

# From stochastic individual-based models to free-boundary Hamilton-Jacobi equations

Nicolas Champagnat\*   Sylvie Méléard†   Sepideh Mirrahimi‡   Viet Chi Tran §

April 20, 2026

## Abstract

We study a stochastic branching model for a population structured by a quantitative phenotypic trait and subject to births, deaths, and mutations. In a regime of large population and small mutations, and in logarithmic scales of size and time, we derive a certain class of free boundary Hamilton-Jacobi equations with state constraints from the stochastic individual-based system. This goes beyond the classical Hamilton-Jacobi equations obtained from deterministic models by taking into account the possible extinction of the system in certain regions of the trait space. The proof is obtained by combining methods for the analysis of Hamilton-Jacobi equations with probabilistic tools from the theory of large deviations and branching processes.

Keywords: stochastic birth-death model, measure-valued process, large population approximation, mutation, Hamilton-Jacobi equations, large deviations.

MSC 2000 subject classification: 92D25, 92D15, 60J80, 60F99, 35F21.

## 1 Introduction

In the mathematical modeling of eco-evolutionary dynamics of phenotypically structured populations with mutation and selection, several methods have been developed over the past two decades, based either on stochastic or deterministic approaches. Among these approaches, some of them aim to describe long-term evolutionary dynamics of dominant evolutionary paths in large populations, based either on stochastic processes or deterministic equations.

The stochastic approach developed in [12, 14] was based on birth and death measure-valued processes [31, 13] involving mutation and competition. The mutation time scale was assumed much slower than that of demographic events and, using slow-fast time scales arguments, some

---

\*Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France; E-mail: [nicolas.champagnat@inria.fr](mailto:nicolas.champagnat@inria.fr)

†Ecole Polytechnique, CNRS, Institut polytechnique de Paris, Inria, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: [sylvie.meleard@polytechnique.edu](mailto:sylvie.meleard@polytechnique.edu)

‡Univ Toulouse, INSA Toulouse, CNRS, IMT, Toulouse, France; E-mail: [sepideh.mirrahimi@math.univ-toulouse.fr](mailto:sepideh.mirrahimi@math.univ-toulouse.fr)

§Univ. Lille, CNRS, Inria, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France; E-mail: [chi.tran@inria.fr](mailto:chi.tran@inria.fr)

dominant evolutionary dynamics were highlighted in the time scale of mutations. However, the hypothesis of very rare mutations, which may be reasonable for certain phenotypes, may seem unrealistic for others, as may be the very long timescale required to describe evolutionary phenomena [48].

At the same time, an approach involving Hamilton-Jacobi equations has been developed to characterize the evolutionary dynamics [20, 5], as limits of integro-differential selection-mutation equations under the assumption of small mutation effect and long time, and using a Hopf-Cole transformation. These integro-differential equations themselves have been derived as large population limits of stochastic individual-based models [31, 13]. The order in which these limiting procedures are applied can influence the outcome, since they do not always commute, potentially leading to some inaccuracies in the deterministic approximations. In particular, the large population approximation prevents phenomena of local extinction, which may lead to an overestimation of the speed of evolution, as this may be slowed down by certain extinction events [45].

More recently [22, 9, 15, 17, 23], stochastic birth and death mutation-selection models were studied under different parameter scalings, with rare mutations, but not so rare as in the previous stochastic approach. They were inspired by some stochastic Hopf-Cole transformation, making them much closer to the deterministic models described above.

In this work, following our seminal work [16] (see also [36]), we reconcile the stochastic and deterministic approaches by deriving a new class of state-constrained Hamilton-Jacobi equations directly from the stochastic system in the time and size scales of [15] but assuming small mutations instead of rare mutations. In [16], we studied the case of a uniformly supercritical branching population with discretized and compact state space using a direct convergence approach and deterministic methods on Hamilton-Jacobi equations [5]. In [36], this result was extended to weaker assumptions in the supercritical case and new results were obtained in the subcritical case, allowing extinction of the population. Here we develop a new and robust mathematical strategy to study the general case where the growth rate can change sign, without requiring a discretized trait space, and where the restrictive assumptions required in [16, 36] are relaxed. This leads to a new limiting object which takes into account the effect of demographic stochasticity and possible extinction of small subpopulations. This approach yields more accurate and informative results with regard to modeling. In particular, the speed of evolution can be slowed down compared to classical Hamilton Jacobi approaches, making this model relevant for understanding the evolutionary dynamics of advantageous traits.

Our main result is based on the variational representation of the solution of the Hamilton-Jacobi equations and on exponential deviations results for branching processes. We were inspired by an analytical approach proposed in [43] and by the study of branching processes in [6, 40].

In Section 1, we develop the model and state our main results (Theorems 1.1 and 1.2), that is large deviations estimates for historical processes on a logarithmic time scale and on the path space, characterized by a variational expression. The latter is related to the solution of state constrained Hamilton-Jacobi equations (Theorem 1.4 and Corollary 1.7). In Section 1.3, we compare our results with existing work in different contexts, all seeking to characterize optimal trajectories in one way or another. Based on Feynman-Kac formula for branching processes, we study in Section 2 the underlying Markovian mutational process for which large deviation results can be easily obtained. We complete these results by proving some uniformity with respect to

the initial condition in the large deviation principle. Sections 3, 4 and 5 are devoted to the proofs of Theorems 1.1 and 1.2. The proof of the lower bound (Theorem 1.2) is more difficult and we use moment methods developed for branching processes as in [40] to control the number of particles whose paths are in a given tube, following the strategy used in [6]. Section 6 studies the link between the variational formulation of the limit and the Hamilton-Jacobi equation.

Notation : The space of finite measures on  $\mathbb{R}$  is denoted by  $\mathcal{M}_F(\mathbb{R})$  and the set of finite point measures on  $\mathbb{R}$  by  $\mathcal{M}_P(\mathbb{R})$ . For a measure  $\mu$  and a  $\mu$ -integrable or positive function  $f$ , we write  $\int_{\mathbb{R}} f(x)\mu(dx) = \langle \mu, f \rangle$ .

For a Polish space  $E$  (in particular for  $E = \mathbb{R}$  or  $E = \mathcal{M}_F(\mathbb{R})$ ), we will denote by  $\mathbb{D}([0, t], E)$  the Skorohod space of right-continuous and left-limited (càdlàg) functions from  $[0, t]$  to  $E$ . We will sometimes use the notation  $\mathbb{D}[0, t]$  or  $\mathbb{D}$  instead of  $\mathbb{D}([0, t], \mathbb{R})$  and  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$  respectively. The space  $\mathbb{D}[0, t]$  is equipped with the Skorohod distance  $d_{\text{Sko}}$  that makes this space Polish, see e.g. [8]: for  $f, g \in \mathbb{D}[0, t]$ ,

$$d_{\text{Sko}}(f, g) = \inf_{\lambda} \left\{ \sup_{s \in [0, t]} |f \circ \lambda(s) - g(s)| + \sup_{s \neq r} \left| \log \frac{\lambda(s) - \lambda(r)}{s - r} \right| \right\}, \quad (1.1)$$

where the supremum is taken over all increasing homeomorphisms of  $[0, t]$ . We denote by  $C[0, t]$  the set of continuous functions from  $[0, t]$  to  $\mathbb{R}$ , by  $C^\infty[0, t]$  the set of infinitely differentiable real functions on  $[0, t]$  and by  $AC[0, t]$  the set of absolutely continuous real functions on  $[0, t]$ .

## 1.1 Model

We consider a stochastic birth-death-mutation process describing an asexual population of individuals (for example, cells or bacteria), characterized by a quantitative trait  $x$  belonging to  $\mathbb{R}$ . Notice that for better legibility we restrict our work to  $x \in \mathbb{R}$ . However, our results can easily be extended to  $x \in \mathbb{R}^d$ . We introduce a parameter  $K \in \mathbb{N}$  scaling the initial population size and the mutation amplitude. We assume that an individual with trait  $x \in \mathbb{R}$  undergoes the following events independently from the other individuals: this individual can

- give birth to a new individual with the same trait  $x$  at rate  $b(x)$ ;
- die at rate  $d(x)$ ;
- give birth to a mutant individual at rate  $p(x)$  and the mutant trait is given by  $x + \frac{Y}{\log K}$  where  $Y$  is distributed as  $G(y)dy$  where  $G$  is a probability density function.

We assume that  $G : \mathbb{R} \rightarrow \mathbb{R}_+^*$  is positive continuous, that it is an even function and that it has all its exponential moments finite.

We also assume that the functions  $b, d, p$  are nonnegative locally Lipschitz-continuous bounded functions, and that there exist positive constants  $\bar{b}, \bar{p}$  and  $\underline{p}$  such that for all  $x \in \mathbb{R}$ ,

$$0 \leq b(x) \leq \bar{b}, \quad \underline{p} \leq p(x) \leq \bar{p}. \quad (1.2)$$

In particular, there exists a positive constant  $\bar{R}$  such that the growth rate satisfies for all  $x \in \mathbb{R}$

$$R(x) := b(x) + p(x) - d(x) \leq \bar{R}. \quad (1.3)$$

In the model, we index individuals using the Ulam-Harris-Neveu numbering. The set of labels is

$$\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^* \times \{0, 1\}^n),$$

where we use the usual notation  $v_1 v_2 \dots v_n$  for the vector  $(v_1, v_2, \dots, v_n) \in \mathcal{U}$ . Individuals initially present in the population form the first generation and are labeled by integers from 1 to  $N_0^K$ , the initial number of individuals. When an individual of label  $v \in \mathcal{U}$  reproduces, we consider that it acquires the label  $v0$  and its offspring is given the label  $v1$ .

The population dynamics at scale  $K$  is the point measure-valued process  $(Z_t^K)_{t \geq 0}$  defined for each  $t \geq 0$  by

$$Z_t^K = \sum_{v \in V_t^K} \delta_{X_t^{K,v}},$$

where  $V_t^K$  denotes the labels of individuals alive at time  $t$  and each individual  $v$  alive at time  $t$  has the trait  $X_t^{K,v}$ . In the sequel, we will sometimes also denote by  $V_{[0,t]}^K$  the set of labels for individuals born before time  $t$  (including those still alive at time  $t$ ), i.e.  $V_{[0,t]}^K = \bigcup_{s \in [0,t]} V_s^K$ .

We assume that

$$Z_0^K \text{ is a Poisson point measure on } \mathbb{R} \text{ with intensity measure } K^{\beta_0^K(x)} dx, \quad (1.4)$$

where for any  $K$ ,  $\beta_0^K$  is a continuous function on  $\mathbb{R}$  which converges uniformly (as  $K$  tends to infinity) to a function  $\bar{\beta}_0$  such that there exist constants  $\bar{\beta}$  and  $\alpha > 0$  such that

$$\bar{\beta}_0 \text{ is locally Lipschitz on } \mathbb{R} \text{ and } \beta_0^K(x) \leq \bar{\beta} - \alpha|x|, \quad \forall x \in \mathbb{R}. \quad (1.5)$$

This implies in particular that the intensity measure  $K^{\beta_0^K(x)} dx$  is finite for any  $K$  and hence that the initial number of individuals  $N_0^K = \langle Z_0^K, 1 \rangle$  is almost surely finite, i.e.  $Z_0^K \in \mathcal{M}_P(\mathbb{R})$ , since

$$E(N_0^K) = \mathbb{E}(\langle Z_0^K, 1 \rangle) = \int_{\mathbb{R}} K^{\beta_0^K(x)} dx < +\infty. \quad (1.6)$$

The process  $Z^K$  can be represented as the unique strong solution of a stochastic differential equation driven by Poisson point measures. Let us consider a Poisson point process  $N(ds, dv, d\theta)$  on  $\mathbb{R}_+ \times \mathcal{U} \times \mathbb{R}_+$  with intensity measure  $ds \otimes n(dv) \otimes d\theta$  where  $ds$  and  $d\theta$  are Lebesgue measures on  $\mathbb{R}_+$  and  $n(dv)$  is the counting measure on the denumerable set  $\mathcal{U}$ . We also introduce a Poisson point measure  $Q(ds, dv, dy, d\theta)$  on  $\mathbb{R}_+ \times \mathcal{U} \times \mathbb{R} \times \mathbb{R}_+$  with intensity measure  $ds \otimes n(dv) \otimes G(y) dy \otimes d\theta$ . We assume that the random measures  $Z_0^K$ ,  $N$  and  $Q$  are independent.

Let us consider a test function  $\varphi \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ , then

$$\begin{aligned} \langle Z_t^K, \varphi \rangle &= \int_{\mathbb{R}} \varphi(x) Z_t^K(dx) = \sum_{v \in V_t^K} \varphi(X_t^{K,v}) \\ &= \langle Z_0^K, \varphi \rangle + \int_0^t \int_{\mathcal{U}} \int_{\mathbb{R}_+} \mathbb{1}_{v \in V_{s-}^K} \varphi(X_{s-}^{K,v}) \left( \mathbb{1}_{\theta \leq b(X_{s-}^{K,v})} \right. \\ &\quad \left. - \mathbb{1}_{b(X_{s-}^{K,v}) < \theta \leq b(X_{s-}^{K,v}) + d(X_{s-}^{K,v})} \right) N(ds, dv, d\theta) \\ &\quad + \int_0^t \int_{\mathcal{U}} \int_{\mathbb{R}} \int_{\mathbb{R}_+} \mathbb{1}_{v \in V_{s-}^K, \theta \leq p(X_{s-}^{K,v})} \varphi \left( X_{s-}^{K,v} + \frac{y}{\log K} \right) Q(ds, dv, dy, d\theta), \end{aligned} \quad (1.7)$$

For  $v \in V_t^K$ , we can define the lineage of  $v$  as follows, and this process will be denoted  $(X_{s \wedge t}^{K,v})_{s \geq 0}$ . It is a path of  $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ , constant after time  $t$  with value  $X_t^{K,v}$ , and that takes at times  $s \leq t$  the trait value of the (unique) ancestor of  $v$  (possibly  $v$  itself) living at this time. In particular, on the event that individual  $v$  reproduces before dying,  $X_t^{K,v} = X_t^{K,v^0} = X_t^{K,v^1}$  for all  $t$  (strictly) smaller than the reproduction time of  $v$ .

Using standard Itô calculus (see [34, 31]), we obtain that:

$$\begin{aligned} \langle Z_t^K, \varphi \rangle &= \langle Z_0^K, \varphi \rangle + \int_0^t \langle Z_s^K, (b + p - d)\varphi \rangle ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} (\varphi(x + \frac{y}{\log K}) - \varphi(x)) p(x) G(y) dy Z_s^K(dx) ds + M_t^{K,\varphi}, \end{aligned} \quad (1.8)$$

where  $M^{K,\varphi}$  is a square integrable martingale with predictable quadratic variation process:

$$\begin{aligned} \langle M^{K,\varphi} \rangle_t &= \int_0^t \int_{\mathbb{R}} (b(x) + p(x) + d(x)) \varphi^2(x) Z_s^K(dx) ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} p(x) (\varphi(x + \frac{y}{\log K}) - \varphi(x))^2 G(y) dy Z_s^K(dx) ds. \end{aligned} \quad (1.9)$$

Using (1.2) and (1.4), it is standard [31] to prove that for any  $T > 0$  and  $n > 0$ ,

$$\mathbb{E}(\sup_{t \leq T} \langle Z_t^K, 1 \rangle^n) < +\infty,$$

and then the process  $Z^K$  is well defined on any time interval  $[0, T]$  and belongs to  $\mathbb{D}(\mathbb{R}_+, \mathcal{M}_P(\mathbb{R}))$ .

We also introduce the associated historical process  $(\Theta_t^K, t \geq 0)$  which is a point measure-valued process taking values in  $\mathcal{M}_P(\mathbb{D})$ , and defined for any  $t \geq 0$  by

$$\Theta_0^K = Z_0^K ; \quad \Theta_t^K = \sum_{v \in V_t^K} \delta_{X_{\cdot \wedge t}^{K,v}}.$$

It is possible to write an SDE in the spirit of (1.7) to describe the dynamics of the historical process, see [10, 33, 42].

In the sequel, because the mutation steps are of order of magnitude  $1/\log K$ , we will consider the process at the time scale  $\log K$ . Hence we define the time-changed historical process  $\tilde{\Theta}^K$  taking values in  $\mathcal{M}_P(\mathbb{D})$  for all  $t > 0$  as

$$\tilde{\Theta}_t^K = \sum_{v \in V_{t \log K}^K} \delta_{X_{(\cdot \wedge t) \log K}^{K,v}}. \quad (1.10)$$

For all  $T > 0$  and all measurable  $A \subset \mathbb{D}[0, T]$ , we define for all  $t \in [0, T]$  the set

$$A_t := \{f_{\cdot \wedge t}, f \in A\} \quad (1.11)$$

of functions in  $A$  stopped at time  $t$ . We define the number of particles living at time  $t \log K$  having their lineage in  $A$  by

$$N_t^{K,A} = \langle \tilde{\Theta}_t^K, \mathbb{1}_{A_t} \rangle. \quad (1.12)$$

As in our seminal works [15] and [16], we are interested in capturing the number of trajectories at logarithmic time scale living in some fixed set, which is of the order of magnitude of a power of  $K$ . To obtain the limiting dynamics (in  $K$ ) of the exponent of such  $K$ -power number, we are led to study the asymptotic behaviour of  $\log N_t^{K,A}/\log K$ , which will be the aim of Theorems 1.1 and 1.2. Note that these quantities can be seen as the stochastic analog of the Hopf-Cole transformation used by the analysts to describe concentration phenomena (see for example [44] and Section 1.3 for more details on the related literature),  $1/\log K$  playing in our setting the role of  $\varepsilon$  in the usual deterministic setting. To achieve this, we use techniques developed in the theory of large deviations and branching Markov processes.

## 1.2 Useful mathematical objects and main results

In this section, we introduce some notations and useful functions that will allow us to state our main results. We define

$$H(\alpha) = \int_{\mathbb{R}} (e^{\alpha y} - 1)G(y) dy, \quad (1.13)$$

$$L(x, v) = \sup_{\alpha \in \mathbb{R}} (\alpha v - p(x)H(\alpha)). \quad (1.14)$$

Our assumptions on  $G$  imply that  $H(\alpha) < \infty$  for all  $\alpha \in \mathbb{R}$  and  $H$  and  $L$  are convex and superlinear functions with respect to  $\alpha$  and  $v$  (see [25, Section 3.3]). In particular, there exist a positive constant  $A$  and a superlinear function  $\mu : \mathbb{R}^+ \rightarrow \mathbb{R}$ , with  $\lim_{r \rightarrow +\infty} \frac{\mu(r)}{r} = +\infty$ , such that for all  $(x, v) \in \mathbb{R}^2$ ,

$$\mu(|v|) - A \leq L(x, v), \quad \lim_{|v'| \rightarrow +\infty} \frac{\partial_v L(x, v') \cdot v'}{|v'|} = +\infty. \quad (1.15)$$

Next, for any  $f \in \mathbb{D}[0, t]$ , we define

$$F_t(f) = \beta_0(f(0)) + \int_0^t R(f_s) ds - I_t(f), \quad (1.16)$$

with

$$I_t(f) = \begin{cases} \int_0^t L(f_s, \dot{f}_s) ds & \text{if } f \in AC[0, t], \\ +\infty & \text{otherwise.} \end{cases} \quad (1.17)$$

The functions  $F_t$  and  $I_t$  will play respectively the roles of a cost function and a good rate function associated with a large deviation principle as we will see later in the article. A non-variational formulation of the rate function  $I_t$  will be given in Section 2.3.

Our main results are asymptotic upper and lower bounds on the logarithm of  $N_t^{K,A}$  defined in (1.12). In the sequel, for any  $a \in \mathbb{R}$  and any sequence of real random variables  $(X_n)_{n \in \mathbb{N}}$ , we will say that

$$\limsup_{n \rightarrow +\infty} X_n \leq a \quad \text{in probability}$$

if, for all  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(X_n \geq a + \varepsilon) = 0.$$

The extension to  $\liminf X_n \geq a$  in probability is straightforward.

