deviation lower bound stated below was originally motivated by the work carried out in Harris and Git [10], which we briefly discuss in section 2. More recent work in Git *et al* [5] used spine techniques to give the first completely rigorous proof of this large deviation lower bound. The spine techniques we use in this paper involve a different change-of-measure on the spine to that found in Git *et al* [5] and, significantly, this new approach should facilitate far more general large-deviation results for branching diffusions, as we hope future work will reveal. Actually, the approach here naturally gives rise to a much stronger result where particles not only arrive at the space-type location $(\beta t, \kappa\sqrt{t})$ at time τ, but are known to have stayed near two specific space-type paths throughout the whole time interval $[0, \tau]$. We state this stronger result as Theorem 4.2, and here give the weaker form as it is required by the work of Harris and Git [10]:

Theorem 1.1 (The short-climb probability) *Let $\tau > 0$ be fixed, and let $\beta, \kappa \in \mathbb{R}$ be given. Then the probability that at least one of the branching particles will be near $(\beta t, \kappa\sqrt{t})$ at time τ (give that the original ancestor was at the space-type origin) has a large-deviations lower bound: for all $\delta, \delta' > 0$,*

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \big|X_u(\tau) - \beta t\big| < \delta t, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\sqrt{t}\Big) \geq -\Theta(\kappa, \beta),$$

where

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

We note that as explained by Harris and Git, and as our spine proof will make clear, this lower-bound of $-\Theta(\beta, \kappa)$ is not exactly optimal and can be marginally improved to a rate of $-J(\tau)$, a constant that we define later; but anyway we would have $-J(\tau) \downarrow -\Theta(\beta, \kappa)$ as $\tau \to \infty$, and therefore for the specific aims of Harris and Git [10] there is not a real loss in the lower bound of $-\Theta(\beta, \kappa)$. We would comment that a large-deviations upper bound is generally easier to obtain than the lower bound, and here a many-to-one theorem could be used much like for the case of branching Brownian motion in a related article by Hardy and Harris [7].

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 to arrive at the point $(\beta t, \kappa\sqrt{t})$. Our improved spine approach which we briefly lay out in section 3 (and which is fully presented in Hardy and Harris [6]) will allow us to explicitly find the Radon-Nikodym derivatives (martingales) of these new measures with respect to the original measure P, and 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.

The layout of this article is as follows: in the next section we discuss the results of Harris and Git [10] in order to give a context to our work. In section 3 we present the foundations of our spine approach, giving definitions of the underlying space and its filtrations and measures. Section 4 contains a heuristic discussion of the large deviations for the model, and can be considered as the motivating arguments behind the choice of a martingale that will be used to carry out a change of measure. In section 5 these strictly-positive martingales Z_t are defined in terms of specific paths that our heuristics will have suggested. As Radon-Nikodym derivatives, these martingales can define the new measures $\mathbb{Q}_t$ (under which they become *sub*martingales) and leaving the proof until a separate later section we go on to state a key result (Theorem 5.6) on their growth under the measures $\mathbb{Q}_t$; this growth result leads directly to the proof of the large-deviations lower bound which we present next in section 6.

The proof of the growth Theorem 5.6 is found in a separate section since it is not particularly short – but neither is it difficult as such. 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 Harris and Git almost-sure result

Before we move on to prove the above theorem, we summarize the main results from Harris and Git [10] and the earlier Harris and Williams [9] so that the reader might understand how Theorem 1.1 fits into the picture.

Work on almost-sure large-deviations results for this typed branching-diffusion began in the paper by Harris and Williams [9], where they proved that if we define a counting function

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

(not to be confused with N_t, the set of individuals alive at time t) for each $\gamma \in \mathbb{R}$, then the limit

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

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

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

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

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

The work of Harris and Git [10] aims at improving this to obtain the almost-sure rate of growth in numbers of particles at certain *spatial and type* positions at large times. They define the following function that counts how many particles occupy a particular region in the type-space domain:

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

Harris and Git's work is directed at proving the following *almost-sure* result:

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,

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

2.1 The almost-sure upper bound

Harris and Williams [9] showed that there are *two* strictly-positive martingales Z_λ^- and Z_λ^+ defined as

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

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 common theme to the upper bounds in much of the work by Harris, Git and Williams is to overestimate indicator functions by an exponential, since it is often the case that this will bring in one of the martingales of the model: for $\lambda \in (\lambda_{\min}, 0)$,

$$\begin{aligned} \sum_{k=1}^{N(t)} \mathbf{1}\big\{X_k(t) \le -\gamma t, Y_k(t)^2 \ge \kappa^2 t\big\} &\le \sum_{k=1}^{N(t)} \exp\big\{\psi_\lambda^+(Y_k(t)^2 - \kappa^2 t)\big\} \exp\big\{\lambda(X_k(t) + \gamma t)\big\} \\ &= e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{(E_\lambda^- + \lambda\gamma - \kappa^2 \psi_\lambda^+)t}. \end{aligned} \tag{3}$$