**Theorem 1.1.** For any  $t > 0$  and any closed set  $A \subset \mathbb{D}[0, t]$ , we have, in probability

$$\limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K,A} \leq \sup\{F_t(f); f \in A, \forall s \in [0, t], F_s(f) \geq 0\}.$$

**Theorem 1.2.** For any  $t > 0$  and any open set  $G \subset \mathbb{D}[0, t]$ , we have, almost surely,

$$\liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K,G} \geq \sup\{F_t(f); f \in G, \forall s \in [0, t], F_s(f) > 0\}.$$

Note that, by definition of  $I_t(f)$ , the upper bound in Theorem 1.1 is equal to

$$\sup\{F_t(f); f \in A \cap AC[0, t], \forall s \in [0, t], F_s(f) \geq 0\}$$

and similarly for Theorem 1.2. Some heuristic explanations for these results are given at the end of this section.

To link these results to a Hamilton-Jacobi equation, let us define for  $a \in \mathbb{R}$ ,

$$u_a(t, x) = \sup\{F_t(f); f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s(f) \geq a\}, \quad (1.18)$$

$$\tilde{\Omega}_a = \{(t, x) \in [0, +\infty) \times \mathbb{R}; \exists f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s(f) \geq a\}.$$

Notice that it is immediate that

$$u_a(t, x) \geq a, \quad \text{for all } (t, x) \in \tilde{\Omega}_a, \quad \text{and} \quad u_a(t, x) = -\infty, \quad \text{for all } (t, x) \in \tilde{\Omega}_a^c, \quad (1.19)$$

where we used the convention that  $\sup \emptyset = -\infty$ .

The next results show that the functions  $u_a(t, x)$  allow us to characterize the asymptotic density of individuals at  $(t, x)$  in the stochastic population process  $Z^K$  and are solutions of some state-constrained Hamilton-Jacobi equations.

Let us define for all  $t \geq 0$ ,  $x \in \mathbb{R}$  and  $\delta > 0$

$$A_t^{x,\delta} = \{f \in \mathbb{D}[0, t], f(t) \in [x - \delta, x + \delta]\},$$

$$G_t^{x,\delta} = \{f \in \mathbb{D}[0, t], f(t) \in (x - \delta, x + \delta)\}.$$

Then, we have the following result.

**Theorem 1.3.** For all  $t \geq 0$  and  $x \in \mathbb{R}$ , in probability

$$\lim_{a \downarrow 0} u_a(t, x) \leq \liminf_{\delta \rightarrow 0} \liminf_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, G_t^{x,\delta}} \leq \limsup_{\delta \rightarrow 0} \limsup_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, A_t^{x,\delta}} \leq u_0(t, x).$$

Moreover, if  $a \mapsto u_a(t, x)$  is continuous at  $a = 0$ , then, in probability

$$\lim_{\delta \rightarrow 0} \lim_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, A_t^{x,\delta}} = \lim_{\delta \rightarrow 0} \lim_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, G_t^{x,\delta}} = u_0(t, x). \quad (1.20)$$

To define the Hamilton-Jacobi equation, we also need to introduce the set

$$\Omega_a = \{(t, x) \in \tilde{\Omega}_a \mid u_a(t, x) > a\},$$

**Theorem 1.4.** *The set  $\Omega_a$  is an open set. The function  $u_a$  belongs to  $C(\text{cl}(\Omega_a))$  and it is the unique locally Lipschitz-continuous and bounded above viscosity solution of the following Hamilton-Jacobi equation*

$$\begin{cases} \partial_t u = p(x)H(\partial_x u) + R(x), & (t, x) \in \Omega_a \\ u(t, x) = a, & (t, x) \in \partial\Omega_a, t > 0, \\ u(0, x) = \beta_0(x), & \text{for all } x \text{ s.t } \beta_0(x) > a. \end{cases} \quad (1.21)$$

The proof of Theorem 1.4 is postponed to Section 6.

Notice that  $u_a$  defined by (1.18), satisfies a state constraint boundary condition in  $\Omega_a$  ([2, Section 5.1.3]), i.e.  $F_s(f)$  has to remain larger than  $a$  for all  $s \in [0, t]$  in the variational formula (1.18). Moreover,  $u_a = -\infty$  in  $\tilde{\Omega}_a^c$  and  $u_a = a$  in  $\tilde{\Omega}_a \setminus \Omega_a$ . We lastly show in the following lemma that, for a.e.  $a$ ,  $\tilde{\Omega}_a \setminus \Omega_a$  is a Lebesgue-null set and  $u_a$  is right continuous with respect to  $a$ . For this, let us define

$$\Gamma_{a_0} = \bigcup_{a > a_0} \tilde{\Omega}_a.$$

One can verify that, for all  $a_1, a_2 \in \mathbb{R}$ , with  $a_1 < a_2$ , we have

$$\tilde{\Omega}_{a_2} \subset \tilde{\Omega}_{a_1}, \quad \Omega_{a_2} \subset \Omega_{a_1}, \quad u_{a_2}(\cdot, \cdot) \leq u_{a_1}(\cdot, \cdot),$$

and for all  $a_0 \in \mathbb{R}$ ,

$$\Gamma_{a_0} \subset \Omega_{a_0} \subset \tilde{\Omega}_{a_0}, \quad \text{and for all } (t, x) \in \tilde{\Omega}_{a_0}^c, \quad \lim_{a \downarrow a_0} u_a(t, x) = u_{a_0}(t, x) = -\infty. \quad (1.22)$$

**Lemma 1.5.** *For almost every  $a_0 \in \mathbb{R}$ , we have*

$$\int_{\mathbb{R}^+ \times \mathbb{R}} \mathbf{1}_{\tilde{\Omega}_{a_0} \setminus \Gamma_{a_0}}(t, x) dt dx = 0, \quad (1.23)$$

and

$$\forall (t, x) \in \Gamma_{a_0} \cup \tilde{\Omega}_{a_0}^c, \quad \lim_{a \downarrow a_0} u_a(t, x) = u_{a_0}(t, x). \quad (1.24)$$

The proof is also deferred to Section 6.

**Remark 1.6.** *Notice that Lemma 1.5 is satisfied for almost every  $a_0$ , but might fail for  $a_0 = 0$ . We explain here that up to perturbing slightly the initial condition, we always have that for all  $(t, x) \in \Gamma_0 \cup \Omega_0^c$  and in particular for a.e.  $(t, x)$ , in probability*

$$\lim_{\delta \rightarrow 0} \lim_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, A_t^{x, \delta}} = u_0(t, x).$$

The exact formulation will appear in (1.26). To precise this statement, let us define for any  $f \in AC[0, t]$

$$\begin{aligned} F_t^\mu(f) &= \beta_0(f(0)) - \mu + \int_0^t (b + p - d)(f(s)) ds - I_t(f), \\ u_0^\mu(t, x) &= \sup \{ F_t^\mu(f); f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s^\mu(f) \geq 0 \}, \end{aligned} \quad (1.25)$$

$$\begin{aligned}\tilde{\Omega}_0^\mu &= \{(t, x) \in [0, +\infty) \times \mathbb{R}; \exists f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s^\mu(f) \geq 0\}, \\ \Omega_0^\mu &= \{(t, x) \in \tilde{\Omega}_0^\mu \mid u_0^\mu(t, x) > 0\}, \\ \Gamma_0^\mu &= \bigcup_{a > a_0} \tilde{\Omega}_a^\mu.\end{aligned}$$

One can verify that

$$\tilde{\Omega}_0^\mu = \tilde{\Omega}_\mu, \quad \Omega_0^\mu = \Omega_\mu, \quad \Gamma_0^\mu = \Gamma_\mu, \quad u_0^\mu(t, x) = u_\mu(t, x) - \mu.$$

Then, Lemma 1.5 implies that for almost every  $\mu_0$  we have

$$\forall (t, x) \in \Gamma_0^{\mu_0} \cup \Omega_0^{\mu_0}, \quad \lim_{\mu \downarrow \mu_0} u_0^\mu(t, x) = u_0^{\mu_0}(t, x).$$

Let us also define

$$N_t^{K, A, \mu} := \langle \tilde{\Theta}_t^{K, \mu}, \mathbb{1}_{A_t} \rangle,$$

where  $\tilde{\Theta}_t^{K, \mu}$  is the historical birth-death process defined in (1.10) and  $A_t$  defined in (1.11), such that  $Z_0^{K, \mu}$  is a Poisson point measure in  $\mathbb{R}$  with intensity measure  $K\beta_0^K(x)^{-\mu}dx$ . We finally deduce from Theorem 1.3 that for a.e.  $\mu$  and all  $(t, x) \in \Gamma_0^\mu \cup \Omega_0^\mu$ , and in particular for a.e.  $(t, x) \in \mathbb{R}^+ \times \mathbb{R}$ , in probability

$$\lim_{\delta \rightarrow 0} \lim_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, A_t^{\delta, \mu}} = u_0^\mu(t, x). \quad (1.26)$$

□

Theorems 1.1 and 1.2 also allow us to identify the typical lineages of the population. As we saw in Theorem 1.3, the density of individuals of trait  $x$  at time  $t$  can be estimated as in (1.20) with

$$u_0(t, x) = \sup \{F_t(f); f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s(f) \geq 0\}.$$

Let us assume that  $f_o$  is an optimal trajectory where the supremum above is attained. It is proved in Lemma 6.1 that such a trajectory necessarily exists. Assume also that  $F_s(f_o) > 0$  for all  $s \in [0, t]$  so that  $u_0(t, x) = \lim_{a \downarrow 0} u_a(t, x)$ . Then, Theorems 1.1 and 1.2 imply that a subpopulation of comparable size (in the logarithmic scale) to the population size close to  $(t, x)$  has indeed followed the trajectory  $f_o$ . More precisely, let us define

$$A_{\delta, f_o} = \{\gamma \in \mathbb{D}[0, t]; d_{\text{Sko}}(\gamma, f_o) \leq \delta\},$$

$$G_{\delta, f_o} = \{\gamma \in \mathbb{D}[0, t]; d_{\text{Sko}}(\gamma, f_o) < \delta\}.$$

Then, Theorems 1.1 and 1.2 imply that, in probability

$$\limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_{\delta, f_o}} \leq \sup \{F_t(f); f \in A_{\delta, f_o}, \forall s \in [0, t], F_s(f) \geq 0\}$$

and

$$\liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, G_{\delta, f_o}} \geq \sup \{F_t(f); f \in G_{\delta, f_o}, \forall s \in [0, t], F_s(f) > 0\}.$$

Letting  $\delta \rightarrow 0$  we then obtain, in probability

$$u_0(t, x) \leq \liminf_{\delta \rightarrow 0} \liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, G_{\delta, f_o}} \leq \limsup_{\delta \rightarrow 0} \limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_{\delta, f_o}} \leq u_0(t, x).$$

where we have used  $u_0(t, x) = F_t^{f_o}$ . We hence deduce the following result.

**Corollary 1.7.** *Let  $(t, x) \in (0, +\infty) \times \mathbb{R}$  and  $f_o$  be an optimal trajectory such that  $f_o(t) = x$  and  $u_0(t, x) = F_t(f_o)$ . Assume also that  $F_s(f_o) > 0$  for all  $s \in [0, t]$ . We then have in probability*

$$\lim_{\delta \rightarrow 0} \lim_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, G_{\delta, f_o}} = \lim_{\delta \rightarrow 0} \lim_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_{\delta, f_o}} = u_0(t, x).$$

Thus any optimal trajectory  $f_o$  can be interpreted as the ancestral lineage of a large part of the population having trait  $x$  at time  $t$ .

Let us end this section with some comments. As explained before, the quantity

$$\beta_t^{K, A} = \frac{1}{\log K} \log N_t^{K, A}$$

is the exponent in  $K$  of the number of particles  $N_t^{K, A}$ , in the sense that  $N_t^{K, A} = K^{\beta_t^{K, A}}$ . The evolution in time of the number of particles around the path  $f$  is approximately  $K^{F_t(f)}$ . This says that the exponent, starting from  $\beta_0(f(0))$ , changes according to the births and deaths along this path as  $\int_0^t R(f_s) ds$ . For comparison, remember that for a branching process without mutation and with constant growth rate  $R$ ,  $\mathbb{E}(N_t^{K, A}) = \mathbb{E}(N_0^{K, A})e^{Rt}$ . The penalization by  $-I_t(f)$ , as we will see later, comes from the fact that the probability for an ancestral lineage to be around  $f$  is of the order of  $K^{-I_t(f)}$ . The state constraint boundary condition tells that only paths  $f$  such that  $F_s(f) > 0$  for all  $s \in [0, t]$  are admissible: the population gets extinct on the way otherwise. Notice also that the assumption that  $F_s(f) > 0$  for all  $s \in [0, t]$  is not contradictory with the fact that for some  $s$  we may have  $R(f(s)) < 0$ .

### 1.3 Comparison with previous works

Our main results take the form of large deviation estimates on population sizes. This question for spatial branching Markov processes for large time goes back to [7] and has been studied by several authors, notably [6], from which a large part of our work is inspired. These works deal with branching Brownian motions on the line and aim at describing the particles that constitute the right front. The method of [6] is based on additive martingales for the branching process and the spine decomposition of [32, 41]. We use a similar approach to obtain large deviations upper bounds on the branching population size. Many other works dealing more specifically with estimates on the position of the rightmost particle in branching Brownian motion used methods based on moment estimates [27, 49]. Similar questions for general branching random walks are studied in [40] and are thus closer to our work. Our proof of the large deviations lower bound on the branching population size is inspired from these works, although we focus on a different question dealing with the study of local population densities in a model with inhomogeneous space dependence of rates. In [39, 47], spatial birth-death processes with interaction are used to study the effect of population sizes and demographic stochasticity on the speed of invasion fronts under different scalings. In particular, in [39, section 1.2] a conjecture has been provided, involving a variational formulation closely related to what we obtain at the limit in Theorems 1.1 and 1.2. This variational problem features a different Lagrangian, which requires a different scaling of the problem.

Hamilton-Jacobi equations have been widely used in the asymptotic analysis of integro-differential equations in evolutionary biology (see for instance [44, 5, 38]) but also in the study

of propagation phenomena (see e.g. [29, 26, 4]). Let us consider the following model

$$\begin{cases} \varepsilon \partial_t n_\varepsilon(t, x) = \int_{\mathbb{R}} p(x+h) n_\varepsilon(x+h) G(h/\varepsilon) dh / \varepsilon + n_\varepsilon(b(x) - d(x)), \\ n_\varepsilon(0, x) = \exp(\beta_0(x)/\varepsilon). \end{cases} \quad (1.27)$$

Here  $n_\varepsilon(t, x)$  stands for the phenotypic density of a population, with  $t \in \mathbb{R}^+$  and  $x \in \mathbb{R}$  corresponding respectively to time and to a phenotypic trait. Similarly to above,  $b(x)$  and  $p(x)$  stand for birth rates without and with mutations and  $d(x)$  corresponds to a death rate. The mutations are distributed as  $\frac{1}{\varepsilon} G(\frac{y}{\varepsilon}) dy$ . The mutational variance scales as  $\varepsilon^2$ , which is assumed to be a small parameter. A change of variable in time  $t \rightarrow t/\varepsilon$  has then been taken into account to accelerate the slow dynamics resulting from small effects of mutations. This change of variable leads to the  $\varepsilon$  coefficient in front of  $\partial_t n_\varepsilon$ .

Such an equation can be related to the stochastic process  $Z_t^K$  introduced in Section 1.1 in two ways. Firstly, it can be obtained as a large population limit, that is  $K \rightarrow +\infty$  of the stochastic process  $Z_t^K/K$  (see [31, 13]), but taking the mutational variance constant of order  $\varepsilon^2$  and independent of  $K$ . Secondly, the expectation of the stochastic process above satisfies (1.27) with  $1/\log K = \varepsilon$ .

The asymptotic behavior of  $n_\varepsilon$  as  $\varepsilon \rightarrow 0$  can be described via a Hopf-Cole transformation:

$$U_\varepsilon(t, x) = \varepsilon \log(n_\varepsilon(t, x)).$$

Notice the analogy of this transformation with the transformation  $\frac{1}{\log K} \log N_t^{K,A}$  used above. It is proved in [5] (in a slightly different setting, taking into account a competition term) that as  $\varepsilon \rightarrow 0$ ,  $U_\varepsilon$  converges to the unique viscosity solution of

$$\begin{cases} \partial_t U = p(x)H(\partial_x U) + R(x), & x \in \mathbb{R} \\ U(0, x) = \beta_0(x), \end{cases} \quad (1.28)$$

with  $H$  and  $R$  defined in (1.13) and (1.3). Notice that this is the same equation as (1.21) but set in the whole domain  $\mathbb{R}$ . When considering the asymptotic behavior of the stochastic process instead of the deterministic integro-differential equation, possible extinction of small subpopulations are taken into account. This leads to a smaller limit  $u_0 \leq U$ . The limit  $u_0$  of the stochastic process can in particular take  $-\infty$  as value. The variational formulation of the problem given in (1.18) provides an intuitive explanation. The maximal trajectories in the variational problem correspond indeed to the typical trajectories of lineages as obtained in Theorems 1.1 and 1.2. If the value function on such a trajectory takes a negative value  $-c < 0$ , then the expected population size approaches the small size of order  $K^{-c}$  as proved later in Theorem 3.1, which results with high probability in extinction, so the trajectory should be excluded. This is a significant difference with the deterministic derivation where all trajectories are allowed, leading to the Hamilton-Jacobi equation (1.28) in the whole domain. In the stochastic derivation the limit  $u_0$  is positive in the set  $\bar{\Omega}_0$  and equal to  $-\infty$  outside this set. The function  $u_0$  satisfies both a Dirichlet boundary condition and a state constraint condition. Several comments are in order.

(i) In deterministic works, one usually considers a slightly more complex model taking into account a nonlocal mortality rate due to competition [44, 5, 38]. In this case, the growth rate is given by  $R(x, I_\varepsilon(t))$ , with  $I_\varepsilon(t) = \int_{\mathbb{R}} n_\varepsilon(t, x) dx$ . Such a mortality rate leads often to a constraint of type

$$\max_{x \in \mathbb{R}} U(t, x) = 0.$$

This constraint on the limit  $U$  might seem confusing when it is combined with the threshold on the trajectories in the stochastic derivation leading to  $u_0 \geq 0$ . Note however that in order to make a relevant comparison between these results we have to divide the population process  $Z_t^K$  by  $K$  (see [13] where (1.27) has been derived from a stochastic model). This means that, when put in a similar framework than usual deterministic works, one has to put the threshold of extinction equal to  $-1$  instead of  $0$ . The expected equation, in presence of a competition term, would then be given by

$$\begin{cases} \partial_t u = H(x, \partial_x u) + R(x, I), & (t, x) \in \Omega^I \\ \max_{x \in \mathbb{R}} u(t, x) = 0, \\ u(t, x) = -1, & (t, x) \in \partial\Omega^I, t > 0, \\ u(0, x) = \beta_0(x) - 1, & \text{for all } x \text{ s.t } (0, x) \in \Omega^I, \end{cases}$$

with a set  $\Omega^I$  which depends on the competition term  $I(\cdot)$  and would be such that  $u$  also satisfies a state constraint boundary condition in  $\Omega^I$ . Obtaining this equation rigorously is the aim of a future work.

(ii) Biological criticisms were made on the Hamilton-Jacobi method because of the so-called tail problem [45]. Artifacts may indeed arise due to an inadequate treatment of small subpopulations. Exponentially small subpopulations which actually may be extinct can have a strong influence on the future of the population. Artificial jumps of the dominant trait may occur. The branching patterns are also too fast. Modifications of the Hamilton-Jacobi equation were suggested in [45, 35, 43] to solve this problem. These modifications were directly made to deterministic models. Here we use a stochastic individual-based approach providing a more biologically relevant justification of the outcome. Note however that we obtain a closely related limit to [43], even though this correction was described in a less direct and less precise way in [43]. Moreover the threshold of extinction that was considered in [43] was arbitrary. Here the threshold is obtained with a direct link to the population size.

(iii) As mentioned above, Hamilton-Jacobi equations are widely used to provide approximations of the phenotypic density of a population in a small mutational variance regime [44, 5, 38]. Here, we go further than characterizing the phenotypic density. We also identify the typical lineages of the population thanks to Corollary 1.7. A previous work [30] had already made a link between the optimal trajectories of the Hamilton-Jacobi equation and the typical lineages of the population. These authors considered a deterministic model in a context of changing environment and they used the neutral fractions approach to study the inside dynamics of the population. Probabilist spinal techniques and historical birth and death processes that allow to link the typical lineages with stochastic individual-based models have been used in [10, 33] to describe the phylogenies but only under a large population limit, whereas here, mutations and time are also rescaled. Some techniques that are used in the present work, such as using Feynman-Kac formulas for characterizing spines, are still taken from these papers (see also [32]).

## 2 Study of an auxiliary jump process

### 2.1 A Feynman-Kac formula for $\mathbb{E}(N_t^{K,A})$

Given  $t > 0$  and  $A \subset \mathbb{D}[0, t]$ , we interpret  $\mathbb{E}(N_t^{K,A})$  as the expectation of a functional of an auxiliary process based on the mutations dynamics.

Based on (1.8), let us introduce a random walk in continuous time  $(X_t^K)_{t \in \mathbb{R}_+}$  with paths in  $\mathbb{D}$  and infinitesimal generator

$$\mathcal{L}^K \varphi(x) = p(x) \int_{\mathbb{R}} \left[ \varphi \left( x + \frac{y}{\log K} \right) - \varphi(x) \right] G(y) dy, \quad (2.1)$$

defined for any measurable bounded function  $\varphi$ .

A pathwise representation similar to (1.7) of the process  $(X_t^K)_{t \geq 0}$  can be obtained as follows: let us give ourselves (on some probability space) a Poisson point measure  $Q(ds, d\theta, dy)$  on  $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$  with intensity  $G(y)dy \otimes d\theta \otimes ds$  and an independent real random variable  $X_0^K$ . We can write

$$X_t^K = X_0^K + \int_0^t \int_{\mathbb{R}_+} \int_{\mathbb{R}} \frac{y}{\log K} \mathbb{1}_{\{\theta \leq p(X_{s-}^K)\}} Q(ds, d\theta, dy). \quad (2.2)$$

Let us also define

$$Y_t^K = X_{t \log K}^K, \quad (2.3)$$

the process in the time scale  $\log K$ . For all  $t > 0$  and  $x \in \mathbb{R}$ , we denote by  $\mu_{x,t}^K$  the law of the process  $(Y_s^K)_{s \in [0,t]}$  on  $\mathbb{D}[0,t]$  conditionally on  $Y_0^K = X_0^K = x$ , and by  $\mathbb{E}_{\mu_{x,t}^K}$  the corresponding expectation.

We have the following classical Feynman-Kac representation of  $N_t^{K,A}$  (also called many-to-one formula).

**Proposition 2.1.** *Let  $x \in \mathbb{R}$ . We consider the birth-death-mutation process  $(Z_t^K, t \geq 0)$  defined as before but started from a unique individual with trait  $x$ , i.e.  $Z_0^K = \delta_x$ , and we denote the corresponding expectation by  $\mathbb{E}_{\delta_x}$ .*

(i) *Let  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  be bounded and measurable. Then, for any  $t > 0$ , we have*

$$\mathbb{E}_{\delta_x} [\langle Z_t^K, \varphi \rangle] = \mathbb{E}_x \left[ \exp \left( \int_0^t R(X_s^K) ds \right) \varphi(X_t^K) \right], \quad (2.4)$$

where  $X^K$  is the process defined in (2.2).

(ii) *For  $t > 0$ ,  $x \in \mathbb{R}$  and for a bounded measurable function  $\Phi : \mathbb{D}[0,t] \rightarrow \mathbb{R}$ ,*

$$\mathbb{E}_{\delta_x} [\langle \Theta_t^K, \Phi \rangle] = \mathbb{E}_{\delta_x} \left[ \sum_{u \in V_t^K} \Phi(X_s^{K,u}, s \leq t) \right] = \mathbb{E}_x \left[ \exp \left( \int_0^t R(X_s^K) ds \right) \Phi(X_s^K, s \leq t) \right]. \quad (2.5)$$

(iii) *For all  $t > 0$  and  $A \subset \mathbb{D}[0,t]$ ,*

$$\mathbb{E}(N_t^{K,A}) = \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,t}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in A} \right] dx.$$

*Proof.* The proof of point (i) for  $\varphi \in C_b(\mathbb{R})$  is given in Section A.1 for the sake of completeness. Since the set of bounded continuous functions on  $\mathbb{R}$  is dense for the bounded pointwise topology in the set of bounded measurable functions [24, Prop. 4.2, Chap. 3], Point (i) extends easily to any  $\varphi$  bounded measurable. Note that the Feynman-Kac formula (2.4) concerns the law of first moments of  $Z_t^K$  (with fixed  $t$ ) issued from one individual with trait  $x$  and can be extended to the whole trajectory using standard techniques (see [10, 41]), providing (ii). Point (iii) then follows from (1.4) changing time  $t$  by  $t \log K$  by change of variables.  $\square$

In what follows, we will also need a many-to-one formula for the whole tree, as in [1, 41]. Recall that  $V_{[0,t]}^K = \bigcup_{s \in [0,t]} V_s^K$  is the set of individuals born before time  $t$  (including those still alive at time  $t$ ). Recall that our continuous-time birth-death process is associated with a binary tree where each node corresponds to a birth or death event. If it is a death event, the node is a leaf. If it is a birth event, then the individual  $v$  is replaced with  $v0$  and  $v1$ , where  $v0$  is the continuation of the mother  $v$  (with the same trait) and  $v1$  is the new offspring (with a possible mutated trait). Let us also denote by  $S_v^0$  the birth time of  $v$  and by  $S_v$  the time at which  $v$  disappears (either by death or reproduction).

**Proposition 2.2.** *We have that for all  $t > 0$ ,  $\Phi : \mathbb{D}[0, t] \times [0, t] \rightarrow \mathbb{R}$  a bounded measurable function and  $x \in \mathbb{R}$ :*

$$\begin{aligned} \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_{[0,t]}^K} \Phi((X_{r \wedge S_v}^{K,v}, r \leq t), S_v \wedge t) \right] \\ = \int_0^t \mathbb{E}_x \left[ \Phi((X_{r \wedge s}^K, r \leq t), s) (b + p + d)(X_s^K) \exp \left( \int_0^s R(X_r^K) dr \right) \right] ds \\ + \mathbb{E}_x \left[ \Phi((X_r^K, r \leq t), t) \exp \left( \int_0^t R(X_r^K) dr \right) \right]. \end{aligned} \quad (2.6)$$

Notice that if there exists a function  $\Psi : \mathbb{D}[0, t] \times [0, t] \rightarrow \mathbb{R}$  such that for all  $(f, s) \in \mathbb{D}[0, t] \times [0, t]$ ,  $\Phi(f, s) = \Psi(f, s) \mathbb{1}_{s < t}$ , then the second term in the right hand side of (2.6) vanishes.

The proof of Proposition 2.2 is deferred to Appendix A.2.

Using Proposition 2.2, we can establish an identity for forks, i.e. sums over pairs of individuals living at time  $t$  (also called many-to-two formula). We do not state this corollary in full generality as this would require extra notations, but will prove a version tailored for our needs in Section 5.1 (see Lemma 5.2).

## 2.2 Large deviation principle for $(\mu_{x,t}^K)_{K \geq 1}$

For all  $t > 0$  and  $x \in \mathbb{R}$ , we define the function  $I_t$  restricted to functions starting from  $x$  as follows: for all  $f \in \mathbb{D}[0, t]$ ,

$$I_{t,x}(f) = \begin{cases} I_t(f) & \text{if } f(0) = x, \\ +\infty & \text{otherwise.} \end{cases}$$

Note that  $I_t(f) = I_{t,f(0)}(f)$ .

The following large deviations principle is a direct application of [21, Theorem 10.2.6]. Indeed, the conditions 10.2.2 and 10.2.4 in this theorem are obviously satisfied thanks to the assumptions on the measure  $p(x)G(y)dy$ .

**Theorem 2.3.** *For all  $t > 0$  and  $x \in \mathbb{R}$ , the family of laws  $(\mu_{x,t}^K)_{K \geq 1}$  satisfies a large deviation principle on  $\mathbb{D}[0, t]$  with rate  $1/\log K$  and good rate function  $I_{t,x}$ : for any subset  $A \subset \mathbb{D}[0, t]$ ,*

$$\begin{aligned} - \inf_{f \in \text{int}(A)} I_{t,x}(f) &\leq \liminf_{K \rightarrow +\infty} \frac{1}{\log K} \mathbb{P}_x((X_{s \log K}^K)_{s \leq t} \in A) \\ &\leq \limsup_{K \rightarrow +\infty} \frac{1}{\log K} \mathbb{P}_x((X_{s \log K}^K)_{s \leq t} \in A) \leq - \inf_{f \in \text{cl}(A)} I_{t,x}(f), \end{aligned} \quad (2.7)$$

where  $\text{int}(A)$  and  $\text{cl}(A)$  are the interior and closure of  $A$  for the Skorohod topology.

In addition, the family of measures  $(\mu_{x,t}^K)_{K \geq 1}$  satisfies the Laplace principle with rate  $1/\log K$  and rate function  $I_{t,x}$  uniformly on compact sets with respect to  $x \in \mathbb{R}$ , as recalled below.

We recall that the rate function  $I_{t,x}$  is good if it is lower semi-continuous and if  $\{I_{t,x} \leq M\}$  is compact for all  $M < \infty$ . We also recall that  $\mu_{x,t}^K$  satisfies the Laplace principle with rate  $1/\log K$  and rate function  $I_{t,x}$  uniformly on compact sets with respect to  $x \in \mathbb{R}$  if, for all compact subset  $A$  of  $\mathbb{R}$  and all constant  $M < \infty$ ,

$$\bigcup_{x \in A} \{f \in \mathbb{D}[0, t], I_{t,x}(f) \leq M\} \quad (2.8)$$

is compact and for all bounded continuous function  $\phi : \mathbb{D}[0, t] \rightarrow \mathbb{R}$

$$\lim_{K \rightarrow \infty} \sup_{x \in A} \left| \frac{1}{\log K} \log \mathbb{E}_{\mu_{x,t}^K} \left( K^{-\phi(Y_s^K, s \leq t)} \right) + \inf_{f \in \mathbb{D}[0, t]} (\phi(f) + I_{t,x}(f)) \right| = 0. \quad (2.9)$$

See [21, Def. 1.2.6].

Let us prove a uniform in  $x$  version of the exponential tightness of the measures  $\mu_{x,t}^K$ .

**Lemma 2.4.** *For all  $t > 0$ , the family of measures  $(\mu_{x,t}^K)_{K,x}$  is exponentially tight, uniformly on compact sets. This means that, for all  $M < \infty$  and all compact subset  $B$  of  $\mathbb{R}$ , there exists a compact subset  $A$  of  $\mathbb{D}[0, t]$  such that*

$$\limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mu_{x,t}^K(A^c) \leq -M.$$

*Proof.* The proof of Lemma 2.4 is inspired from [19, Exercice 4.1.10].

*Step 1: asymptotic bound.* Our first goal is to prove that, for all  $M_0 < \infty$ , all  $\eta > 0$  and all compact subset  $B$  of  $\mathbb{R}$ , there exist an integer  $m \geq 1$ , functions  $f_1, \dots, f_m \in \mathbb{D}[0, t]$  and an integer  $K_0$  such that, for all  $K \geq K_0$ ,

$$\sup_{x \in B} \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_\eta(f_i) \right)^c \right] \leq K^{-M_0}, \quad (2.10)$$

where  $B_\eta(f)$  is the open ball of radius  $\eta$  centered at  $f$  for Skorohod's distance  $d_{\text{Sko}}$ .

Fix  $\eta > 0$ ,  $M_0 < \infty$  and  $B \subset \mathbb{R}$  compact. The set  $\mathbb{D}[0, t]$  equipped with the Skorohod topology is a Polish space, so in particular there exists a dense sequence  $(f_n)_{n \geq 1}$ . Let  $m \in \mathbb{N}$ . For all  $k \geq 1$ , we define for  $f \in \mathbb{D}[0, t]$

$$\phi_k(f) = k d_{\text{Sko}}(f, (\cup_{i=1}^m B_\eta(f_i))^c). \quad (2.11)$$

Since the function  $\phi_k$  is bounded and continuous, it follows from the Laplace principle uniformly on compact sets (2.9) that

$$\limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_\eta(f_i) \right)^c \right] \leq \limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mathbb{E}_{\mu_{x,t}^K} (K^{-\phi_k}) \quad (2.12)$$

$$\leq - \inf_{x \in B} \inf_{f \in \mathbb{D}[0, t]} (\phi_k(f) + I_{t,x}(f)). \quad (2.13)$$

If  $f \in \cup_{i=1}^m B_{\eta/2}(f_i)$ ,  $\phi_k(f) \geq k\eta/2$ , hence

$$\limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right)^c \right] \leq - \left( \frac{k\eta}{2} \wedge \inf_{x \in B} \inf_{f \notin \cup_{i=1}^m B_{\eta/2}(f_i)} I_{t,x}(f) \right).$$

Now, (2.8) implies that

$$\bigcup_{x \in B} \{f \in \mathbb{D}[0, t], I_{t,x}(f) \leq M_0 + 1\}$$

is compact, hence there exists  $m \geq 1$  such that this set is included in  $\cup_{i=1}^m B_{\eta/2}(f_i)$ . Therefore, for such a value of  $m$ ,

$$\limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right)^c \right] \leq -\frac{k\eta}{2} \wedge (M_0 + 1).$$

Choosing  $k > 2M_0/\eta$  ends the proof of (2.10).

*Step 2: uniform bound.* We now prove the following stronger version of (2.10): for all  $M_0 < \infty$ ,  $\eta > 0$  and compact subset  $B$  of  $\mathbb{R}$ , there exists an integer  $m \geq 1$  and functions  $f_1, \dots, f_m \in \mathbb{D}[0, t]$  such that, for all  $K \geq 1$ ,

$$\sup_{x \in B} \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right)^c \right] \leq K^{-M_0}. \quad (2.14)$$

In view of Step 1, to prove this, increasing  $m$  if necessary, it is sufficient to prove that for any  $K \leq K_0 - 1$ , for all  $m$  large enough,

$$\sup_{x \in B} \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right)^c \right] \leq K^{-M_0}.$$

So we fix  $K \leq K_0 - 1$  and  $M_0 < \infty$  until the end of Step 2. The last claim follows from a continuity property of the process  $Y^K$  constructed from (2.2) and (2.3) with respect to its initial value: given  $x \in \mathbb{R}$  and a Poisson point measure  $Q^K(du, d\theta, dy)$  on  $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$  with intensity  $(\log K)dud\theta G(y)dy$ , we define  $Y^{K,x}$  as the solution to

$$Y_t^K = x + \int_0^t \int_{\mathbb{R}_+} \int_{\mathbb{R}} \frac{y}{\log K} \mathbb{1}_{\{\theta \leq p(Y_{u-}^K)\}} Q^K(du, d\theta, dy)$$

and we define  $\tilde{Y}^{K,x}$  as the solution to

$$\tilde{Y}_t^K = x + \int_0^t \int_{\mathbb{R}_+} \int_{\mathbb{R}} \frac{y}{\log K} \mathbb{1}_{\{\theta < p(\tilde{Y}_{u-}^K)\}} Q^K(du, d\theta, dy).$$

Notice that the inequality in the definition of  $\tilde{Y}_t^K$  is strict, contrary to the one in the definition of  $Y_t^K$ . By standard properties of Poisson point measures,  $Y^{K,x} = \tilde{Y}^{K,x}$  almost surely. In addition, by continuity of the function  $p$ , for any  $\omega$  in the event  $\{Y^{K,x} = \tilde{Y}^{K,x}\}$ , the map

$$y \mapsto (Y_s^{K,y}(\omega) - y)_{s \leq t}$$

is constant for  $y$  in a neighborhood of  $x$ . Therefore, it follows from Lebesgue's theorem that, for any measurable set  $G \subset \mathbb{D}[0, t]$ , defining for all  $y \in \mathbb{R}$

$$G_y := \{f + y, f \in G\}, \quad (2.15)$$

the map

$$y \mapsto \mu_{y,t}^K(G_y)$$

is continuous at  $x$ . Since  $x$  was arbitrary, we deduce that this map is continuous. More precisely, for all  $x \in \mathbb{R}$  and  $G \subset \mathbb{D}[0, t]$ ,

$$|\mu_{y,t}^K(G_y) - \mu_{x,t}^K(G_x)| \leq \mathbb{P}(Y_s^{K,y} - y \neq Y_s^{K,x} - x, \text{ for some } s \in [0, t]) \xrightarrow{y \rightarrow x} 0.$$

In particular, there exists  $\delta_x > 0$  such that for all  $y \in [x - \delta_x, x + \delta_x]$ ,

$$\mathbb{P}(Y_s^{K,y} - y \neq Y_s^{K,x} - x, \text{ for some } s \in [0, t]) \leq \frac{K^{-M_0}}{2}. \quad (2.16)$$

We can assume without loss of generality that  $\delta_x \leq \eta/2$  for all  $x \in \mathbb{R}$ . Since the compact set  $B$  is included in the union of the intervals  $(x - \delta_x, x + \delta_x)$  there exist  $N < \infty$  and  $x_1, \dots, x_N \in \mathbb{R}$  such that

$$B \subset \bigcup_{j=1}^N (x_j - \delta_{x_j}, x_j + \delta_{x_j}).$$

For any  $j \in \{1, \dots, N\}$ , since  $(f_i)_{i \geq 1}$  is dense in  $\mathbb{D}[0, t]$ , there exists  $m_j$  large enough such that

$$\mu_{x_j,t}^K \left[ \left( \bigcup_{i=1}^{m_j} B_{\eta/2}(f_i) \right)^c \right] \leq \frac{K^{-M_0}}{2}.$$

Our goal is now to extend this estimate to any  $x \in B$ . Let  $x \in B$ , there exists  $j \in \{1, \dots, N\}$  such that  $|x - x_j| \leq \delta_{x_j}$ . Recalling that  $\delta_{x_j} < \eta/2$ , for all  $m \geq \sup_{1 \leq j \leq N} m_j$ , we have the inclusion,

$$\left( \bigcup_{i=1}^m B_{\eta/2}(f_i) \right)_{x-x_j} \subset \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right).$$

where the notation  $(\ )_{x-x_j}$  has been defined in (2.15). Therefore,

$$\begin{aligned} \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta}(f_i) \right)^c \right] &\leq \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta/2}(f_i) \right)_{x-x_j}^c \right] \\ &\leq \left| \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta/2}(f_i) \right)_{x-x_j}^c \right] - \mu_{x_j,t}^K \left[ \left( \bigcup_{i=1}^m B_{\eta/2}(f_i) \right)^c \right] \right| \\ &\quad + \mu_{x_j,t}^K \left[ \left( \bigcup_{i=1}^{m_j} B_{\eta/2}(f_i) \right)^c \right] \\ &\leq \mathbb{P}(Y_s^{K,y} - y \neq Y_s^{K,x} - x, \text{ for some } s \in [0, t]) + \frac{K^{-M_0}}{2}. \end{aligned}$$

Hence, by (2.16), we have proved (2.14) for  $m \geq \sup_{1 \leq j \leq N} m_j$ .