(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 (2) – explains why $\Delta(\gamma, \kappa)$ relates to (3) above: by choosing λ at the infimum we get

$$N_t(\gamma, \kappa) \le e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{\Delta(\gamma,\kappa)t}. \tag{4}$$

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

$$\limsup_{t\to\infty} e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) \le 0, \qquad \text{for each } \lambda \in (\lambda_{\min}, 0). \tag{5}$$

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 $\Delta(\gamma, \kappa) \ge 0$, (4) and (5) 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, Harris and Git 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 [9];

the short climb: Following this, over a short period of time $[t, t+\tau]$ with τ a fixed small 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 $\{\gamma\kappa^2/(\theta+\kappa^2)\}t$ in spatial position.

Harris and Git have shown that the combination of these two phases will present us with *sufficiently many* particles at the space-type position $(\gamma T, \kappa\sqrt{T})$ at the large time $T = t+\tau$, concluding their proof of Theorem 2.1 – we refer the reader to their work for further details. Theorem 1.1 that we are going to prove makes up the *short climb* phase.

3 The spine-approach foundations

Before moving on to the proofs, we briefly review the formal constructions on which our spine analysis is based – full details are laid out in Hardy and Harris [6]. The reader who is familiar with the work of Lyons *et al* [18, 13, 19] and other recent work based on these (examples are Kyprianou [14], Kyprianou and Sani [15], Athreya [1], Olofsson [22] amongst others) will notice significant differences in our approach via our use of the filtrations on the single underlying space.

The branching diffusion and its probabilities are going to be built on an underlying filtered space of sample trees with spines, which we now define in terms of Ulam-Harris notation and marked trees.

The set of Ulam-Harris labels is to be equated with the set Ω of finite sequences of strictly-positive integers:

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

where we take $\mathbb{N} = \{1, 2, \ldots\}$. For two words $u, v \in \Omega$, uv denotes the concatenated word ($u\emptyset = \emptyset u = u$), and therefore Ω contains elements like '213' (or '$\emptyset$213'), which we read as 'the individual being the 3rd child of the 1st child of the 2nd child of the initial ancestor $\emptyset$'. For two labels $v, u \in \Omega$ the notation $v < u$ means that v is an *ancestor* of u, and $|u|$ denotes the length of u. The set of all ancestors of u can be expressed in two equivalent ways:

$$\{v : v < u\} = \{v : \exists w \in \Omega \text{ such that } vw = u\}.$$

Collections of labels, ie. subsets of Ω, will therefore be groups of individuals. In particular, a subset $\tau \subset \Omega$ will be called a *Galton-Watson tree* if:

1. $\emptyset \in \tau$,
2. if $u, v \in \Omega$, then $uv \in \tau$ implies $u \in \tau$,
3. for all $u \in \tau$, $uj \in \tau$ if and only if $1 \le j \le 2$; we are supposing that each particle produces only two offspring.

The set of all Galton-Watson trees will be called $\mathbb{T}$. Typically we use the name τ for a particular tree, and whenever possible we will use the letters u or v or w to refer to the labels in τ, which we may also refer to as *nodes of* τ or *individuals in* τ or just as *particles.*

Each individual should have a *space-type location* $(X_u(t), Y_u(t)) \in \mathbb{R}\times\mathbb{R}$ at each moment of its *lifetime* $\sigma_u \in [0, \infty)$. Since a Galton-Watson tree $\tau \in \mathbb{T}$ in itself can express only the *family* structure of the individuals in the branching model, in order to give them these extra features we suppose that each individual $u \in \tau$ has a mark $((X_u, Y_u), \sigma_u)$ associated with it which we read as:

- $\sigma_u \in \mathbb{R}^+$ is the *lifetime* of u, which determines the *fission time* of particle u as $S_u := \sum_{v \leq u} \sigma_v$ (with $S_\emptyset := \sigma_\emptyset$). The times S_u may also be referred to as the *death* times;
- $X_u : [S_u - \sigma_u, S_u) \to \mathbb{R}$ gives the *space-location* and $Y_u : [S_u - \sigma_u, S_u) \to \mathbb{R}$ gives the *type-location* of u at time $t \in [S_u - \sigma_u, S_u)$.

To avoid ambiguity, it is always necessary to decide whether a particle is in existence or not at its death time.