*Step 3: Conclusion of the proof.* We fix  $M < \infty$  and a compact subset  $B$  of  $\mathbb{R}$ . For all integer  $k \geq 1$ , we apply (2.14) with  $\eta = 1/k$  and  $M_0 = kM$ : there exists  $m_k < \infty$  such that for all  $K \geq 1$

$$\sup_{x \in B} \mu_{x,t}^K \left[ \left( \bigcup_{i=1}^{m_k} B_{1/k}(f_i) \right)^c \right] \leq K^{-kM}.$$

This implies that

$$\sup_{x \in B} \mu_{x,t}^K \left[ \left( \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{m_k} B_{1/k}(f_i) \right)^c \right] \leq \frac{K^{-M}}{1 - K^{-M}}.$$

Now, the set

$$\bigcap_{k=1}^{\infty} \bigcup_{i=1}^{m_k} B_{1/k}(f_i)$$

is precompact. Indeed, considering any sequence  $(\phi_j)_{j \geq 1}$  in this set, since it belongs to  $\cup_{i=1}^{m_1} B_1(f_i)$ , there exists  $i_1 \in \{1, \dots, m_1\}$  such that  $\phi_j \in B_1(f_{i_1})$  infinitely often. Using the diagonal extraction procedure, we deduce that there exists a subsequence, still denoted  $(\phi_j)_{j \geq 1}$  for convenience, and integers  $i_k \in \{1, \dots, m_k\}$  such that  $\phi_j \in B_{1/k}(f_{i_k})$  for all  $j$  large enough. This implies that the sequence  $(\phi_j)_{j \geq 1}$  is Cauchy, hence the conclusion.

Therefore, the compact set

$$\overline{\bigcap_{k=1}^{\infty} \bigcup_{i=1}^{m_k} B_{1/k}(f_i)}$$

satisfies the claim of Lemma 2.4. □

### 2.3 Non-variational form and domain of the rate function $I_t$

Note that the function  $H$  is a convex function and  $H'$  is a  $C^1$ -diffeomorphism from  $\mathbb{R}$  to itself. We have the following result, providing an alternative, non-variational expression for the rate function  $I_t$  and characterizing the set of functions  $f$  such that  $I_t(f) < +\infty$ .

**Lemma 2.5.** (i) For all  $t$  and  $f \in C^2[0, t]$ ,

$$I_t(f) = \psi_f(t)f(t) - \psi_f(0)f(0) - \int_0^t (f_s \psi'_f(s) + p(f_s)H(\psi_f(s))) ds, \quad (2.17)$$

where for all  $s \in [0, t]$ ,

$$\psi_f(s) = (H')^{-1} \left( \frac{\dot{f}_s}{p(f_s)} \right).$$

(ii) For all  $f \in AC[0, t]$ ,  $I_t(f) < +\infty$  iff

$$\int_0^t \dot{f}_s (H')^{-1} \left( \frac{\dot{f}_s}{\bar{p}} \right) ds = \int_0^t |\dot{f}_s| (H')^{-1} \left( \frac{|\dot{f}_s|}{\bar{p}} \right) ds < +\infty.$$

*Proof.* Let  $f \in C^2[0, t]$ . From (1.14) and (1.17), we have that

$$I_t(f) = \int_0^t \sup_{\alpha \in \mathbb{R}} (\alpha \dot{f}_s - p(f_s)H(\alpha)) ds.$$

For a given  $s \in [0, t]$ , the optimization problem  $\sup_{\alpha \in \mathbb{R}} (\alpha \dot{f}_s - p(f_s)H(\alpha))$  reaches its supremum for  $\alpha$  that solves  $H'(\alpha) = \dot{f}_s/p(f_s)$ , i.e. for  $\alpha = \psi_f(s)$ . Hence

$$I_t(f) = \int_0^t (\psi_f(s) \dot{f}_s - p(f_s)H(\psi_f(s))) ds.$$

For  $f \in C^2[0, t]$ , the function  $\psi_f$  is  $C^1$ , hence integration by parts yields (2.17).

We now proceed with the proof of (ii). Observe that  $L(x, \beta) = p(x)\tilde{L}(\beta/p(x))$ , with  $\tilde{L}(v) := \sup_{\alpha \in \mathbb{R}} \{\alpha v - H(\alpha)\}$  convex, non-decreasing on  $[0, +\infty)$  and non-increasing on  $(-\infty, 0]$  with  $\tilde{L}(0) = 0$ . Since the optimization problem in the definition of  $\tilde{L}$  has a unique solution, we obtain  $\tilde{L}(v) = v(H')^{-1}(v) - H(H'^{-1}(v))$ .

We next deduce from the change of variable  $y \rightarrow -y$  and from the fact that  $G$  is symmetric that for all  $\alpha \in \mathbb{R}$

$$H'(\alpha) = \int_{\mathbb{R}} \frac{ye^{\alpha y} - ye^{-\alpha y}}{2} G(y) dy$$

and

$$H''(\alpha) = \int_{\mathbb{R}} \frac{y^2 e^{\alpha y} + y^2 e^{-\alpha y}}{2} G(y) dy.$$

Hence, using that  $x \sinh x \leq x^2 \cosh x$  for all  $x \in \mathbb{R}^+$ ,

$$|H'(\alpha)| \leq \int_{\mathbb{R}} |y| \sinh(|\alpha y|) G(y) dy \leq \frac{1}{|\alpha|} \int_{\mathbb{R}} (\alpha y)^2 \cosh(\alpha y) G(y) dy = |\alpha| H''(\alpha). \quad (2.18)$$

This implies that for all  $a > 1$  and all  $x \in \mathbb{R}$

$$|(H')^{-1}(x)| \leq |(H')^{-1}(ax)| \leq a |(H')^{-1}(x)|. \quad (2.19)$$

Indeed,  $H'$  is an increasing homeomorphism from  $\mathbb{R}$  to itself and, since  $G$  is symmetric,  $H'(x)$  and  $(H')^{-1}(x)$  have the same sign as  $x$ , so the first inequality is clear and it is enough to check the second inequality of (2.19) for  $x > 0$ . In this case, we have

$$\begin{aligned} \ln(H')^{-1}(ax) - \ln(H')^{-1}(x) &= \int_1^a \frac{x((H')^{-1})'(rx)}{(H')^{-1}(rx)} dr = \int_1^a \frac{x}{(H')^{-1}(rx)H''((H')^{-1}(rx))} dr \\ &\leq \int_1^a \frac{dr}{r} = \ln a, \end{aligned}$$

using (2.18) with  $\alpha = (H')^{-1}(rx)$ . Hence (2.19) is proved.

Using the fact that there exists some  $x_0 > 0$  such that, for all  $x \geq x_0$  or  $x \leq -x_0$ ,

$$e^x - 1 \leq \frac{x}{2} e^x,$$

and thus, using that  $xe^x \geq -e^{-1}$  for all  $x \in \mathbb{R}$ , we deduce that for all  $x \in \mathbb{R}$ ,

$$e^x - 1 \leq \frac{x}{2} e^x + \frac{e^{-1}}{2} + e^{x_0} - 1.$$

This implies that, for all  $\alpha \in \mathbb{R}$ ,

$$H(\alpha) \leq \frac{\alpha}{2} H'(\alpha) + \frac{e^{-1}}{2} + e^{x_0} - 1.$$

Since  $H(\alpha) \geq 0$  for all  $\alpha$ , we deduce that, for all  $x, \beta \in \mathbb{R}$ ,

$$\beta(H')^{-1} \left( \frac{\beta}{p(x)} \right) \geq L(x, \beta) \geq \frac{\beta}{2} (H')^{-1} \left( \frac{\beta}{p(x)} \right) - \bar{p} \left( \frac{e^{-1}}{2} + e^{x_0} - 1 \right). \quad (2.20)$$

Then, it follows from (2.19) that

$$|\beta|(H')^{-1} \left( \frac{|\beta|}{\bar{p}} \right) \leq |\beta|(H')^{-1} \left( \frac{|\beta|}{p(x)} \right) = \beta(H')^{-1} \left( \frac{\beta}{p(x)} \right) \leq \frac{\bar{p}}{\underline{p}} |\beta|(H')^{-1} \left( \frac{|\beta|}{\bar{p}} \right).$$

Therefore, Point (ii) follows.  $\square$

### 3 Large deviations estimates on $\mathbb{E}(N_t^{K,A})$

In order to prove the large deviations estimates on  $N_t^{K,A}$  of Theorems 1.1 and 1.2, we first need large deviations estimates on  $\mathbb{E}(N_t^{K,A})$ .

**Theorem 3.1.** *For any  $t > 0$  and for any closed  $A \subset \mathbb{D}[0, t]$ ,*

$$\limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log (\mathbb{E}[N_t^{K,A}]) \leq \sup_{f \in A} F_t(f), \quad (3.1)$$

and for any open  $G \subset \mathbb{D}[0, t]$

$$\liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log (\mathbb{E}[N_t^{K,G}]) \geq \sup_{f \in G} F_t(f), \quad (3.2)$$

with the usual convention that  $\sup \emptyset = -\infty$ .

Note that, in contrast with Theorems 1.1 and 1.2, there is no state constraint in the right-hand sides of (3.1) and (3.2). This shows the fundamental difference between the behavior of the stochastic process and its expectation, already observed in [7, 6].

*Proof of the upper bound in Theorem 3.1.* Let  $A \subset \mathbb{D}[0, t]$  be closed. Assume first that

$$\sup_{f \in A} F_t(f) > -\infty.$$

Since  $\beta_0(x) \rightarrow -\infty$  when  $x \rightarrow \pm\infty$  and  $\beta_0^K$  converges to  $\beta_0$  for the uniform norm, there exists a compact subset  $B$  of  $\mathbb{R}$  such that

$$\forall x \in B^c, \quad \beta_0^K(x) \leq -t(\bar{b} + \bar{p}) - 1 + \sup_{f \in A} F_t(f). \quad (3.3)$$

Let  $A_0$  be the compact set given by Lemma 2.4 with this choice of  $B$  and with the constant

$$M = \sup_{K \geq 1} \sup_{x \in \mathbb{R}} \beta_0^K(x) + (\bar{b} + \bar{p})t + 1 - \sup_{f \in A} F_t(f). \quad (3.4)$$

Fix  $\varepsilon > 0$  and  $f \in A \cap A_0$ . Since  $\beta_0, b, d$  and  $p$  are continuous, there exists  $\delta_f > 0$  such that, for all  $K$  large enough, for all  $g \in B_{\delta_f}(f)$ ,

$$\beta_0^K(g(0)) + \int_0^t R(g(s))ds \leq \beta_0(f(0)) + \int_0^t R(f(s))ds + \varepsilon.$$

In addition, since the function  $I_t$  is lower semi-continuous on  $\mathbb{D}[0, t]$  (cf. [28, Thm. 3.2.1]), we can also assume reducing  $\delta_f > 0$  if necessary that, for all  $x \in [f(0) - \delta_f, f(0) + \delta_f]$  and all  $g \in B_{2\delta_f}(f)$ ,

$$I_{t,x}(g) \geq I_{t,f(0)}(f) - \varepsilon. \quad (3.5)$$

We introduce for all  $k \geq 1$  the function  $\phi_{k,f}$  on  $\mathbb{D}[0, t]$  defined for all  $g \in \mathbb{D}[0, t]$  by

$$\phi_{k,f}(g) = k [(d_{\text{Sko}}(g, B_{\delta_f}(f))) \wedge 1].$$

Then

$$\begin{aligned} \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,T}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_{\delta_f}(f)} \right] dx \\ \leq K^{\beta_0(f(0)) + \int_0^t R(f(s))ds + \varepsilon} \int_{f(0) - \delta_f}^{f(0) + \delta_f} \mu_{x,t}^K(B_{\delta_f}(f)) dx \\ \leq 2\delta_f K^{\beta_0(f(0)) + \int_0^t R(f(s))ds + \varepsilon} \sup_{x \in [f(0) - \delta_f, f(0) + \delta_f]} \mathbb{E}_{\mu_{x,t}^K} (K^{-\phi_{k,f}}). \end{aligned} \quad (3.6)$$

Since the function  $\phi_{k,f}$  is bounded continuous, it follows from the Laplace principle uniform on compact sets of Theorem 2.3 that

$$\begin{aligned} \limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \left( \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,T}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_{\delta_f}(f)} \right] dx \right) \\ \leq \beta_0(f(0)) + \int_0^t R(f(s))ds + \varepsilon - \inf_{g \in \mathbb{D}[0,t]} (\phi_{k,f}(g) + I_{t,x}(g)). \end{aligned}$$

Now, using the definition of  $\phi_k$  and (3.5)

$$\inf_{g \in \mathbb{D}[0,t]} (\phi_{k,f}(g) + I_{t,x}(g)) \geq \inf_{g \notin B_{2\delta_f}(f)} \phi_{k,f}(g) \wedge \inf_{g \in B_{2\delta_f}(f)} I_{t,x}(g) \geq ((\delta_f \wedge 1)k) \wedge (I_t(f) - \varepsilon).$$

Since  $k$  was arbitrary, choosing it large enough entails

$$\begin{aligned} \limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \left( \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,T}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_{\delta_f}(f)} \right] dx \right) \\ \leq F_t(f) + 2\varepsilon. \end{aligned} \quad (3.7)$$

Since  $A \cap A_0$  is compact, there exists  $m < \infty$  and  $f_1, \dots, f_m \in A \cap A_0$  such that

$$A \cap A_0 \subset \bigcup_{j=1}^m B_{\delta_{f_j}}(f_j).$$

Now, it follows from Proposition 2.1 that

$$\begin{aligned} \mathbb{E}N_t^{K,A} &\leq \sum_{j=1}^m \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,T}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_{\delta_{f_j}}(f_j)} \right] dx \\ &\quad + \int_{B^c} K^{\beta_0^K(x) + (\bar{b} + \bar{p})t} \mu_{x,T}^K(A_0^c) dx + \text{Leb}(B) K^{\sup_{y \in \mathbb{R}} \beta_0^K(y) + (\bar{b} + \bar{p})t} \sup_{x \in B} \mu_{x,T}^K(A_0^c). \end{aligned}$$

Therefore, using Lemma 2.4 and Eq. (1.5), (3.3), (3.4) and (3.7),

$$\begin{aligned} \limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \mathbb{E}N_t^{K,A} &\leq \max\{F_t(f_j) + 2\varepsilon, 1 \leq j \leq m\} \vee \left( \sup_{f \in A} F_t(f) - 1 \right) \\ &\quad \vee \left( \sup_{y \in \mathbb{R}, K \geq 1} \beta_0^K(y) + (\bar{b} + \bar{p})t - M \right) \leq \sup_{f \in A} F_t(f) + 2\varepsilon. \end{aligned}$$

Since  $\varepsilon > 0$  was arbitrary, the proof is completed in the case where  $\sup_{f \in A} F_t(f) > -\infty$ .

In the case where  $\sup_{f \in A} F_t(f) = -\infty$ , let  $C > 0$  be fixed. For all  $f \in A$ , there exists  $\delta_f > 0$  such that, for all  $x \in [f(0) - \delta_f, f(0) + \delta_f]$  and all  $g \in B_{2\delta_f}(f)$ ,  $I_{t,x}(g) \geq M$ , where

$$M = \sup_{K \geq 1} \sup_{x \in \mathbb{R}} \beta_0^K(x) + (\bar{b} + \bar{p})t + C.$$

Following the same argument as in (3.6) (with  $k = 1$ ), we deduce that

$$\limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \left( \int_{\mathbb{R}} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{x,T}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_{\delta_f}(f)} \right] dx \right) \leq -C$$

We conclude as above that

$$\limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \mathbb{E}N_t^{K,A} \leq -C.$$

Since  $C > 0$  was arbitrary, the proof of the upper bound in Theorem 3.1 is complete.  $\square$

*Proof of the lower bound in Theorem 3.1.* We fix  $t > 0$  and  $G \subset \mathbb{D}[0, t]$  open. The proof is divided into 3 steps.

*Step 1.* We first prove using the Laplace principle uniform on compacts the following property: for all  $f \in \mathbb{D}[0, t]$  such that  $I_{t,f(0)}(f) < +\infty$  and all  $\delta > 0$ ,

$$\liminf_{K \rightarrow \infty} \inf_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \frac{1}{\log K} \log \mu_{t,x}^K(B_\delta(f)) \geq - \sup_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \inf_{g \in B_{2\delta/3}(f)} I_{t,x}(g). \quad (3.8)$$

For all integer  $k \geq 1$ , we define the function  $\phi_k$  on  $\mathbb{D}[0, t]$  as: for all  $g \in \mathbb{D}[0, t]$

$$\phi_k(g) = k \left( \frac{3}{\delta} d_{\text{sko}}(g, B_{2\delta/3}(f)) \wedge 1 \right).$$

Since  $\phi_k$  is bounded continuous, it follows from the Laplace principle uniform on compacts that

$$\begin{aligned} \liminf_{K \rightarrow \infty} \inf_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \frac{1}{\log K} \log \mathbb{E}_{\mu_{t,x}^K} (K^{-\phi_k}) &\geq - \sup_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \inf_{g \in \mathbb{D}[0, t]} (\phi_k(g) + I_{t,x}(g)) \\ &\geq - \sup_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \inf_{g \in B_{2\delta/3}(f)} I_{t,x}(g), \end{aligned}$$

where we used that  $\phi_k = 0$  on  $B_{2\delta/3}(f)$ . We observe that, for all  $x \in \mathbb{R}$ ,

$$\mathbb{E}_{\mu_{t,x}^K}(K^{-\phi_k}) \leq K^{-k} + \mu_{t,x}^K(B_\delta(f)),$$

so we deduce that

$$\begin{aligned} \max \left\{ \liminf_{K \rightarrow \infty} \inf_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \frac{1}{\log K} \log \mu_{t,x}^K(B_\delta(f)); -k \right\} \\ \geq - \sup_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \inf_{g \in B_{2\delta/3}(f)} I_{t,x}(g). \end{aligned}$$

Letting  $k \rightarrow \infty$  yields (3.8).

*Step 2.* We now prove the following weak continuity property of  $I_{t,x}$ : for all  $f \in \mathbb{D}[0, t]$  such that  $I_{t,f(0)}(f) < +\infty$  and all  $\varepsilon > 0$ , there exists  $\delta > 0$  such that, for all  $x \in [f(0) - \delta/2, f(0) + \delta/2]$ , there exists  $g \in B_\delta(f)$  with  $g(0) = x$  such that  $I_{t,x}(g) \leq I_{t,f(0)}(f) + \varepsilon$ .

To prove this, we recall from Step 1 of the proof of Lemma 2.5 (ii) that  $I_{t,f(0)}(f) < +\infty$  implies that  $|\dot{f}|(H')^{-1}(|\dot{f}|/\bar{p}) \in L^1[0, t]$ . Defining for all  $x \in \mathbb{R}$  and  $s \in [0, t]$   $f_s^{(x)} = x - f_0 + f_s$  and observing that  $\dot{f}^{(x)} = \dot{f}$ , it then follows from (2.19) and (2.20) that there exists a constant  $C$  such that, for all  $x$  in a neighborhood of  $f(0)$ ,  $L(f_s^{(x)}, \dot{f}_s^{(x)}) \leq C|\dot{f}|(H')^{-1}(|\dot{f}|/\bar{p})$ . Since  $L(\cdot, \cdot)$  is continuous with respect to both variables, we deduce from Lebesgue's theorem that  $I_{t,x}(f^{(x)}) \rightarrow I_{t,f(0)}(f)$  when  $x \rightarrow f(0)$ , hence the result.