Remark 3.1 *Our convention throughout will be that a particle u dies 'just before' its death time S_u (which explains why we have defined X_u and Y_u as paths on the open-ended interval $[S_u - \sigma_u, S_u)$). Thus at the time S_u the particle u has* disappeared*, replaced by its 2 children which are both alive and ready to go.*

We denote a single marked tree by $(\tau, (X, Y), \sigma)$ or (τ, M) for shorthand, and the set of all marked Galton-Watson trees by $\mathcal{T}$:

- $\mathcal{T} := \Big\{(\tau, X, \sigma) : \text{where } \tau \in \mathbb{T}, \sigma_u \in \mathbb{R}^+, (X_u, Y_u) : [S_u - \sigma_u, S_u) \to \mathbb{R}^2 \text{ for each } u \in \tau\Big\}.$

Each marked tree will have a subset of its particles that are alive at a given time:

- the set of particles that are alive at time t is given by $N_t := \big\{u \in \tau : S_u - \sigma_u \leq t < S_u\big\}$.

For any given marked tree $(\tau, M) \in \mathcal{T}$ we can identify distinguished lines of descent from the initial ancestor: $\emptyset, u_1, u_2, u_3, \ldots \in \tau$, in which u_3 is a child of u_2, which itself is a child of u_1 which is a child of the original ancestor $\emptyset$. We'll call such a subset of τ a *spine*, and will refer to it as ξ:

- a spine ξ is a subset of nodes $\big\{\emptyset, u_1, u_2, u_3, \ldots\big\}$ in the tree τ that make up a unique line of descent. We use ξ_t to refer to the unique node in ξ that that is alive at time t.

In a more formal definition, which can for example be found in the paper by Rouault and Liu [17], a spine is thought of as a point on $\partial\tau$ the boundary of the tree – in fact the boundary is *defined* as the set of all infinite lines of descent. This explains the notation $\xi \in \partial\tau$ in the following definition: we augment the space $\mathcal{T}$ of marked trees to become

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

It is natural to speak of the *space-type location of the spine at time t* which think of just as the space-type position of the unique node that is in the spine and alive at time t:

- we define the time-t position of the spine as $(\xi_t, \eta_t) := \big(X_u(t), Y_u(t)\big)$, where $u \in \xi \cap N_t$.

By using the notation ξ_t to refer to both the node in the tree and that node's spatial position we are introducing potential ambiguity, but in practice the context will make clear which we intend. However, in case of needing to emphasize, we shall give the node a longer name:

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

which may also be written as $\mathrm{node}_t(\xi)$. At this point it is intuitively reasonable to believe that the spine is going to behave just like any 'typical' particle in the branching diffusion: its type-diffusion η_t will be an Ornstein-Uhlenbeck process with generator (1), whilst its space-diffusion ξ_t will satisfy $\mathrm{d}\xi_t = \sqrt{a}\eta_t \mathrm{d}W_t$ for a Brownian-motion W_t.

As the spine (ξ_t, η_t) diffuses, at the fission times S_u for $u \in \xi$ it gives birth to some offspring, one of which continues the spine whilst the others go off to create subtrees like copies of the original branching diffusion. These times are especially important for the later spine decomposition of the martingale Z_λ, and we therefore give them a name:

- the sequence of random times $\{S_u : u \in \xi\}$ are known as the *fission times on the spine*;

Finally, it will later be important to know how many fission times there have been in the spine, or what is the same, to know which generation of the family tree the node ξ_t is in (where the original ancestor $\emptyset$ is considered to be the 0th generation)

Definition 3.2 *We define the counting function*

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

or equivalently,

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

which tells us which generation the spine node is in, or equivalently how many fission times there have been on the spine. For example, if $\xi = (\emptyset, u_1, u_2, u_3, \ldots)$ with $node_t(\xi) = u_2$ then both $\emptyset$ and u_1 have died and so $n_t = 2$.

The collection of all marked trees with a distinguished spine $(\hat{\tau}, \xi)$ is given the label $\tilde{\mathcal{T}}$. On this space we define four filtrations of key importance that encapsulate different knowledge, but see Hardy and Harris [6] for more precise details:

- $\mathcal{F}_t$ knows everything that has happened to all the branching particles up to the time t, *but does not know which one is the spine*;
- $\tilde{\mathcal{F}}_t$ knows everything that $\mathcal{F}_t$ knows and also knows which line of descent is the spine (it is in fact the finest filtration);
- $\mathcal{G}_t$ knows only about the spine's motion in J up to time t, but does not actually know which line of descent in the family tree makes up the spine;
- $\tilde{\mathcal{G}}_t$ knows about the spine's motion and also knows which nodes it is composed of. Furthermore it knows about the fission times of these nodes and how many children were born at each time.

Having now defined the underlying space for our probabilities, we can more clearly define the probability measures:

Definition 3.3 *For each $x \in \mathbb{R}$, let $P^{x,y}$ be the measure on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ such that the filtered probability space $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t \ge 0}, P^\cdot)$ makes the $\mathbb{R}^2$-valued point process $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$ the canonical model for the branching diffusion described in the first section.*

For details of how the measures $P^{x,y}$ are formally constructed on the underlying space of trees, we refer the reader to the work of Neveu [20] and Chauvin [3, 2].

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

$$\text{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{2}. \tag{6}$$

This observation is the key to our method for extending the measures, and for this we make use of the following representation found in Lyons [18].

Theorem 3.4 *If f is a $\tilde{\mathcal{F}}_t$-measurable function then we can write:*

$$f = \sum_{u \in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{7}$$

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

As just a simple example of this we would have

$$\eta_t^2 e^{\xi_t} = \sum_{u \in N_t} Y_u(t)^2 e^{X_u(t)} \mathbf{1}_{(\xi_t = u)},$$

which makes it clear how a spine expression can be mapped over onto the branching particles. We use this representation to extend the measures $P^{x,y}$.

Definition 3.5 *Given the measure $P^{x,y}$ on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ we extend it to the probability measure $\tilde{P}^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by defining*

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P}^{x,y} := \int_{\mathcal{T}} \sum_{u \in N_t} f_u \prod_{v < u} \frac{1}{2} \, \mathrm{d}P^{x,y}, \tag{8}$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (7).

The previous approach to spines, exemplified in Lyons [18], used the idea of *fibres* to get a measure analogous to our $\tilde{P}$ that could measure the spine. However, a weakness in this approach was that the corresponding measure did not have a finite mass and therefore could not be normalized to become a probability measure like our $\tilde{P}$. Our new idea of using the down-weighting term of (6) in the definition of $\tilde{P}$ is crucial in ensuring that we do not get an infinite-mass measure, and leads to the very useful situation in which *all* measure changes in our formulation are carried out by *martingales*.

Theorem 3.6 *This measure $\tilde{P}^{x,y}$ really is an extension of $P^{x,y}$, in that $P = \tilde{P}|_{\mathcal{F}_\infty}$.*

Proof: If $f \in m\mathcal{F}_t$ then the representation (7) is trivial and therefore by definition

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

However, it can be shown that $\sum_{u \in N_t} \prod_{v<u} \frac{1}{2} = 1$ by retracing the sum back through the lines of ancestors to the original ancestor $\emptyset$, factoring out the product terms as each generation is passed. Thus

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

$\square$

The above foundations to the spine approach are laid out in a more general form in Hardy and Harris [6] where the reader will also find other interesting results that follow from this new formalization using different filtrations rather than using marginalizing as was the case in Lyons [18] for example.

4 Large deviations heuristics

We suppose that the number $\tau > 0$ is given and fixed; all our large-deviations results in this chapter will be considered as occurring over the fixed time-interval $[0, \tau]$. We here present some arguments concerning the large-deviations behaviour of the branching diffusion which the reader should take just as the *intuition* behind our later *rigourous* proofs.

By definition, under the measure $\tilde{P}$ $(= \tilde{P}^{0,0})$ the *spine* (ξ_s, η_s) satisfies

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

for two independent $\tilde{P}$-Brownian motions B_s and W_s. 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 4.1 *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 fixed time-interval $[0, \tau]$.

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

$$\exp\Big(-\frac{t}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2\,\mathrm{d}s - \frac{t}{2}\int_0^\tau \frac{\dot{x}_s^2}{ay_s^2}\,\mathrm{d}s\Big), \tag{9}$$

for large enough t.

The reader who is familiar with the large-deviations principle for branching Brownian motion (see Hardy and Harris [7] for a spine proof, or Lee [16] for a classical proof) would make the reasonable guess that the probability that *at least one* of the *re-scaled* branching particles $\big(X_u(s)/t, Y_u(s)/\sqrt{t}\big)$ follows the type-path y_s and space-path x_s closely over the time interval $[0, \tau]$ is roughly

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

when t is large. By standard optimization arguments (Harris and Git [10] 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

$$\exp\Big\{-\inf_{x,y}\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2 + \frac{1}{2}\frac{\dot{x}_s^2}{ay_s^2} - ry_s^2\,\mathrm{d}s\Big)t - \rho w\Big]\Big\}, \tag{10}$$

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

$$y(0) = 0, y(\tau) = \kappa, x(0) = 0, x(\tau) = \beta. \tag{11}$$

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 (10) – 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.