*Step 3.* We now conclude the proof as follows: first, if  $\inf_{g \in G} I_{t,g(0)}(g) = +\infty$ , there is nothing to prove. So assume the converse and fix  $\varepsilon > 0$ . Then  $\sup_{g \in G} F_t(g) > -\infty$ , so we can take  $f \in G$  such that

$$F_t(f) \geq \sup_{g \in G} F_t(g) - \varepsilon.$$

Since  $G$  is open and in view of Steps 1 and 2, there exists  $\delta > 0$  such that  $B_\delta(f) \subset G$  and

$$\liminf_{K \rightarrow \infty} \inf_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \frac{1}{\log K} \log \mu_{t,x}^K(B_\delta(f)) \geq -I_{t,f(0)}(f) - \varepsilon.$$

Reducing  $\delta > 0$  if necessary, we can assume that, for  $K$  large enough, for all  $g \in B_\delta(f)$ ,

$$\beta_0^K(g(0)) + \int_0^t R(g(s))ds \geq \beta_0^K(f(0)) + \int_0^t R(f(s))ds - \varepsilon.$$

Now, the Feynman-Kac formula of Proposition 2.1 implies that

$$\begin{aligned} & \frac{1}{\log K} \log \mathbb{E} N_t^{K,G} \\ & \geq \frac{1}{\log K} \log \int_{f(0) - \delta/3}^{f(0) + \delta/3} K^{\beta_0^K(x)} \mathbb{E}_{\mu_{t,x}^K} \left[ \exp \left( \log K \int_0^t R(Y_s^K) ds \right) \mathbb{1}_{(Y_s^K)_{s \in [0,t]} \in B_\delta(f)} \right] \\ & \geq \beta_0^K(f(0)) + \int_0^t R(f(s))ds - \varepsilon + \frac{1}{\log K} \log \left( \frac{2\delta}{3} \inf_{x \in [f(0) - \delta/3, f(0) + \delta/3]} \mu_{t,x}^K(B_\delta(f)) \right). \end{aligned}$$

Therefore,

$$\liminf_{K \rightarrow \infty} \frac{1}{\log K} \log \mathbb{E} N_t^{K,G} \geq F_t(f) - 2\varepsilon \geq \sup_{g \in G} F_t(g) - 3\varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, the lower bound in Theorem 3.1 is proved.  $\square$

## 4 Proof of Theorem 1.1

The proof relies on the following proposition.

**Proposition 4.1.** *For any  $t > 0$  and all closed set  $A \subset \mathbb{D}[0, t]$ , in probability*

$$\limsup_{K \rightarrow \infty} \frac{\log N_t^{K,A}}{\log K} \leq \sup\{F_t(f); f \in A\}.$$

*Proof.* We introduce the notation

$$F_t(A) = \sup\{F_t(f); f \in A\}.$$

Using Markov inequality and the upper bound in Theorem 3.1, we have for all fixed  $\delta > 0$

$$\limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \mathbb{P}(N_t^{K,A} \geq K^{F_t(A)+\delta}) \leq \limsup_{K \rightarrow \infty} \frac{1}{\log K} \log \left( \frac{\mathbb{E}(N_t^{K,A})}{K^{F_t(A)+\delta}} \right) \leq -\delta.$$

Therefore, for  $K$  large enough,

$$\mathbb{P}(N_t^{K,A} \geq K^{F_t(A)+\delta}) \leq K^{-\delta/2}.$$

Hence it follows that, in probability,

$$\limsup_{K \rightarrow \infty} \frac{\log N_t^{K,A}}{\log K} \leq F_t(A) + \delta.$$

Since  $\delta$  was arbitrary, the proof is complete.  $\square$

*Proof of Theorem 1.1. Step 1: Proof for  $A$  compact.*

Fix  $k \geq 1$ . For any  $f \in \mathbb{D}[0, t]$  such that  $I_t(f) < \infty$ , there exists  $\delta_f > 0$  such that, for all  $g \in \text{adh}(B_{\delta_f}(f))$ ,

$$\forall s \in [0, t], \quad F_s(g) \leq F_s(f) + \frac{1}{k}. \quad (4.1)$$

Indeed, if it was not true, there would exist  $s_n \in [0, t]$  and  $g_n \in \mathbb{D}[0, t]$  converging to  $f$  such that  $F_{s_n}(g_n) > F_{s_n}(f) + 1/k$ . After extraction, we can assume that  $s_n \rightarrow s$ . We then deduce from the lower semi-continuity of  $I_{(s-\delta)\vee 0}$  (see [28, Thm. 3.2.1]) that, for all  $\delta > 0$ ,

$$\begin{aligned} F_{(s-\delta)\vee 0}(f) &\geq \limsup_n F_{(s-\delta)\vee 0}(g_n) \\ &\geq \limsup_n F_{s_n}(g_n) - \bar{R}\delta \\ &\geq \limsup_n F_{s_n}(f) - \bar{R}\delta + \frac{1}{k} = F_s(f) - \bar{R}\delta + \frac{1}{k}. \end{aligned}$$

Therefore,

$$F_s(f) = \lim_{\delta \rightarrow 0} F_{(s-\delta)\vee 0}(f) \geq F_s(f) + \frac{1}{k},$$

which is absurd.

Similarly, if  $f \in \mathbb{D}[0, t]$  is such that  $I_t(f) = +\infty$ , there exists  $\delta_f > 0$  such that, for all  $g \in \text{adh}(B_{\delta_f}(f))$ ,

$$\forall s \in [0, t], \quad F_s(g) \leq -k.$$

Now, assume that  $A$  is compact. Since

$$A \subset \bigcup_{f \in A} B_{\delta_f}(f),$$

there exists  $n \geq 1$  and  $f_1, \dots, f_n \in A$  such that, denoting  $\delta_i = \delta_{f_i}$ ,

$$A \subset \bigcup_{i=1}^n B_{\delta_i}(f_i).$$

Given  $i \in \{1, \dots, n\}$ , if there exists  $s \leq t$  such that  $F_s(f_i) < -1/k$ , we can apply Proposition 4.1 to show that, in probability,

$$\limsup_{K \rightarrow \infty} \frac{\log N_s^{K, \text{cl}(B_{\delta_i}(f_i))}}{\log K} \leq F_s(f_i) + 1/k < 0.$$

Since  $N_s^{K, \text{cl}(B_{\delta_i}(f_i))}$  is integer-valued, this implies that,  $N_s^{K, \text{cl}(B_{\delta_i}(f_i))} = 0$  for  $K$  large enough, hence with probability converging to one,

$$\frac{\log N_s^{K, \text{cl}(B_{\delta_i}(f_i))}}{\log K} = -\infty$$

and, since  $t \geq s$ ,

$$\frac{\log N_t^{K, \text{cl}(B_{\delta_i}(f_i))}}{\log K} = -\infty.$$

If  $F_t(f_i) > -\infty$  and for all  $s \leq t$ ,  $F_s(f_i) \geq -1/k$ , we deduce from Proposition 4.1 that, in probability,

$$\limsup_{K \rightarrow \infty} \frac{\log N_t^{K, \text{cl}(B_{\delta_i}(f_i))}}{\log K} \leq F_t(f_i) + 1/k.$$

Finally, if  $F_t(f_i) = -\infty$ , we deduce from Proposition 4.1 that, in probability,

$$\limsup_{K \rightarrow \infty} \frac{\log N_t^{K, \text{cl}(B_{\delta_i}(f_i))}}{\log K} \leq -k.$$

Observing that

$$N_t^{K, A} \leq \sum_{i=1}^n N_t^{K, \text{cl}(B_{\delta_i}(f_i))},$$

we obtain in probability

$$\begin{aligned} \limsup_{K \rightarrow \infty} \frac{\log N_t^{K, A}}{\log K} &\leq \sup \left\{ F_t(f_i) + \frac{1}{k}, i \in \{1, \dots, n\} \text{ s.t. } F_s(f_i) \geq -\frac{1}{k} \forall s \in [0, t] \right\} \vee (-k) \\ &\leq \left( \frac{1}{k} + \sup \left\{ F_t(f), f \in A \text{ s.t. } F_s(f) \geq -\frac{1}{k} \forall s \in [0, t] \right\} \right) \vee (-k). \end{aligned} \quad (4.2)$$

Now, when  $k$  converges to  $+\infty$ ,

$$\sup \left\{ F_t(f), f \in A \text{ s.t. } F_s(f) \geq -\frac{1}{k} \forall s \in [0, t] \right\} \rightarrow \sup \{ F_t(f), f \in A \text{ s.t. } F_s(f) \geq 0 \forall s \in [0, t] \},$$

since otherwise, there would exist  $\eta > 0$  and a sequence  $(f_k)_{k \geq 1}$  in  $A$  such that  $F_s(f_k) \geq -1/k$  for all  $s \in [0, t]$  and  $F_t(f_k) \geq \sup\{F_t(f), f \in A, F_s(f) \geq 0 \forall s \in [0, t]\} + \eta$ . Since  $A$  is compact, we can assume after extraction that  $f_k$  converges to some  $g \in A$ . Since  $F_s$  is upper semi-continuous for all  $s \in [0, t]$ ,  $F_s(g) \geq \limsup_k F_s(f_k) \geq 0$  for all  $s \in [0, t]$  and  $F_t(g) \geq \sup\{F_t(f), f \in A, F_s(f) \geq 0 \forall s \in [0, t]\} + \eta$ , which is a contradiction.

Hence, letting  $k \rightarrow \infty$  in (4.2) ends the proof of Theorem 1.1 for  $A$  compact.

*Step 2: Proof of Theorem 1.1 for  $A$  closed.* Following standard arguments in large deviations, to conclude, it only remains to prove that, for all  $M_0$ , there exists a compact set  $C_{M_0} \subset \mathbb{D}[0, t]$  such that, almost surely

$$\limsup_{K \rightarrow +\infty} \frac{\log N_t^{K, C_{M_0}^c}}{\log K} \leq -M_0. \quad (4.3)$$

This can be deduced from Proposition 2.1 and Lemma 2.4 as follows: let  $C_{M_0} \subset \mathbb{D}[0, t]$  be the compact set of Lemma 2.4 for  $B = \{x \in \mathbb{R}, \beta_0(x) \geq -2 - t\bar{R} - M_0\}$  (which is compact by Assumption (1.5)), and  $M = M_0 + t\bar{R} + \bar{\beta} + 2$ , i.e.

$$\limsup_{K \rightarrow \infty} \sup_{x \in B} \frac{1}{\log K} \log \mu_{x,t}^K(C_{M_0}^c) \leq -M_0 - \bar{\beta} - t\bar{R} - 2.$$

By Proposition 2.1, for  $K$  large enough,

$$\mathbb{E} N_t^{K, C_{M_0}^c} \leq \int_{B^c} K^{(-2-t\bar{R}-M_0) \wedge (\bar{\beta} - \alpha|x|)} K^{t\bar{R}} dx + \text{Leb}(B) \sup_{x \in B} K^{\bar{\beta} + t\bar{R} + \frac{1}{2}} \mu_{x,t}^K(C_{M_0}^c),$$

so

$$\limsup_{K \rightarrow \infty} \frac{\log \mathbb{E} N_t^{K, C_{M_0}^c}}{\log K} \leq -M_0 - \frac{3}{2}.$$

To conclude, we proceed as in the proof of Proposition 4.1:

$$\mathbb{P} \left( N_t^{K, C_{M_0}^c} \geq K^{-M_0} \right) \leq \frac{\mathbb{E} N_t^{K, C_{M_0}^c}}{K^{-M_0}} \leq K^{-5/4}$$

for  $K$  large enough. Then Borel-Cantelli's lemma entails (4.3).  $\square$

## 5 Proof of Theorem 1.2

We follow the general approach of [6], using moment estimates following ideas in [40]. We start with considering for  $t \geq 0$ ,  $\varepsilon > 0$  and  $f \in \mathbb{D}[0, t]$  such that  $I_t(f) < +\infty$ . Recall that  $B_\varepsilon(f)$  is the open ball  $\{g \in \mathbb{D}[0, t], d_{\text{SkO}}(g, f) < \varepsilon\}$ . Let us denote by

$$N_t^{K, \varepsilon, f} = N_t^{K, B_\varepsilon(f)} = \sum_{u \in V_{t \log K}^K} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K, u}, s \leq t),$$

the number of particles remaining in the tube of width  $\varepsilon$  around  $f$  until time  $t$ . The key ingredient of the proof is the following lemma.

**Lemma 5.1.** *Let us consider  $t \geq 0$  and  $f \in AC[0, t]$  such that  $I_t(f) < +\infty$  and such that for all  $s \leq t$ ,  $F_s(f) > 0$ . Almost surely, for all sufficiently small  $\varepsilon > 0$ ,*

$$\liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N^{K, \varepsilon, f} \geq F_t(f). \quad (5.1)$$

Assuming Lemma 5.1, we can then complete the proof of Theorem 1.2 as follows. Let  $t > 0$  and  $G$  be an open subset of  $\mathbb{D}[0, t]$ . If

$$\sup\{F_t(g); g \in G, \forall s \in [0, t], F_s(g) > 0\} = -\infty,$$

then we have nothing to prove. Otherwise, for any  $\delta > 0$ , we can find  $f \in G$  such that  $F_s(f) > 0$  for all  $s \in [0, t]$  and

$$F_t(f) \geq \sup\{F_t(g); g \in G, \forall s \in [0, t], F_s(g) > 0\} - \delta.$$

Since  $G$  is open, there exists  $\varepsilon > 0$  such that  $B_\varepsilon(f) \subset G$ . Hence, reducing  $\varepsilon > 0$  if necessary, we deduce from Lemma 5.1 that, almost surely,

$$\begin{aligned} \liminf_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, G} &\geq \liminf_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K, \varepsilon, f} \\ &\geq F_t(f) \\ &\geq \sup\{F_t(g); g \in G, \forall s \in [0, t], F_s(g) > 0\} - \delta. \end{aligned}$$

Since  $\delta > 0$  was arbitrary, the proof of Theorem 1.2 is completed.

The proof of Lemma 5.1 is divided into two subsections. In Section 5.1, we establish an upper bound of  $\mathbb{E}[(N_t^{K, \varepsilon, f})^2]$  with the idea of [40] in mind, by establishing a many-to-two formula (or formula for forks). We conclude in Section 5.2.

### 5.1 Estimates for the second moment of $N^{K, \varepsilon, f}$

A key ingredient to compare  $N_t^{K, \varepsilon, f}$  with its expectation is to control its second moment. For this purpose, we first establish the next lemma.

**Lemma 5.2.** *Let us consider  $t \geq 0$  and  $f \in AC[0, t]$  such that  $I_t(f) < +\infty$  and such that for all  $s \leq t$ ,  $F_s(f) > 0$ . For all  $\delta' > 0$ , for all sufficiently small  $\varepsilon > 0$  and sufficiently large  $K$ , we have*

$$\mathbb{E}[(N_t^{K, \varepsilon, f})^2] \leq \mathbb{E}^2[N_t^{K, \varepsilon, f}] \times K^{8\delta'}. \quad (5.2)$$

*Proof.* Recall that  $V_t^K$  is the collection of labels of individuals alive at time  $t$ , and that  $V_{[0, t]}^K$  consists of the labels of individuals alive between times 0 and  $t$ .

First, we have  $\mathbb{E}[(N_t^{K,\varepsilon,f})^2] = A + B + C$  with

$$\begin{aligned} A &= \mathbb{E} \left( \sum_{v \in V_{t \log K}^K} \mathbb{1}_{B_\varepsilon(f)}^2(X_{s \log K}^{K,v}, s \leq t) \right) \\ B &= \mathbb{E} \left( \sum_{\substack{v_1 \neq v_2, v_1, v_2 \in V_{t \log K}^K \\ \exists w \in V_{[0, t \log K]}^K, w \prec v_1, w \prec v_2}} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_1}, s \leq t) \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_2}, s \leq t) \right) \\ C &= \mathbb{E} \left( \sum_{\substack{v_1 \neq v_2, v_1, v_2 \in V_{t \log K}^K \\ \nexists w \in V_{[0, t \log K]}^K, w \prec v_1, w \prec v_2}} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_1}, s \leq t) \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_2}, s \leq t) \right). \end{aligned}$$

In  $B$ , the two individuals  $v_1$  and  $v_2$  have a common ancestor  $w \in V_{[0, t \log K]}^K$  and hence a common ancestor at time 0, whereas in  $C$ , the two individuals are descended from different ancestors at time 0.

**Term A.** Let  $\delta' > 0$ . For the term  $A$ , we notice that  $A = \mathbb{E}(N_t^{K,\varepsilon,f})$ . Then, using Theorem 3.1, we obtain that for  $\varepsilon$  small enough and  $K$  large enough,

$$K^{F_t(f) - \delta'} \leq K^{\sup_{g \in B_\varepsilon(f)} F_t(g) - \delta'} \leq A \leq K^{\sup_{g \in \overline{B}_\varepsilon(f)} F_t(g) + \delta'} \leq K^{F_t(f) + 2\delta'}, \quad (5.3)$$

where  $\overline{B}_\varepsilon(f)$  denotes here the closure in  $\mathbb{D}$  of  $B_\varepsilon(f)$ , and where we used (3.5) for the right-most inequality.

**Term B.** Summing over all pair of leaves  $(v_1, v_2) \in (V_{t \log K}^K)^2$ ,  $v_1 \neq v_2$ , that have a common ancestor at time 0 amounts to summing over all the internal nodes  $w$  of the tree, and then over all pairs made of one descendant at  $t \log K$  of  $w_0$  and one of  $w_1$ . Then, denoting by  $\tilde{S}_w \log K$  the time at which the individual  $w$  gives birth,

$$B = \mathbb{E} \left[ \sum_{\substack{w \in V_{[0, t \log K]}^K \\ \tilde{S}_w < t}} \left( \sum_{\substack{v_1 \in V_{t \log K}^K \\ v_1 \succeq w_0}} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_1}, s \leq t) \right) \times \left( \sum_{\substack{v_2 \in V_{t \log K}^K \\ v_2 \succeq w_1}} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K,v_2}, s \leq t) \right) \right].$$

Conditionally on  $\mathcal{F}_{\tilde{S}_w \log K}$ , the two populations descending from  $w_0$  and  $w_1$  are independent. The idea is to condition on  $\mathcal{F}_{\tilde{S}_w \log K}$  and consider the time intervals  $[0, \tilde{S}_w \log K]$  and  $[\tilde{S}_w \log K, t \log K]$ . It is then natural to introduce, for any time  $s \in [0, t]$  and  $f \in AC[0, t]$ , the notation

$$F_{s,t}(f) := F_t(f) - F_s(f) = \int_s^t R(f(u)) du - I_{s,t}(f), \quad (5.4)$$

where  $I_{s,t}(f) = I_t(f) - I_s(f) = \int_s^t L(f_u, \dot{f}_u) du$  can be interpreted as the rate function of the large deviations for  $Y^K$  on the time interval  $[s, t]$ .

A difficulty arises from the Skorohod distance as,

$$d_{\text{Sko}}(f, g) \leq \max \left( d_{\text{Sko}}(f|_{[0,s]}, g|_{[0,s]}), d_{\text{Sko}}(f|_{[s,t]}, g|_{[s,t]}) \right), \quad (5.5)$$

with an inequality that can be strict as the right hand side can be associated only with a time deformation  $\lambda$  in (1.1) that keeps  $s$  fixed. This implies that

$$B_\varepsilon(f) \supset \{g \in \mathbb{D}[0, t] : g|_{[0, s]} \in B_\varepsilon(f|_{[0, s]})\} \cap \{g \in \mathbb{D}[0, t] : g|_{[s, t]} \in B_\varepsilon(f|_{[s, t]})\},$$

and not the reverse inclusion, which raises difficulties when separating the time interval  $[0, t]$  into  $[0, s]$  and  $[s, t]$ .