Harris and Git [10] state that for any given type-path y, the optimal space-path x for (10) under the constraint $x(\tau) = \beta$ will always be given by

$$x_s = \lambda\int_0^s ay_w^2\,\mathrm{d}w, \qquad \text{for } s \in [0, \tau], \tag{12}$$

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 (12). 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:

$$\lambda = \frac{\beta}{a\int_0^\tau y_s^2\,\mathrm{d}s}. \tag{13}$$

Thus for the event being considered in Theorem 1.1, the optimal spatial-path x is determined *uniquely* by (12) together with (13). 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

$$\exp\Big\{-\inf_{y}\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2 + \frac{a\lambda^2}{2}y_s^2 - ry_s^2\,\mathrm{d}s\Big)t - \rho w\Big]\Big\}, \tag{14}$$

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

$$y(0) = 0, y(\tau) = \kappa, \quad \lambda := \frac{\beta}{a \int_0^\tau y_s^2 \, \mathrm{d}s}. \tag{15}$$

Harris and Git [10] presented alternative heuristic arguments based on birth-death processes to arrive at the expression (14). Using Euler-Lagrange techniques they showed that the specific path

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0, \tau] \tag{16}$$

is optimal for this expression, where

$$\mu := \mu_\lambda = \frac{\sqrt{\theta(\theta - 8r - 4a\lambda^2)}}{2}, \tag{17}$$

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

$$\frac{\beta}{a\lambda} = \kappa^2 \Big(\frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2 \sinh^2 \mu_\lambda \tau} \Big). \tag{18}$$

We refer the reader to Harris and Git [10] 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 (16) together with x_s defined at (12), 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 would actually follow as a corollary.

Theorem 4.2 *Let $\tau > 0$ be fixed. For any $\kappa \in \mathbb{R}$ and $\lambda \in (\lambda_{\min}, 0)$ we define two continuous paths on $[0, \tau]$,*

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad x_s := a\lambda \int_0^s y_w^2 \, \mathrm{d}w, \qquad s \in [0, \tau], \tag{19}$$

and note that at time τ these paths reach the points $y_\tau = \kappa$ and $x_\tau = \beta$ where

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

Then, for large t, the probability that at least one of the typed branching particles $(X_u(s), Y_u(s))$ will stay near $(tx_s, \sqrt{t} y_s)$ throughout $s \in [0, \tau]$ (given that the original ancestor was at the space-type origin) has a large-deviations lower bound: for all $\delta, \delta' > 0$,

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

We remark that it should be possible to develop the ideas that we use in this article to obtain proofs of large-deviations principles for many branching-diffusion models (see Harris and Git [10] for a discussion of such principles), 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.

5 Martingales and measures

Although we have already indicated a specific path at (16), 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 (16). 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.

Estimates on a martingale can give us a lower-bound for the large-deviations events of our branching diffusion, and the expressions that we considered in the heuristics of the previous section can be taken as a starting point. 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 (14) above.

The process n_w (defined at Definition 3.2) 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 a $\tilde{P}$-martingale too (a similar martingale was used by Kyprianou [14] for BBM). We use the product of these two martingales to define a new measure:

Theorem 5.1 *For $t > 0$ we define a measure $\tilde{\mathbb{Q}}_t$ on the filtered space $\big(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\tau, (\tilde{\mathcal{F}}_w)_{w\in[0,\tau]}\big)$*

$$\frac{\mathrm{d}\tilde{\mathbb{Q}}_t}{\mathrm{d}\tilde{P}}\bigg|_{\tilde{\mathcal{F}}_w} := e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w} \times \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_s - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big) \\ \times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big). \quad (20)$$

Under the measure $\tilde{\mathbb{Q}}_t$ we can give a pathwise construction of the branching-diffusion $(\mathbb{X}_s)_{s\in[0,\tau]}$:

- *the spine process (ξ_s, η_s) starts at $(0,0)$ and diffuses as a solution to*

$$\mathrm{d}(\eta_s^t - y_s) = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\tilde{B}_s - \frac{\theta}{2}\big(\eta_s^t - y_s\big)\mathrm{d}s \quad (21)$$

  and

$$\mathrm{d}\xi_s^t = \frac{\sqrt{a}}{\sqrt{t}}\eta_s^t\,\mathrm{d}\tilde{W}_s + a\lambda y_s \eta_s^t\,\mathrm{d}s, \quad (22)$$

  so that the type process η_s^t will be an $OU(\theta/t, \theta/2)$ along the path y, and the spatial motion ξ_s has a drift component added;

- *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 $R(\cdot)$.*

We give the name $\tilde{\zeta}_t(w)$ to this $(\tilde{\mathcal{F}}_t, \tilde{P})$-martingale defined at (20).

Because of 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.

Definition 5.2 *We define a measure on $\mathcal{F}_\infty$:*

$$\mathbb{Q}_t := \tilde{\mathbb{Q}}_t|_{\mathcal{F}_\infty}.$$

Under $\mathbb{Q}_t$ the above construction still holds for the branching diffusion, the only difference being that we cannot use it to measure the spine – one of the particles in the branching diffusion is still the spine, but it is not known which. A very useful consequence of our new approach with filtrations is that we can directly use the $\tilde{\mathbb{Q}}_t$-martingale $\tilde{\zeta}_t$ to define a $\mathbb{Q}_t$-martingale:

Definition 5.3 *For each $t > 0$ we define a $\big((\mathcal{F}_w)_{0\le w\le\tau}, P\big)$-martingale as*

$$Z_t(w) = \tilde{P}\big(\tilde{\zeta}_t(w)|\mathcal{F}_w\big), \quad \textit{for } w \in [0,\tau].$$

It is trivial to check that such a conditional-expectation operation produces a martingale. Using the representation (7) we can explicitly determine the form of this martingale.

Theorem 5.4 *For each $w \in [0,\tau]$,*

$$Z_t(w) = \sum_{u\in N_w} e^{-r\int_0^w Y_u(s)^2\,\mathrm{d}s-\rho w}\exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (y_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_u(s) - \frac{t}{2\theta}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big)$$
$$\times\exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_u(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big), \quad (23)$$

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

Proof: Bearing in mind that $2^{n_w} = \prod_{v<\xi_w} 2$, if we use the representation (7) we get

$$\tilde{\zeta}_t(w) = \sum_{u\in N_t} e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}e^{G_u(w;t)}e^{H_u(w;t)} \times \prod_{v<u} 2\times \mathbf{1}_{(\xi_w=u)},$$

where for shorthand we define, for any $u\in N_w$ and $w\in[0,\tau]$, and any $t>0$:

$$G_u(w;t) := \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (y_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_u(s) - \frac{t}{2\theta}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s\Big),$$
$$H_u(w;t) := \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_u(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big).$$

Then because G_u and H_u are $\mathcal{F}_w$-measurable we have

$$\begin{aligned}\tilde{P}\big(\tilde{\zeta}_t(w)|\mathcal{F}_w\big) &= \sum_{u\in N_t} e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}e^{G_u(w)}e^{H_u(w)}\times\prod_{v<u}2\times\tilde{P}\big(\xi_w=u|\mathcal{F}_w\big)\\ &= \sum_{u\in N_t} e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}e^{G_u(w)}e^{H_u(w)}\times\prod_{v<u}2\times\prod_{v<u}\frac{1}{2}\\ &= \sum_{u\in N_t} e^{-\int_0^w R(Y_u(s))\,\mathrm{d}s}e^{G_u(w;t)}e^{H_u(w;t)} = Z_t(w).\end{aligned}$$

In the above we used $\tilde{P}\big(\xi_t = u|\mathcal{F}_t\big) = \prod_{v<u} \frac{1}{2}$, which is a direct consequence of the fact that the spine is chosen uniformly, as mentioned at (6). $\square$

The conditional-expectation relationship between $\tilde{\zeta}_t$ and Z_t gives us the *very important* result that the martingale Z_t is the Radon-Nikodym derivative between the branching-particle measures $\mathbb{Q}_t$ and P:

Theorem 5.5 *For each* $w \in [0, \tau]$*,*

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

Proof: The definition of the measures $\mathbb{Q}_t$ and $\tilde{\mathbb{Q}}_t$ says that for $F \in \mathcal{F}_w$,

$$\mathbb{Q}_t(F) = \tilde{\mathbb{Q}}_t(F) = \tilde{P}\big(\tilde{\zeta}_t(w)\mathbf{1}_F\big),$$

and from the definition of the conditional expectation we have

$$\tilde{P}\big(\tilde{\zeta}_t(w)\mathbf{1}_F\big) = \tilde{P}\big(Z_t(w)\mathbf{1}_F\big).$$

Thus $\mathbb{Q}_t(F) = \tilde{P}\big(Z_t(w)\mathbf{1}_F\big) = P\big(Z_t(w)\mathbf{1}_F\big)$, which concludes the proof. $\square$

For our proof of Theorem 1.1 (and its stronger version of Theorem 4.2) it is important to know how quickly $Z_t(\tau)$ grows under the measure $\mathbb{Q}_t$. The following result is the main application of spines in this article:

Theorem 5.6 *For the specific path* y *defined at* (16)*, 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

$$J(w) := \int_0^w \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2 - ry_s^2\Big]\,\mathrm{d}s, \tag{24}$$

and

$$M(w) := \int_0^w \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2\Big]\,\mathrm{d}s. \tag{25}$$

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 and Git [9] for their simpler martingale Z_λ, where their classical approach succeeded mainly thanks to the (ad-hoc) inequality of Neveu [21] and the fact that they had a relatively accessible $\mathcal{L}^2$-theory available. It is notoriously difficult to deal with operations like $Z_t(\tau)^\alpha$ since these martingales Z_t are defined via *sums*, and classical inequalities will tend to not be 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 5.6, we can therefore use Doob's submartingale inequality to prove the following:

Theorem 5.7 *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 5.6 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 (16), it can be shown that