However, we will be only interested in the case where  $f \in AC[0, t]$ , for which there exists a constant  $\eta(\varepsilon)$  depending on  $\varepsilon$ ,  $f$  and  $t$  (the dependencies on  $f$  and  $t$  are omitted in the notation for the sake of readability), such that

$$B_\varepsilon(f) \subset B_{\eta(\varepsilon)}^\infty(f) := \{g \in \mathbb{D}[0, t], \sup_{r \in [0, t]} |f(r) - g(r)| < \eta(\varepsilon)\}. \quad (5.6)$$

This inclusion is proved in Appendix B. Now, for the uniform norm, we have an equality in (5.5) and

$$B_{\eta(\varepsilon)}^\infty(f) = \left\{g \in \mathbb{D}[0, t] : g|_{[0, s]} \in B_{\eta(\varepsilon)}^\infty(f|_{[0, s]})\right\} \cap \left\{g \in \mathbb{D}[0, t] : g|_{[s, t]} \in B_{\eta(\varepsilon)}^\infty(f|_{[s, t]})\right\}. \quad (5.7)$$

With these notations, we have on the set  $\{\tilde{S}_w < t\}$ :

$$\begin{aligned} & \mathbb{E} \left( \sum_{\substack{v \in V_{t \log K}^K \\ v \succeq w 0}} \mathbb{1}_{B_\varepsilon(f)}(X_{s \log K}^{K, v}, s \leq t) \middle| \mathcal{F}_{\tilde{S}_w \log K} \right) \\ & \leq \mathbb{E} \left( \sum_{\substack{v \in V_{t \log K}^K \\ v \succeq w 0}} \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f)}(X_{s \log K}^{K, v}, s \leq t) \middle| \mathcal{F}_{\tilde{S}_w \log K} \right) \\ & = \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K, w}, r \leq \tilde{S}_w) \\ & \quad \times \mathbb{E}_{\delta_{X_{\tilde{S}_w \log K}^{K, w 0}}} \left( \sum_{v \in V_{(t-s) \log K}^K} \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f(s+\cdot))}(X_{r \log K}^{K, v}, r \leq t-s) \right) \Big|_{s=\tilde{S}_w} \\ & = \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K, w}, r \leq \tilde{S}_w) \\ & \quad \times \mathbb{E}_{X_{\tilde{S}_w \log K}^{K, w 0}} \left( K \int_0^{t-s} R(X_{r \log K}^K) dr \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f(s+\cdot))}(X_{r \log K}^K, r \leq t-s) \right) \Big|_{s=\tilde{S}_w} \\ & \leq \mathbb{1}_{B_{\eta(\varepsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K, w}, r \leq \tilde{S}_w) K \int_{\tilde{S}_w}^t R(f(r)) dr + C(f, R) t \eta(\varepsilon) \\ & \quad \times \mathbb{P}_{X_{\tilde{S}_w \log K}^{K, w 0}} \left( \sup_{r \leq t-s} |X_{r \log K}^K - f(s+r)| < \eta(\varepsilon) \right) \Big|_{s=\tilde{S}_w}, \end{aligned} \quad (5.8)$$

where we have used Proposition 2.1(ii) to obtain the fourth line, and where the constant  $C(f, R)$  is the Lipschitz norm of  $R$  on  $[\inf f - 1, \sup f + 1]$ .

We would like to use Theorem 2.3 to upper bound the probability appearing in the right hand side of (5.8). But because  $\tilde{S}_w$  is a random time, we have to discretize the time interval  $[0, t]$ . It will be useful to consider  $h > 0$  a small time mesh such that  $t/h \in \mathbb{N}$ . For any  $x \in \mathbb{R}$  and

$s < s' < t$ ,

$$\begin{aligned}
& \mathbb{P}_x \left( \sup_{r \leq t-s} |X_{r \log K}^K - f(s+r)| < \eta(\epsilon) \right) \\
& \leq \mathbb{E}_x \left( \mathbb{1}_{X_{(s'-s) \log K}^K \in [f(s')-\eta(\epsilon), f(s')+\eta(\epsilon)]} \mathbb{P}_{X_{(s'-s) \log K}^K} \left( \sup_{r \leq t-s'} |X_{r \log K}^K - f(s'+r)| < \eta(\epsilon) \right) \right) \\
& \leq \sup_{y \in [f(s')-\eta(\epsilon), f(s')+\eta(\epsilon)]} \mathbb{P}_y \left( \sup_{r \leq t-s'} |X_{r \log K}^K - f(s'+r)| < \eta(\epsilon) \right) \\
& \leq \sup_{y \in [f(s')-\eta(\epsilon), f(s')+\eta(\epsilon)]} \mu_{y, t-s'}^K(B_{\eta(\epsilon)}(f|_{[s', t]}))
\end{aligned}$$

where we notice that we have returned to a ball with respect to the Skorohod distance in the last line, permitting to use our large deviation result. Hence, for  $k \in \{0, \dots, t/h - 1\}$ , on the event  $\{\tilde{S}_w \in [kh, (k+1)h]\}$ ,

$$\begin{aligned}
& \mathbb{E} \left( \sum_{\substack{v \in V_{t \log K}^K \\ v \succeq w 0}} \mathbb{1}_{B_\epsilon(f)}(X_{s \log K}^{K,v}, s \leq t) \Big| \mathcal{F}_{\tilde{S}_w \log K} \right) \\
& \leq \mathbb{1}_{B_{\eta(\epsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K,w}, r \leq \tilde{S}_w) K^{\int_{\tilde{S}_w}^t R(f(r)) dr + C(f,R)t\eta(\epsilon)} \\
& \quad \times \sup_{y \in [f((k+1)h)-\eta(\epsilon), f((k+1)h)+\eta(\epsilon)]} \mu_{y, t-(k+1)h}^K(B_{\eta(\epsilon)}(f|_{[(k+1)h, t]})).
\end{aligned}$$

Using the Laplace principle uniformly on compact sets of Theorem 2.3 as in (2.11) and (2.13), we deduce that, for  $K$  large enough, for all  $y \in [f((k+1)h) - \eta(\epsilon), f((k+1)h) + \eta(\epsilon)]$ ,

$$\begin{aligned}
\frac{\log \mu_{y, t-(k+1)h}^K(B_{\eta(\epsilon)}(f|_{[(k+1)h, t]}))}{\log K} & \leq - \inf_{\substack{g \in \overline{B}_{\eta(\epsilon)}(f|_{[(k+1)h, t]}) \\ g(0) \in [f((k+1)h)-\eta(\epsilon), f((k+1)h)+\eta(\epsilon)]}} I_{t-(k+1)h}(g) \\
& \leq -I_{t-(k+1)h}(f|_{[(k+1)h, t]}) + \delta',
\end{aligned}$$

by the lower semi-continuity of  $I_{t-(k+1)h}$ , for  $\epsilon$  small enough. Hence, on the event  $\{\tilde{S}_w \in [kh, (k+1)h]\}$ ,

$$\begin{aligned}
& \mathbb{E} \left( \sum_{\substack{v \in V_{t \log K}^K \\ v \succeq w 0}} \mathbb{1}_{B_\epsilon(f)}(X_{s \log K}^{K,v}, s \leq t) \Big| \mathcal{F}_{\tilde{S}_w \log K} \right) \\
& \leq \mathbb{1}_{B_{\eta(\epsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K,w}, r \leq \tilde{S}_w) K^{\int_{\tilde{S}_w}^t R(f(r)) dr + C(f,R)t\eta(\epsilon) - I_{t-(k+1)h}(f|_{[(k+1)h, t]}) + \delta'} \\
& \leq \mathbb{1}_{B_{\eta(\epsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K,w}, r \leq \tilde{S}_w) K^{F_{\tilde{S}_w, t}(f) + C(f,R)t\eta(\epsilon) + \delta' + I_{kh, (k+1)h}(f)} \\
& \leq \mathbb{1}_{B_{\eta(\epsilon)}^\infty(f|_{[0, \tilde{S}_w])}}(X_{r \log K}^{K,w}, r \leq \tilde{S}_w) K^{F_{\tilde{S}_w, t}(f) + 2\delta'}, \tag{5.9}
\end{aligned}$$

reducing  $\epsilon > 0$  if needed and choosing  $h$  such that  $I_{kh, (k+1)h}(f) \leq \delta'/2$  for all  $k$ . This property holds true on the event  $\{\tilde{S}_w \in [kh, (k+1)h]\}$  for all  $K$  large enough and all  $\epsilon > 0$  small enough. Since there are only finitely many  $k \in \{0, \dots, t/h - 1\}$ , we deduce that (5.9) holds true almost surely for all  $K$  large enough and all  $\epsilon > 0$  small enough.

A similar expression holds for the population stemming from  $w_1$ . Thus, using Proposition 2.2,

$$\begin{aligned}
B &\leq \int_{\mathbb{R}} K^{\beta_0(x)} \mathbb{E}_{\delta_x} \left( \sum_{\substack{w \in V_{[0,t \log K]}^K \\ \tilde{S}_w < t}} \mathbb{1}_{\sup_{r \leq \tilde{S}_w} |X_{r \log K}^{K,w} - f(r)| < \varepsilon} K^{2F_{\tilde{S}_w,t}(f) + 4\delta'} \right) dx \\
&\leq \int_{f(0)-\varepsilon}^{f(0)+\varepsilon} K^{\beta_0(x)} \int_0^t \mathbb{E}_x \left( \mathbb{1}_{\sup_{r \leq s} |X_{r \log K}^K - f(r)| < \varepsilon} \right. \\
&\quad \left. \times (b+p+d)(X_{s \log K}^K) K^{2F_{s,t}(f) + \int_0^s R(X_{r \log K}^K) dr + 4\delta'} \right) ds dx \\
&\leq (\bar{b} + \bar{p} + \bar{d}) 2\varepsilon K^{\beta_0(f(0)) + C\varepsilon} \int_0^t K^{-I_s(f) + 2F_{s,t}(f) + \int_0^s R(f(r)) dr + 5\delta'} ds \\
&\leq K^{F_t(f) + \sup_{s \in [0,t]} F_{s,t}(f) + 6\delta'}, \tag{5.10}
\end{aligned}$$

for  $\varepsilon$  sufficiently small and  $K$  sufficiently large, and where we used (1.16) and (5.4) in the fourth line. Note that:

$$F_t(f) + \sup_{s \in [0,t]} F_{s,t}(f) = F_t(f) + \sup_{s \in [0,t]} (F_t(f) - F_s(f)) = 2F_t(f) - \inf_{s \in [0,t]} F_s(f).$$

Thus,

$$B \leq K^{2F_t(f) - \inf_{s \in [0,t]} F_s(f) + 6\delta'} \leq K^{2F_t(f) + 6\delta'}, \tag{5.11}$$

as  $\inf_{s \in [0,t]} F_s(f) \geq 0$  by the assumptions of Lemma 5.2.

**Term C.** By the branching property,

$$\begin{aligned}
C &= \mathbb{E} \left( \sum_{\substack{w_1 \neq w_2 \\ w_1, w_2 \in V_0^K}} \left( \sum_{\substack{v_1 \in V_{t \log K}^K \\ v_1 \succeq w_1}} \sum_{\substack{v_2 \in V_{t \log K}^K \\ v_2 \succeq w_2}} \mathbb{1}_{B_\varepsilon}(f)(X_{s \log K}^{K,v_1}, s \leq t) \mathbb{1}_{B_\varepsilon}(f)(X_{s \log K}^{K,v_2}, s \leq t) \right) \right) \\
&= \mathbb{E} \left( \sum_{\substack{w_1 \neq w_2 \\ w_1, w_2 \in V_0^K}} \mathbb{E}_{\delta_{X_0^{K,w_1}}} \left( \sum_{v_1 \in V_{t \log K}^K} \mathbb{1}_{B_\varepsilon}(f)(X_{s \log K}^{K,v_1}, s \leq t) \right) \right. \\
&\quad \left. \times \mathbb{E}_{\delta_{X_0^{K,w_2}}} \left( \sum_{v_2 \in V_{t \log K}^K} \mathbb{1}_{B_\varepsilon}(f)(X_{s \log K}^{K,v_2}, s \leq t) \right) \right) \\
&= \mathbb{E} \left( \sum_{\substack{w_1 \neq w_2 \\ w_1, w_2 \in V_0^K}} \mathbb{E}_{\delta_{X_0^{K,w_1}}} \left( N_t^{K,\varepsilon,f} \right) \mathbb{E}_{\delta_{X_0^{K,w_2}}} \left( N_t^{K,\varepsilon,f} \right) \right)
\end{aligned}$$

Using (5.3) for the internal expectations and accounting that we start from a single particle and not from a point process of intensity  $K^{\beta_0(x)} dx$ ,

$$\begin{aligned}
C &\leq \mathbb{E} \left( \iint_{(f(0)-\varepsilon, f(0)+\varepsilon)^2} \mathbb{1}_{x \neq y} K^{2F_t(f) - 2\beta_0(f(0)) + 4\delta'} Z_0^K(dx) Z_0^K(dy) \right) \\
&= K^{2F_t(f) - 2\beta_0(f(0)) + 4\delta'} \iint_{(f(0)-\varepsilon, f(0)+\varepsilon)^2} K^{\beta_0(x)} K^{\beta_0(y)} dx dy \leq K^{2F_t(f) + 5\delta'}, \tag{5.12}
\end{aligned}$$

where we used the multivariate Mecke's formula (e.g. [37, Theorem 4.4]) in the second line.

Gathering the upper bounds in (5.3), (5.11) and (5.12) gives that:

$$\mathbb{E}[(N_t^{K,\varepsilon,f})^2] \leq K^{2F_t(f)+6\delta'} \leq \mathbb{E}^2[N_t^{K,\varepsilon,f}] \times K^{8\delta'}, \quad (5.13)$$

where the last inequality comes from the lower bound in (5.3). This ends the proof of Lemma 5.2.  $\square$

Notice that if the initial condition has not the intensity  $K^{\beta_0(x)} dx$  but  $K^{\beta_0(x)-\delta''} dx$  for example, the proof above would provide instead of (5.13) that:

$$\mathbb{E}[(N_t^{K,\varepsilon,f})^2] \leq K^{2F_t(f)+6\delta'-2\delta''} \leq \mathbb{E}^2[N_t^{K,\varepsilon,f}] \times K^{8\delta'}. \quad (5.14)$$

For the last inequality, notice that  $\beta_0(x)$  appears in the definition of  $F_t(f)$  so that for this new initial condition  $\beta_0(x) - \delta''$ , the lower bound in (5.3) becomes:

$$K^{(\beta_0(x)-\delta'')+ \int_0^t R(f(s))ds - I_t(f) - \delta'} = K^{F_t(f) - \delta'' - \delta'} \leq \mathbb{E}(N_t^{K,\varepsilon,f}). \quad (5.15)$$

These inequalities will be useful in the following.

## 5.2 Proof of Lemma 5.1

We are now ready to prove Lemma 5.1. Let us introduce

$$\delta_0 = \min_{s \in [0,t]} \{F_s(f)\} > 0. \quad (5.16)$$

Let us fix  $\delta \in (0, \delta_0)$ . Our purpose is to prove that almost surely, for all sufficiently small  $\varepsilon > 0$ ,

$$\liminf_{K \rightarrow \infty} \frac{1}{\log K} \log N_t^{K,\varepsilon,f} > F_t(f) - \delta. \quad (5.17)$$

A direct use of Lemma 5.2 would only show that  $\frac{1}{\log K} \log N_t^{K,\varepsilon,f} > F_t(f) - \delta$  holds with a probability converging to 0 as  $K \rightarrow +\infty$  (see (5.20) below). To obtain the almost sure lower bound we make use of the branching property by dividing the initial population into several groups. For this, let us consider  $\delta'' \in (0, \delta_0)$  so that  $\beta_0(f(0)) - \delta'' > 0$ . This  $\delta''$  as well as the  $\delta'$  appearing in (5.3) will be fixed depending on  $\delta$  in the end of the proof, and  $\varepsilon$  (resp.  $K$ ) are chosen small (resp. large) enough according to these choices.

By the form of the initial condition (1.4), and by the superposition principle, we can write

$$Z_0^K(dx) = \sum_{i=1}^{\lfloor K^{\delta''} \rfloor} \tilde{Z}_0^{K,i}(dx)$$

where  $\tilde{Z}_0^{K,i}(dx)$  are i.i.d. Poisson point measures with the modified intensity measure  $(K^{\beta_0^K(x)}/\lfloor K^{\delta''} \rfloor)dx$ . This decomposition combined with the branching property will provide that (5.1) holds almost surely by Borel-Cantelli's lemma. For  $i \in \{1, \dots, \lfloor K^{\delta''} \rfloor\}$ , we will denote by  $\tilde{N}_t^{K,\varepsilon,f,i}$  the number of particles, among those started from  $\tilde{Z}_0^{K,i}$ , that remain in the tube of width  $\varepsilon$  around  $f$  until time  $t$ . Note that the random variables  $\tilde{N}_t^{K,\varepsilon,f,i}$  are i.i.d.

Let us consider the branching process started from  $\tilde{Z}_0^{K,1}$  and let  $\delta \in (0, \delta_0)$ . We have by the Cauchy-Schwarz inequality that

$$\mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \mathbb{1}_{\tilde{N}_t^{K,\varepsilon,f,1} > K^{F_t(f)-\delta}} \right] \leq \sqrt{\mathbb{E}[(\tilde{N}_t^{K,\varepsilon,f,1})^2] \mathbb{P}(\tilde{N}_t^{K,\varepsilon,f,1} > K^{F_t(f)-\delta})}. \quad (5.18)$$

Moreover,

$$\begin{aligned} \mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \mathbb{1}_{\tilde{N}_t^{K,\varepsilon,f,1} > K^{F_t(f)-\delta}} \right] &= \mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \right] - \mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \mathbb{1}_{\tilde{N}_t^{K,\varepsilon,f,1} \leq K^{F_t(f)-\delta}} \right] \\ &\geq \mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \right] - K^{F_t(f)-\delta} \\ &\geq \mathbb{E} \left[ \tilde{N}_t^{K,\varepsilon,f,1} \right] (1 - K^{\delta'+\delta''-\delta}), \end{aligned} \quad (5.19)$$

using (5.15). As a consequence, using (5.14),

$$\begin{aligned} \mathbb{P}(\tilde{N}_t^{K,\varepsilon,f,1} > K^{F_t(f)-\delta}) &\geq \frac{\mathbb{E}^2[\tilde{N}_t^{K,\varepsilon,f,1}] (1 - K^{\delta'+\delta''-\delta})^2}{\mathbb{E}[(\tilde{N}_t^{K,\varepsilon,f,1})^2]} \\ &\geq K^{-8\delta'} (1 - K^{\delta'+\delta''-\delta})^2. \end{aligned}$$

If  $\delta'$  and  $\delta''$  are small enough so that  $\delta' + \delta'' - \delta < 0$ , we obtain for  $K$  large enough,

$$\mathbb{P}(\tilde{N}_t^{K,\varepsilon,f,1} > K^{F_t(f)-\delta}) \geq \frac{1}{2} K^{-8\delta'} \geq K^{-9\delta'} \quad (5.20)$$

and the same inequality also holds for any  $\tilde{N}_t^{K,\varepsilon,f,i}$ ,  $i \in \{1, \dots, \lfloor K^{\delta''} \rfloor\}$ .

We can now conclude, using the branching property. Let us consider the full branching process started from  $Z_0^K(dx)$ . We have:

$$\begin{aligned} \mathbb{P}(N_t^{K,\varepsilon,f} \leq K^{F_t(f)-\delta}) &\leq \prod_{i=1}^{K^{\delta''}} \mathbb{P}(\tilde{N}_t^{K,\varepsilon,f,i} \leq K^{F_t(f)-\delta}) \\ &\leq (1 - K^{-9\delta'})^{\lfloor K^{\delta''} \rfloor} \sim \exp(-K^{\delta''-9\delta'}), \end{aligned} \quad (5.21)$$

for  $K$  sufficiently large. This bound converges to zero provided  $\delta'' - 9\delta' > 0$ .

Choosing  $\delta' = \delta/20$  and  $\delta'' = \delta/2$ , we have that  $\delta' + \delta'' - \delta = -9\delta/20 < 0$  and  $\delta'' - 9\delta' = \delta/20 > 0$ , in accordance to what we wanted in (5.20) and (5.21).

Moreover, from (5.21), we obtain by Borel-Cantelli's lemma, that (5.17) holds true almost surely. Since  $\delta$  was arbitrary, we have proved Lemma 5.1.  $\square$

## 6 The link between the variational formulation of the limit and the Hamilton-Jacobi equation (1.21)

In this section, we study the link between the variational formulation (1.18) and the Hamilton-Jacobi equation (1.21). In particular, we prove Theorem 1.4, Lemma 1.5, and Theorem 1.3. To this end, we first provide some preliminary lemmas.