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

where we recall that this $\lambda \in (\lambda_{\min}, 0)$ was specifically determined by (18). In fact, Harris and Git [10] explain that this choice of λ was optimal in that

$$\lambda\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth\mu_\lambda\tau\Big) = \sup_\gamma\Big\{\gamma\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta}\coth\mu_\gamma\tau\Big)\Big\}. \tag{26}$$

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

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

where we recall that

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

In this way it can be deduced from (26) that

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

which in turn more than implies $J(\tau) < \Theta(\beta,\kappa)$, and therefore gives the following corollary to Theorem 5.7:

Corollary 5.8 *Let $\tau > 0$ be fixed. Then for all $\varepsilon > 0$,*

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

6 Proving the large-deviations lower bound

Barring the proof of Theorem 5.6 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 1.1, which we recall as stating that for large t, the probability that at least one of the branching particles will be near $(\beta t, \kappa\sqrt{t})$ at time τ (where $\tau, \beta, \kappa \in \mathbb{R}$ with $\beta < 0$ and $\tau > 0$ are given and fixed) has a large-deviations lower bound: for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \big|X_u(\tau) - \beta t\big| < \delta t, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\sqrt{t}\Big) \geq -\Theta(\beta,\kappa).$$

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 5.5. The upper bound that we have derived for Z_t at Corollary 5.8 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.

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 (18). 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 $s \in [0,\tau]$:

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

where $\delta, \delta' > 0$ are given and fixed. Then for any $\varepsilon > 0$,

$$\begin{aligned} P_t\Big(\exists u \in N_\tau \text{ such that } A_t(u)\Big) &= \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u)\Big) \\ &\geq \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) \\ &\geq e^{-(\Theta(\beta,\kappa)+\varepsilon)t}\, \mathbb{Q}_t\Big(\exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) \\ &\geq e^{-(\Theta(\beta,\kappa)+\varepsilon)t}\, \tilde{\mathbb{Q}}_t\Big(A_t(\xi); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big). \end{aligned} \tag{27}$$

Given (21) and (22), 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:

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

which is exactly the statement that

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

At the same time, Corollary 5.8 says,

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

and since $\varepsilon > 0$ was arbitrary, it follows from (27) that for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1}\log P\Big(\forall s\in[0,\tau], \big|X_u(s) - a\lambda t\int_0^s y_w^2\,\mathrm{d}w\big| < \delta, \big|Y_u(s) - \sqrt{t}y_s\big| < \delta'\Big) \ge -\Theta(\beta,\kappa),$$

which gives the proof of the stronger version at Theorem 4.2. The constraints of (15) state that for the particular y and λ we chose above, we have

$$\sqrt{t}y(\tau) = \kappa\sqrt{t}, \quad \text{and} \quad a\lambda t\int_0^\tau y_s^2\,\mathrm{d}w = \beta t,$$

and therefore we can also deduce the weaker result to complete our proof of Theorem 1.1:

$$\liminf_{t\to\infty} t^{-1}\log P\Big(\exists u\in N_\tau, \big|X_u(\tau) - \beta t\big| < \delta, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\Big) \ge -\Theta(\beta,\kappa).$$

This completes the proof of the short-climb large-deviations lower bound. □

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 5.6. It is Jensen's inequality that immediately allows us to concentrate on the spine decomposition:

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

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

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

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

$$\mathrm{d}B_s = \mathrm{d}\tilde{B}_s + \frac{\sqrt{t}}{\sqrt{\theta}}\Big(\dot{y}_s + \frac{\theta}{2}y_s\Big)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}W_s = \mathrm{d}\tilde{W}_s + \sqrt{at}\lambda y_s\,\mathrm{d}s, \tag{28}$$

which can both be substituted into the spine term to give,

$$\begin{aligned}\textbf{spine term} &= e^{t\int_0^\tau \frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s-\rho\tau} \times e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]} \\ &= e^{tJ(\tau)-\rho\tau} \times e^{rt\int_0^\tau[(\eta_s^t)^2-y_s^2]\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.\end{aligned} \tag{29}$$

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 (where $\bar{B}_s$ is a Brownian motion)

$$\mathrm{d}\big(\eta_s^t-(1+\alpha)y\big) = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\bar{B}_s - \frac{\theta}{2}\big(\eta_s^t-(1+\alpha)y_s\big)\mathrm{d}s \tag{30}$$

which is to say that η^t is an $\mathrm{OU}(\frac{\theta}{t},\frac{\theta}{2})$ 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:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}e^{\alpha^2 tM(\tau)} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big), \\ &\le e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\end{aligned} \tag{31}$$

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 $[\eta_s^t - (1+\alpha)y_s]$ 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*

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

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*

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

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

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

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 (31) we now want to understand the behaviour of the expectation term $\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t\int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big)$ 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 [4].

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 $(X^t)_{t\ge 0}$.*

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 $\gamma > 1$,

$$\limsup_{t\to\infty} t^{-1} \log \mathbb{E}\Big[e^{\gamma t \phi(X^t)}\Big] < \infty. \tag{32}$$

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:*

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big) < 0. \tag{33}$$

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

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

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 (32) 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

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big) = \sup_{z\in C_0[0,\tau]}\Big\{\Big(\int_0^\tau \alpha r\big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\}. \tag{34}$$

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

$$\ddot{z}_s - \Big(\frac{\theta^2}{4} + 2\theta\alpha r\Big)z_s = (1+\alpha)\ddot{y}_s - \frac{\theta^2}{4}(1+\alpha)y_s, \tag{35}$$

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

With the *specific* path (16) that resulted from the Harris and Git optimizations of the large-deviations heuristics, it is relatively simple to solve (35) and find that the optimal path z is just a constant multiple of the path y:

$$z_s = K_\alpha y_s, \qquad \text{where } K_\alpha := \frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r}(1+\alpha). \tag{36}$$

Substituting for z into (34) we find that

$$\begin{aligned}&\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\big)\\ &= \alpha(1-K_\alpha^2)\Big[\int_0^\tau r y_s^2\,\mathrm{d}s\Big] - (K_\alpha - (1+\alpha))^2\Big[\frac{1}{2\theta}\int_0^\tau \big(\dot{y}_s + \frac{\theta}{2}y_s\big)^2\,\mathrm{d}s\Big],\end{aligned} \tag{37}$$

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

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

$$1 < K_\alpha < 1+\alpha. \tag{38}$$

This small lemma can be proved with simple algebra from the definition of μ_λ given at (17): 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 $\square$

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 (28) and immediately factor out the term $J(S_k)$ by over-estimating:

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

For the particular path y that we chose at (16), it was shown by Harris and Git [10] that

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

and therefore we have

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

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 $(u+v)^\alpha\leq u^\alpha+v^\alpha$,*

This proposition implies that for $0\leq\alpha\leq 1$,

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

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 5.1:

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

Fubini's theorem can be applied to this, and the transformations that worked on the spine term to give (31) can here too be applied to arrive at

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

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,*

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

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

$$\frac{\partial}{\partial w}\,\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) = \tilde{\mathbb{Q}}_t^\alpha\Big(\alpha rt[y_w^2-(\eta_w^t)^2]e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big), \tag{39}$$

it follows that

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

Integration by parts now proves

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

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,

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

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

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

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

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

which we can use to get:

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

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

$$\tilde{\mathbb{Q}}_t\big(\textbf{sum term}^\alpha\big) \le 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM(\tau)}\bigg\{\frac{1}{\alpha}+\big[rt\kappa^2+\rho\big]\tau\,\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\left(\sup_w\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\Big)\bigg\}. \tag{40}$$

7.0.4 A second application of Varadhan's lemma

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

Once again the observation

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

shows that the moment condition (32) is satisfied and therefore from Varadhan's lemma, Theorem 7.3, it follows that:

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

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:

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

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:

$$\begin{aligned}\sup_z\Big\{\Big(\sup_w\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\Big\} &\le \sup_z\sup_w\Big\{\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\Big\} \qquad (41)\\ &= \sup_{0\le w\le\tau}\sup_z\Big\{\Big(\int_0^w \alpha r\big[y_s^2-z_s^2\big]\,\mathrm{d}s\Big) - I(z,w)\Big\}.\end{aligned}$$

If we compare the term

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

with (34) 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 (36). Furthermore, evaluating the left-hand side of (41) shows that we actually have the equality:

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

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

$$\lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}^\alpha_t\Big(e^{\alpha r t\big(\sup_w \int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\big)}\Big) = \lim_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}^\alpha_t\Big(e^{\alpha r t\int_0^\tau [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big) < 0. \qquad (42)$$

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

$$\begin{aligned}\tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) &\le \tilde{\mathbb{Q}}_t\big(\textbf{spine term}^\alpha\big) + \tilde{\mathbb{Q}}_t\big(\textbf{sum term}^\alpha\big) \\ &\le e^{\alpha t J(\tau)} e^{\alpha^2 t M(\tau)} \bigg\{ \Big(1 + 2\big[rt\kappa^2 + \rho\big]\tau\Big) \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \big(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s\big)}\Big) + \frac{2}{\alpha}\bigg\}. \qquad (43)\end{aligned}$$

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 5.6 is completed. □

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

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

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

[5] Y. Git, J.W.Harris, and S.C.Harris, *Exponential growth rates & large deviations for a typed branching diffusion*, (2004), no. 0406, Mathematics Preprint, University of Bath.

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

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

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

[9] 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

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

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

[18] 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

[19] 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

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

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

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