**Lemma 6.1.** *Let  $(t, x) \in \tilde{\Omega}_a$ . There exists an optimal trajectory  $f \in \text{AC}([0, t])$  in the maximizing problem in (1.18), such that  $f(t) = x$  and that for all  $s \in [0, t]$ ,  $F_s(f) \geq a$ . Moreover,  $\|f\|_{W^{1,\infty}([0,t])}$  is uniformly bounded for all  $(t, x) \in [0, T] \times B_M(0)$ .*

*Proof.* Let  $f_n \in \text{AC}[0, t]$  be such that  $F_t(f_n) \rightarrow u_a(t, x)$  and  $F_s(f_n) \geq a$ , for all  $s \in [0, t]$ . Since  $L(x, v)$  is strictly convex and superlinear with respect to  $v$ , and since  $R(x)$  and  $\beta_0(x)$  are bounded above, it can be shown (see [28, Section 3.2]) that  $f_n$  converges, as  $n \rightarrow +\infty$ , to an absolutely continuous trajectory  $f_0$ , with  $f_0(t) = x$  and

$$F_s(f_n) \rightarrow F_s(f_0), \quad \text{for all } s \in [0, t].$$

Since  $F_s(f_n) \geq a$ , for all  $s \in [0, t]$ , we also have  $F_s(f_0) \geq a$  and hence  $(s, f_0(s)) \in \tilde{\Omega}_a$ . Moreover, one can prove (see [28] and [11, Lemma 6]) that this optimal trajectory is bounded in  $W^{1,\infty}([0, t])$  for all  $(t, x)$  in a compact set  $[0, T] \times B_M(0)$ . Note that here we can use the results of [11] since our assumptions in (1.2)–(1.3) lead to the assumptions made in the latter article (see [11, Corollary 4]).  $\square$

**Lemma 6.2.** *The set  $\Omega_a$  is an open set. Furthermore,  $u_a$  is bounded above locally in  $t$  and globally in  $x$ , and it is locally Lipschitz continuous in  $\bar{\Omega}_a$ , and consequently in  $\tilde{\Omega}_a$ . Moreover, the Lipschitz bound in  $t$  and  $x$ , is locally uniform with respect to  $a$ .*

*Proof.* i)  **$u_a$  is lower semicontinuous in  $\Omega_a$  and  $\Omega_a$  is an open set.** Let  $(t_1, x_1) \in \Omega_a$ . We prove that for all  $\eta > 0$ , there exists  $r > 0$ , small enough, such that  $B_r(t_1, x_1) \in \Omega_a$  and that, for all  $(t_2, x_2) \in B_r(t_1, x_1)$ , we have

$$u_a(t_2, x_2) \geq u_a(t_1, x_1) - \eta.$$

Let  $(t_2, x_2) \in B_r(t_1, x_1)$ . We define  $\delta = |t_1 - t_2| + |x_1 - x_2|$  and note that  $t_2 - \delta < t_1$ . We also recall that, due to Lemma 6.1, there exists an optimal trajectory  $f_1(\cdot)$  such that  $f_1(t_1) = x_1$ ,  $u(t_1, x_1) = F_{t_1}(f_1)$  and  $F_s(f) \geq a$ , for all  $s \in [0, t_1]$ . We define

$$\bar{f}(s) = \begin{cases} f_1(s) & \text{for all } s \in [0, t_2 - \delta], \\ f_1(t_2 - \delta) + \frac{s-t_2+\delta}{\delta}(x_2 - f_1(t_2 - \delta)) & \text{for all } s \in [t_2 - \delta, t_2]. \end{cases} \quad (6.1)$$

We will prove that, for  $r$  small enough,  $F_\tau(\bar{f}) \geq a$  for all  $\tau \in [0, t_2]$  and that  $F_{t_2}(\bar{f}) \geq u_a(t_1, x_1) - \eta$ . Note that, for all  $\tau \in [0, t_2 - \delta]$ ,  $F_\tau(\bar{f}) = F_\tau(f_1) \geq a$ . We next consider the case  $\tau \in [t_2 - \delta, t_2]$  and write

$$F_\tau(\bar{f}) = F_{t_1}(f_1) - \int_{t_2-\delta}^{t_1} [R(f_1(s)) - L(f_1(s), \dot{f}_1(s))] ds + \int_{t_2-\delta}^{\tau} [R(\bar{f}(s)) - L(\bar{f}(s), \dot{\bar{f}}(s))] ds. \quad (6.2)$$

Note that for all  $\tau \in [t_2 - \delta, t_2]$ ,  $\bar{f}(\tau) \in [x_2 \wedge f_1(t_2 - \delta), x_2 \vee f_1(t_2 - \delta)]$ . Hence,  $\bar{f}(\tau)$  is uniformly bounded in  $[t_2 - \delta, t_2]$ , for fixed  $(t_1, x_1)$  and  $r$ . We next note that

$$|\dot{\bar{f}}(\tau)| = \frac{|x_2 - f_1(t_2 - \delta)|}{\delta} \leq \frac{|x_1 - x_2|}{\delta} + \frac{2|f_1(t_1) - f_1(t_2 - \delta)|}{t_1 - t_2 + \delta}, \quad \text{for all } \tau \in [t_2 - \delta, t_2].$$

Since  $|\dot{f}_1|$  is uniformly bounded in  $[0, t_1]$  thanks to Lemma 6.1 and since  $|x_1 - x_2| \leq \delta$ , we deduce that there exists a constant  $C$  such that

$$|\dot{\bar{f}}(\tau)| \leq C.$$

Therefore, the integrand terms in the r.h.s. of (6.2) are bounded.

Since  $R$  and  $L$  are locally bounded, up to choosing  $r$  to be a smaller constant, we obtain that, for all  $\tau \in [t_2 - \delta, t_2]$  and,

$$F_\tau(\bar{f}) \geq u_a(t_1, x_1) - C'\delta \geq u_a(t_1, x_1) - C'r > a.$$

We deduce that  $(t_1, x_1) \in \Omega_a$  and hence  $\Omega_a$  is an open set. Furthermore, we have, for  $r$  small enough

$$u_a(t_2, x_2) \geq F_{t_2}(\bar{f}) \geq u_a(t_1, x_1) - C'r \geq u_a(t_1, x_1) - \eta.$$

(ii)  $u_a$  is continuous on  $\partial\Omega_a$ . We recall from (1.19) that  $u \geq a$  in  $\tilde{\Omega}_a$ . From the definition of  $\Omega_a$ , it is then immediate that  $u_a = a$  on  $\partial\Omega_a \subset \tilde{\Omega}_a \setminus \Omega_a$ . We prove that  $u_a$  is continuous on  $\partial\Omega_a$ . Let  $(\bar{t}, \bar{x}) \in \partial\Omega_a$  and  $(t_n, x_n) \in \Omega_a$  such that, as  $n \rightarrow +\infty$ ,  $(t_n, x_n) \rightarrow (\bar{t}, \bar{x})$ . Then, there exist optimal trajectories  $f_n : [0, t_n] \rightarrow \mathbb{R}$ , such that  $f_n(t_n) = x_n$ ,  $F_s(f_n) \geq a$ , for all  $s \in [0, t_n]$ , and  $u_a(t_n, x_n) = F_{t_n}(f_n)$ . Similarly to the proof of Lemma 6.1, we deduce that  $f_n$  converges along subsequences, as  $n \rightarrow +\infty$ , to an absolutely continuous trajectory  $\bar{f}$  such that  $\bar{f}(\bar{t}) = \bar{x}$ , and  $F_s(f_n) \rightarrow F_s(\bar{f})$ , for all  $s \in [0, \bar{t}]$ . Consequently,  $F_s(\bar{f}) \geq a$  for all  $s \in [0, \bar{t}]$  and  $u(\bar{t}, \bar{x}) \geq F_{\bar{t}}(\bar{f}) = \lim_{n \rightarrow +\infty} u_a(t_n, x_n) \geq a$ . Since  $(\bar{t}, \bar{x}) \notin \Omega_a$ , we deduce that  $u_a(\bar{t}, \bar{x}) = a$ , and hence  $\lim_{n \rightarrow +\infty} u_a(t_n, x_n) = u_a(\bar{t}, \bar{x}) = a$ .

(iii)  $u_a$  is locally Lipschitz continuous in  $\bar{\Omega}_a$ . We first prove that  $u_a$  is locally Lipschitz continuous in  $\Omega_a$ , for all  $a > 0$ , in the following sense. Let  $B_r \subset \Omega_a$  be a ball of radius  $r$ . We will prove that for  $r$  chosen small enough, there exists a constant  $C$  such that for all  $(t_1, x_1) \in \bar{B}_r$  and  $(t_2, x_2) \in \bar{B}_r$ ,

$$|u_a(t_1, x_1) - u_a(t_2, x_2)| \leq C(|t_1 - t_2| + |x_1 - x_2|).$$

The proof follows similar arguments as in the proof of (i). We first choose  $\delta = |t_1 - t_2| + |x_1 - x_2|$  and notice that  $t_2 - \delta < t_1$ . From step (i),  $u_a$  is lower semi-continuous and hence  $u_m := \min_{(t,x) \in \bar{B}_r} u_a(t, x) > a$ . Let  $f_1$  be the optimal trajectory such that  $f_1(t_1) = x_1$  and  $F_s(f_1) \geq a$  for all  $s \in [0, t_1]$  and  $F_{t_1}(f_1) = u_a(t_1, x_1) > a$ . We define  $\bar{f}$  as in (6.1) and notice similarly to above that  $\bar{f}$  and  $|\dot{\bar{f}}|$  are bounded. We also notice that, for  $\tau \in [0, t_2 - \delta]$ ,  $F_\tau(\bar{f}) \geq a$ . We next use (6.2) and the boundedness of  $\bar{f}$ ,  $\dot{\bar{f}}$ ,  $f_1$  and  $\dot{f}_1$  to obtain that, for  $\tau \in [t_2 - \delta, t_2]$ ,

$$F_\tau(\bar{f}) \geq F_{t_1}(f_1) - C\delta = u_a(t_1, x_1) - C\delta \geq u_m - C\delta.$$

We deduce on the one hand that, for  $r$  small enough,  $F_\tau(\bar{f}) \geq a$ , for all  $\tau \in [t_2 - \delta, t_2]$ . On the other hand, we have

$$u_a(t_2, x_2) \geq F_{t_2}(\bar{f}) \geq u_a(t_1, x_1) - C\delta.$$

The opposite inequality can be proved following similar arguments. We conclude that  $u_a$  is Lipschitz continuous in  $\bar{B}_r$ .

We next notice that the Lipschitz bound above only depends on the local bounds on  $L$  and  $R$ . From the continuity of  $u_a$  up to the boundary of  $\Omega_a$ , we deduce that  $u_a$  is indeed locally Lipschitz continuous in  $\bar{\Omega}_a$ . Since  $u_a(t, x) = a$  in  $\tilde{\Omega}_a \setminus \Omega_a$ , we deduce that  $u_a$  is also locally Lipschitz in  $\tilde{\Omega}_a$ . Finally, since the Lipschitz bound above only depends on the local bounds on  $L$  and  $R$ , we deduce that there exist Lipschitz bounds which are locally uniform with respect to  $a$ .

(v) **The bound from above.** From the definition of  $u_a$  in (1.18) and the fact that  $R$ ,  $-L$  and  $\beta_0$  are bounded from above, thanks to assumptions (1.3) and (1.5), we obtain a uniform bound from above on  $u_a$ , locally in  $t$  and globally in  $x$ .  $\square$

**Lemma 6.3.** *Let  $(t, x) \in \Omega_a$  and  $0 < \tau < t$ , with  $\tau$  small enough. Then, we have*

$$u_a(t, x) = \sup_{\substack{f \in \text{AC}[t-\tau, t] \\ f(t)=x}} \int_{t-\tau}^t [R(f(s)) - L(f(s), \dot{f}(s))] ds + u_a(t - \tau, f(t - \tau)). \quad (6.3)$$

*Proof.* Let  $f_0$  be an optimal trajectory such that  $f_0(t) = x$  and  $F_t(f_0) = x$ , and  $F_s(f_0) \geq a$ , for all  $s \in [0, t]$ . We have

$$u_a(t, x) = \int_{t-\tau}^t [R(f_0(s)) - L(f_0(s), \dot{f}_0(s))] ds + \int_0^{t-\tau} [R(f_0(s)) - L(f_0(s), \dot{f}_0(s))] ds + \beta_0(f_0(0)).$$

Since  $F_s(f_0) \geq a$ , for all  $s \in [0, t]$  and in particular for all  $s \in [0, t - \tau]$ , we deduce that

$$\begin{aligned} u_a(t, x) &\leq \int_{t-\tau}^t [R(f_0(s)) - L(f_0(s), \dot{f}_0(s))] ds + u_a(t - \tau, f_0(t - \tau)) \\ &\leq \sup_{\substack{f \in \text{AC}[t-\tau, t] \\ f(t)=x}} \int_{t-\tau}^t [R(f(s)) - L(f(s), \dot{f}(s))] ds + u_a(t - \tau, f(t - \tau)). \end{aligned}$$

Let us now assume that  $f_1 : [t - \tau, t] \rightarrow \mathbb{R}$  is such that  $f_1(t) = x$  and

$$u_a(t, x) < \int_{t-\tau}^t [R(f_1(s)) - L(f_1(s), \dot{f}_1(s))] ds + u_a(t - \tau, f_1(t - \tau)). \quad (6.4)$$

Let also  $f_2$  be an optimal trajectory such that  $f_2(t - \tau) = f_1(t - \tau)$ ,  $u_a(t - \tau, f_1(t - \tau)) = F_{t-\tau}(f_2)$  and  $F_s(f_2) \geq a$  for all  $s \in [0, t - \tau]$ . We then define

$$f_3(s) = \begin{cases} f_2(s) & s \in [0, t - \tau], \\ f_1(s) & s \in [t - \tau, t]. \end{cases}$$

Then it is immediate that  $u_a(t, x) < F_t(f_3)$ . We will prove that  $F_s(f_3) \geq a$ , for all  $s \in [0, t]$ , which leads to a contradiction with the latter inequality and the definition of  $u_a$ . Notice that since  $F_s(f_2) \geq a$ , for all  $s \in [0, t - \tau]$ , we deduce that  $F_s(f_3) \geq a$ , for all  $s \in [0, t - \tau]$ . To prove this property for  $s \in [t - \tau, t]$ , we write

$$F_s(f_3) > u_a(t, x) - \int_s^t [R(f_1(s)) - L(f_1(s), \dot{f}_1(s))] ds.$$

Since  $R$  is bounded from above and  $L$  is bounded from below, we deduce that, for all  $s \in [t - \tau, t]$  and  $C$  a positive constant,

$$F_s(f_3) > u_a(t, x) - C(t - s) \geq u_a(t, x) - C(t - \tau).$$

Since  $u_a(t, x) > a$ , choosing  $\tau$  small enough, we obtain that  $F_s(f_3) \geq a$  for all  $s \in [t - \tau, t]$  and hence  $F_s(f_3) \geq a$  for all  $s \in [0, t]$ .  $\square$

**Proof of Theorem 1.4.** Thanks to Lemma 6.2,  $u_a$  is locally Lipschitz continuous in  $\overline{\Omega}_a$  and due to Lemma 6.3 it satisfies (6.3). It is then immediate (see [3, Section 3.3]) that  $u$  is a viscosity solution to the Hamilton-Jacobi equation (1.21) in  $\Omega_a$ . We conclude the proof using the uniqueness of locally Lipschitz and bounded from above viscosity solutions to (1.21) [18]. Note that the uniqueness result in [18] is given for a Hamilton-Jacobi equation, with a convex

Hamiltonian, in the whole domain. However, the proof can be adapted to Hamilton-Jacobi equations, with a convex Hamiltonian, and with Dirichlet boundary conditions. Note also that when  $\Omega_a$  is bounded, the uniqueness follows from more standard arguments as in [3, Section 5].  $\square$

We next prove Lemma 1.5.

**Proof of Lemma 1.5.** (i) We first prove (1.23). Let  $\mathcal{K} \subset \mathbb{R}^+ \times \mathbb{R}$  be a compact set. We define

$$f^{\mathcal{K}}(a) = \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_a}(t, x) dt dx.$$

Since  $(\tilde{\Omega}_a)_a$  is a decreasing family of sets, we deduce that  $f_a^{\mathcal{K}}$  is a decreasing function with respect to  $a$ . Since a decreasing function has at most a countable set of discontinuity points, we deduce that  $f^{\mathcal{K}}(a)$  is continuous at almost every point  $a$ . At a continuity point  $a_0$  of  $f^{\mathcal{K}}$  we have

$$\lim_{a \rightarrow a_0} \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_a}(t, x) dt dx = \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_{a_0}}(t, x) dt dx.$$

Moreover, since  $\Gamma_{a_0} = \bigcup_{a > a_0} \tilde{\Omega}_a$ , we deduce that

$$\int_{\mathcal{K}} \mathbb{1}_{\Gamma_{a_0}}(t, x) dt dx = \lim_{a \downarrow a_0} \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_a}(t, x) dt dx = \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_{a_0}}(t, x) dt dx.$$

It follows that, for a.e.  $a_0 \in \mathbb{R}$ ,

$$\int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_{a_0} \setminus \Gamma_{a_0}}(t, x) dt dx = 0.$$

Since this equality holds a.e. for all compact set  $\mathcal{K} \subset \mathbb{R}^+ \times \mathbb{R}$ , we deduce that, for a.e.  $a_0 \in \mathbb{R}$ ,

$$\int_{\mathbb{R}^+ \times \mathbb{R}} \mathbb{1}_{\tilde{\Omega}_{a_0} \setminus \Gamma_{a_0}}(t, x) dt dx = 0.$$

(ii) We next prove (1.24). Let  $\mathcal{K} \subset \mathbb{R}^+ \times \mathbb{R}$ , be a compact set. We then define

$$g(a) = \int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_a} u_a(t, x) dt dx.$$

Notice that  $\mathbb{1}_{\tilde{\Omega}_a} u_a(t, x)$  is decreasing with respect to  $a$ . It hence converges, as  $a \downarrow a_0$ , to  $\mathbb{1}_{\Gamma_{a_0}} v(t, x)$ , for a certain function  $v(t, x)$ . Moreover, since  $u_a$  is locally Lipschitz in  $\tilde{\Omega}_a$  with respect to  $t$  and  $x$ , with a locally uniform dependence on  $a$ , we deduce that  $v(t, x)$  is indeed a continuous function in  $\Gamma_{a_0}$ . We also have

$$\lim_{a \downarrow a_0} g(a) = \int_{\mathcal{K}} \mathbb{1}_{\Gamma_{a_0}} v(t, x) dt dx.$$

We also notice that  $g$  is a decreasing function with respect to  $a$ . Consequently,  $g$  is continuous with respect to  $a$ , for almost every  $a$ . We deduce that, for almost every  $a_0 \in \mathbb{R}$ ,

$$\int_{\mathcal{K}} \mathbb{1}_{\tilde{\Omega}_{a_0}} u_{a_0}(t, x) dt dx = g(a_0) = \lim_{a \downarrow a_0} g(a) = \int_{\mathcal{K}} \mathbb{1}_{\Gamma_{a_0}} v(t, x) dt dx.$$

From the monotonicity of  $\Omega_a$  and  $u_a$  we also obtain

$$\mathbb{1}_{\tilde{\Omega}_{a_0}} u_{a_0}(\cdot, \cdot) \geq \mathbb{1}_{\Gamma_{a_0}} v(\cdot, \cdot).$$

Combining the lines above we obtain that, for almost every  $a_0 \in \mathbb{R}$  and  $(t, x) \in \mathcal{K}$ ,

$$\mathbb{1}_{\tilde{\Omega}_{a_0}} u_{a_0}(t, x) = \mathbb{1}_{\Gamma_{a_0}} v(t, x),$$

and consequently, for almost every  $a_0 \in \mathbb{R}$  and  $(t, x) \in \mathcal{K} \cap \Gamma_{a_0}$ ,

$$u_{a_0}(t, x) = v(t, x).$$

Finally, since  $u_{a_0}$  and  $v$  are both continuous in  $\Gamma_{a_0}$ , we deduce that for almost every  $a_0 \in \mathbb{R}$  and for all  $(t, x) \in \mathcal{K} \cap \Gamma_{a_0}$ ,

$$u_{a_0}(t, x) = v(t, x) = \lim_{a \downarrow a_0} u_a(t, x).$$

Since this equality holds in  $\mathcal{K} \cap \Gamma_{a_0}$  for a.e.  $a_0$  and any compact set  $\mathcal{K}$ , we deduce that it also holds for a.e.  $a_0$  and for all  $(t, x) \in \Gamma_{a_0}$ . Combining this property with (1.22) we obtain (1.24).  $\square$

**Proof of Theorem 1.3.** (i) We first prove the lower bound. From Theorem 1.2 we deduce that

$$\liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, G_t^{x, \delta}} \geq \sup\{F_t(f); f \in G_t^{x, \delta}, \forall s \in [0, t], F_s(f) > 0\}.$$

Let  $a > 0$ . Then, we have for all  $a > 0$  and  $\delta > 0$ ,

$$\{f \in AC[0, t], f(t) = x, \forall s \in [0, t], F_s(f) \geq a\} \subset \{f \in G_t^{x, \delta}, \forall s \in [0, t], F_s(f) > 0\}.$$

Using (1.19) we obtain that, for all  $a > 0$  and  $\delta > 0$ ,

$$u_a(t, x) \leq \sup\{F_t(f); f \in G_t^{x, \delta}, \forall s \in [0, t], F_s(f) > 0\}.$$

Combining the properties above we conclude that

$$\lim_{a \downarrow 0} u_a(t, x) \leq \liminf_{\delta \rightarrow 0} \liminf_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, G_t^{x, \delta}}.$$

(ii) We next prove the upper bound. From Theorem 1.1 we obtain that

$$\limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_t^{x, \delta}} \leq \sup\{F_t(f); f \in A_t^{x, \delta}, \forall s \in [0, t], F_s(f) \geq 0\}.$$

Similarly to Lemma 6.1, and since the set  $A_t^{x, \delta}$  is a closed set, there exists an optimal trajectory  $f^\delta$  which maximizes  $F_t(\cdot)$  in the set above, with  $F_s(f^\delta) \geq 0$  for all  $s \in [0, t]$ . Moreover,  $\|f^\delta\|_{W^{1, \infty}([0, t])}$  is bounded uniformly with respect to  $\delta$ . Note that the  $W^{1, \infty}$  bound is proved in [28] and [11, Lemma 6] in the maximization problem with a fixed ending point. Here, we consider the trajectories with ending points in  $[x - \delta, x + \delta]$ . Therefore, if  $f_0$  is the optimal trajectory,  $f_0$  is also an optimal trajectory with ending point at  $f_0(t)$  and the same result applies. We deduce that

$$\limsup_{\delta \rightarrow 0} \limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_t^{x, \delta}} \leq \limsup_{\delta \rightarrow 0} F_t(f^\delta).$$

Let  $(f^{\delta_n})_n$  be a sequence of trajectories such that

$$\limsup_{\delta \rightarrow 0} F_t(f^\delta) = \lim_{n \rightarrow \infty} F_t(f^{\delta_n}).$$

From the uniform bound in  $W^{1,\infty}$ , we deduce that, up to considering a subsequence,  $f^{\delta_n}$  converges to a trajectory  $f^0 \in AC[0, t]$  such that

$$\limsup_{\delta \rightarrow 0} F_t(f^\delta) = F_t(f^0), \quad f^0(t) = x, \quad F_s(f^0) \geq 0, \quad \forall s \in [0, t].$$

It follows that

$$\limsup_{\delta \rightarrow 0} \limsup_{K \rightarrow +\infty} \frac{1}{\log K} \log N_t^{K, A_t^{x, \delta}} \leq u_0(t, x).$$

□

## A Proof of the many-to-one formulas

### A.1 Proof of Proposition 2.1 (i)

Let us give a simple proof based on Itô's formula. Let us first note that the intensity measure of  $Z_t^K$ ,  $\nu_t^K(dy) = \mathbb{E}_{\delta_x} [Z_t^K(dy)]$  defined for any  $\varphi$  in  $C_b(\mathbb{R})$  by

$$\langle \nu_t^K, \varphi \rangle = \mathbb{E}_{\delta_x} [\langle Z_t^K, \varphi \rangle]$$

is the unique weak solution of

$$\begin{cases} \partial_t \nu_t = \nu_t \mathcal{L}^K + R \nu_t, \\ \nu_0 = \delta_x, \end{cases} \quad (\text{A.1})$$

where we denote by  $\nu \mathcal{L}^K$  the adjoint of the operator  $\mathcal{L}^K$  applied to the probability measure  $\nu$ . Uniqueness of such a solution is proven as in Theorem 2.2 in [33].

Let us show that the r.h.s. term of (2.4) also satisfies (A.1). Uniqueness will yield the result. Let  $\varphi$  in  $C_b^1(\mathbb{R})$ . Applying Itô's formula with jumps (e.g. [34, Th.5.1]) to the semimartingale  $\exp\left(\int_0^t R(X_s^K) ds\right) \varphi(X_t^K)$ , we have

$$\begin{aligned} \exp\left(\int_0^t R(X_s^K) ds\right) \varphi(X_t^K) - \varphi(X_0^K) - \int_0^t \exp\left(\int_0^s R(X_r^K) dr\right) \mathcal{L}^K \varphi(X_s^K) ds \\ - \int_0^t \varphi(X_s^K) R(X_s^K) \exp\left(\int_0^s R(X_r^K) dr\right) ds. \end{aligned}$$

is a square integrable martingale since  $R$  is bounded. Taking the expectation, we obtain that

$$\begin{aligned} \mathbb{E}_x \left[ \exp\left(\int_0^t R(X_s^K) ds\right) \varphi(X_t^K) \right] = \varphi(x) \\ + \mathbb{E}_x \left[ \int_0^t \exp\left(\int_0^s R(X_r^K) dr\right) \left\{ R(X_s^K) \varphi(X_s^K) + \mathcal{L}^K \varphi(X_s^K) \right\} ds \right]. \quad (\text{A.2}) \end{aligned}$$

If we define the measure  $\mu_t$  for any test function  $\varphi \in C_b^1(\mathbb{R})$  by

$$\langle \mu_t, \varphi \rangle = \mathbb{E}_x \left[ \exp\left(\int_0^t R(X_s^K) ds\right) \varphi(X_t^K) \right],$$

we obtain from (A.2) that

$$\langle \mu_t, \varphi \rangle = \langle \delta_x, \varphi \rangle + \int_0^t \langle \mu_s, R\varphi + \mathcal{L}^K \varphi \rangle ds.$$

This proves that the flow  $(\mu_t, t \geq 0)$  is a weak solution of (A.1) and the conclusion follows by uniqueness of solution of the equation.

## A.2 Proof of Proposition 2.2

Recall that  $V_{[0,t]}^K$  is the set of individuals born before time  $t$  and that  $V_t^K$  is the set of individuals still alive at time  $t$ .

Note that in our model, the trait  $x$  of an individual  $v$  remains constant during their life. The total rate of event for such an individual will be denoted here by

$$\Lambda(x) = b(x) + p(x) + d(x),$$

and the time  $S_v$  where they dies or gives birth has a density  $\Lambda(x) \exp(\Lambda(x)(s - S_v^0)) \mathbb{1}_{s > S_v^0}$  with respect to Lebesgue's measure, conditionally on its birth time  $S_v^0$ . Let us note that

$$v \in V_s^K \iff S_v^0 \leq s < S_v.$$

First, we have for all  $v \in \mathcal{U}$ ,

$$\begin{aligned} & \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \Phi((X_{r \wedge S_v}^{K,v}, r \leq t), S_v \wedge t) \right] \\ &= \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \int_{S_v^0}^{+\infty} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s \wedge t) \Lambda(X_s^{K,v}) e^{-\int_{S_v^0}^s \Lambda(X_r^{K,v}) dr} ds \right] \\ &= \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \int_{S_v^0}^{+\infty} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s \wedge t) \Lambda(X_s^{K,v}) \left( \int_s^{+\infty} \Lambda(X_\tau^{K,v}) e^{-\int_{S_v^0}^\tau \Lambda(X_r^{K,v}) dr} d\tau \right) ds \right] \\ &= \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \int_{S_v^0}^{+\infty} \left( \int_{S_v^0}^\tau \Phi((X_{r \wedge s}^{K,v}, r \leq t), s \wedge t) \Lambda(X_s^{K,v}) ds \right) \Lambda(X_\tau^{K,v}) e^{-\int_{S_v^0}^\tau \Lambda(X_r^{K,v}) dr} d\tau \right] \\ &= \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \int_{S_v^0}^{S_v} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s \wedge t) \Lambda(X_s^{K,v}) ds \right], \end{aligned} \tag{A.3}$$

where we used Fubini's theorem at the third equality, and where we recognized the distribution

of  $S_v$  to obtain the last equality. Then, summing (A.3) over  $v \in \mathcal{U}$  entails

$$\begin{aligned}
& \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_{[0,t]}^K} \Phi((X_{r \wedge S_v}^{K,v}, r \leq t), S_v \wedge t) \right] \\
&= \sum_{v \in \mathcal{U}} \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_{[0,t]}^K} \int_{S_v^0}^{S_v} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s \wedge t) \Lambda(X_s^{K,v}) ds \right] \\
&= \sum_{v \in \mathcal{U}} \mathbb{E}_{\delta_x} \left[ \int_0^t \mathbb{1}_{v \in V_s^K} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s) \Lambda(X_s^{K,v}) ds \right] \\
&\quad + \sum_{v \in \mathcal{U}} \mathbb{E}_{\delta_x} \left[ \mathbb{1}_{v \in V_t^K} \Phi((X_r^{K,v}, r \leq t), t) \int_t^{S_v} \Lambda(X_s^{K,v}) ds \right] \\
&= \int_0^t \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_s^K} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s) \Lambda(X_s^{K,v}) \right] ds \\
&\quad + \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_t^K} \Phi((X_r^{K,v}, r \leq t), t) \Lambda(X_t^{K,v}) (S_v - t) \right] \\
&= \int_0^t \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_s^K} \Phi((X_{r \wedge s}^{K,v}, r \leq t), s) \Lambda(X_s^{K,v}) \right] ds + \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_t^K} \Phi((X_r^{K,v}, r \leq t), t) \right]
\end{aligned}$$

where we used in the third line that for  $v \in V_t^K$  and  $t \leq s < S_v$ ,  $X_s^{K,v} = X_t^{K,v}$ . For the last equality, we notice that

$$\mathbb{E}[S_v - t \mid \mathcal{F}_t] = \frac{1}{\Lambda(X_t^{K,v})}.$$

Using Proposition 2.1(ii), we obtain

$$\begin{aligned}
& \mathbb{E}_{\delta_x} \left[ \sum_{v \in V_{[0,t]}^K} \Phi((X_{r \wedge S_v}^{K,v}, r \leq t), S_v \wedge t) \right] \\
&= \int_0^t \mathbb{E}_x \left[ \Phi((X_{r \wedge s}^K, r \leq t), s) \Lambda(X_s^K) \exp \left( \int_0^s R(X_r^K) dr \right) \right] ds \\
&\quad + \mathbb{E}_x \left[ \Phi((X_r^K, r \leq t), t) \exp \left( \int_0^t R(X_r^K) dr \right) \right].
\end{aligned}$$

This ends the proof of Proposition 2.2.

## B On Skorohod balls around absolutely continuous functions

We prove here the inclusion stated in (5.6). Let  $f \in AC[0, t]$ . Note that for any homeomorphism  $\lambda$  of  $[0, t]$ , we have

$$\sup_{r \in [0,t]} |f(r) - g(r)| \leq \sup_{r \in [0,t]} |f \circ \lambda(r) - g(r)| + \sup_{r \in [0,t]} |f \circ \lambda(r) - f(r)|. \quad (\text{B.1})$$

Thus, if  $d_{\text{sko}}(f, g) < \varepsilon$ , the first term in the right hand side can be made smaller than  $2\varepsilon$  for a good choice of homeomorphism  $\lambda$ . By (1.1), the latter homeomorphism can be chosen such that

$$\sup_{r \in [0, t]} |\lambda(r) - r| \leq (e^\varepsilon - 1)t.$$

Then, the second term in the right hand side of (B.1) is upper-bounded by the modulus of continuity  $\omega(f, (e^\varepsilon - 1)t)$ , where

$$\omega(f, \eta) = \sup_{|t-s| < \eta} |f(t) - f(s)|.$$

Since  $f$  is absolutely continuous, and since  $(e^\varepsilon - 1)t$  converges to zero for  $\varepsilon \rightarrow 0$ , the second term also converges to zero. Gathering these equations gives (5.6), for  $\eta(\varepsilon) = 2\varepsilon + \omega(f, (e^\varepsilon - 1)t)$ .

## Acknowledgements

We thank B. Mallein and P. Maillard for enlightening discussions and references. This work is funded by the European Union (ERC, SINGER, 101054787 and ERC-2024-COG MUSEUM-101170884). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. This work has also been supported by the Chair “Modélisation Mathématique et Biodiversité” of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X. V.C.T. acknowledge the support of the R-CDP-24-004-C2EMPI project, funded by the French State under the France-2030 programme, the University of Lille, the Initiative of Excellence of the University of Lille, the European Metropolis of Lille.

## References

- [1] V. Bansaye, J.F. Delmas, L. Marsalle and V.C. Tran. Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *The Annals of Applied Probability*, Vol. 21, No. 6, 2263-2314, (2011).
- [2] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*. Springer-Verlag Berlin Heidelberg, 1994.
- [3] G. BARLES, *An Introduction to the Theory of Viscosity Solutions for First-Order Hamilton-Jacobi Equations and Applications*, Lecture Notes in Mathematics 2074, 2013.
- [4] G. Barles, L. C. Evans, and P. E. Souganidis. Wavefront propagation for reaction-diffusion systems of PDE. *Duke Math. J.*, 61(3):835–858, 1990.
- [5] G. Barles, S. Mirrahimi, and B. Perthame. Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result. *Methods and Applications of Analysis*, 16(3):321–340, 2009.
- [6] J. Berestycki and E. Brunet and J.W. Harris and J. W. Harris and C. Simon and M.I. Roberts. Growth rates of the population in a branching Brownian motion with an inhomogeneous breeding potential. *Stochastic Process. Appl.* 125 (2015), no. 5, 2096–2145.
- [7] J.D. Biggins. The growth and spread of the general branching random walk. *The Annals of Applied Probability*, 5(4):1008-1024, (1995).
- [8] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York (1968).
- [9] A. Bovier, L. Coquille, C. Smadi. Crossing a fitness valley as a metastable transition in a stochastic population model. *Ann. Appl. Probab.* 29(6), 3541–3589 (2019).

- [10] V. Calvez, B. Henry, S. Méléard, V.C. Tran. Dynamics of lineages in adaptation to a gradual environmental change. *Annales Henri Lebesgue*, 5:729–777, 2022.
- [11] V. CALVEZ, V. AND K.-Y. LAM, *Uniqueness of the viscosity solution of a constrained Hamilton-Jacobi equation*, *Calc. Var. Partial Differ. Equ.*, **59(5)** (2020) pp. 163.
- [12] N. Champagnat. A microscopic interpretation for adaptative dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116:1127–1160, 2006.
- [13] N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models via timescale separation. *Theoretical Population Biology*, 69:297–321, 2006.
- [14] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151(1-2):45–94, 2011.
- [15] N. Champagnat, S. Méléard and V.C. Tran. Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. *Ann. Appl. Probab.*, 31(4), 1820–1867, 2021.
- [16] N. Champagnat, S. Méléard, S. Mirrahimi and V.C. Tran. Filling the gap between individual-based evolutionary models and Hamilton-Jacobi equations. *J. Ec. Polytechnique*, 10, 1247–1275, 2023.
- [17] L. Coquille and A. Kraut and C. Smadi. Stochastic individual-based models with power law mutation rate on a general finite trait space. *Electron. J. Probab.*, 26, 1–37, 2021.
- [18] G. DAL MASO, H. FRANKOWSKA, *Value functions for Bolza problems with discontinuous Lagrangians and Hamilton-Jacobi inequalities*, *ESAIM: Control, Optimisation and Calculus of Variations* 5, 369-393, 2000.
- [19] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Vol 38., Springer, second edition, 1998.
- [20] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame. The dynamics of adaptation: an illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67, 257–271, 2005.
- [21] P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics, 1997.
- [22] R. Durrett and J. Mayberry. Travelling waves of selective sweeps. *Annals of Applied Probability*, 21(2), 699–744, 2011.
- [23] M. Esser and A. Kraut. A general multi-scale description of metastable adaptive motion across fitness valleys. *J. Math. Biol.*, 89(46), 2024.
- [24] Stewart N. Ethier and Thomas G, Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 1986.
- [25] L. C. Evans *Partial differential equations Graduate Studies in Mathematics Vol. 19*, American Mathematical Society, 1998.
- [26] L. C. Evans and P. E. Souganidis. A PDE approach to geometric optics for certain semilinear parabolic equations. *Indiana Univ. Math. J.*, 38(1):141–172, 1989.
- [27] M. Fang and O. Zeitouni. Slowdown for Time Inhomogeneous Branching Brownian Motion. *J. Statistical Physics*, 149:1-9, (2012).
- [28] A. Fathi. *Weak Kam Theorem in Lagrangian Dynamics*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2016.
- [29] M Freidlin. Limit theorems for large deviations and reaction-diffusion equations. *The Annals of Probability*, 13(3):639–675, 1985.
- [30] R. Forien, J. Garnier and F. Patout. Ancestral lineages in mutation-selection equilibria with moving optimum. *Bull Math Biol.*, 84(93), 2022.
- [31] N. Fournier and S. Méléard. A Microscopic Probabilistic Description of a Locally Regulated Population and Macroscopic Approximations. *Ann. Appl. Probab.*, 14(4):1880-1919, 2004.
- [32] R. Hardy and S.C. Harris. *A Spine Approach to Branching Diffusions with Applications to Lp-Convergence of Martingales*. *Séminaire de Probabilités XLII*, Lecture Notes in Mathematics, 1979:281-330, Springer Nature, 2009.
- [33] B. Henry, S. Méléard and V.C. Tran. Time reversal of spinal processes for linear and non-linear branching processes near stationarity. *Electronic Journal of Probability*, 28(32):1–27, 2023.

- [34] N. Ikeda and S. Watanabe. Stochastic Differential Equations and Diffusion Processes, North-Holland Publishing Company, 1989.
- [35] P.-E. Jabin. Small populations corrections for selection-mutation models. Netw. Heterog. Media, 7(4), 805–836, 2012.
- [36] A. Jeddi. Asymptotic behavior of some stochastic models in population dynamics: a Hamilton-Jacobi approach. ArXiv arXiv:2602.20825, 2026.
- [37] G. Last and M. Penrose. Lectures on the Poisson Process. IMS Textbook by Cambridge University Press (2017)
- [38] A. Lorz, S. Mirrahimi and B Perthame. Dirac mass dynamics in a multidimensional nonlocal parabolic equation. Communications in Partial Differential Equations, 36, 1071–1098 (2011).
- [39] P. Maillard and G. Raoul and J. Tourniaire. Spreading speed of locally regulated population models in macroscopically heterogeneous environments (2024). ArXiv:2105.06985.
- [40] B. Mallein. Maximal displacement of a branching random walk in time- inhomogeneous environment. Stochastic Processes and Their Applications, 125(10), 3958–4019 (2015).
- [41] A. Marguet. Uniform sampling in a structured branching population. Bernoulli, 25, 4A, 2649-2695, 2019.
- [42] S. Méléard and V.C. Tran. Nonlinear historical superprocess approximations for population models with past dependence. Electronic Journal of Probability, 17(47):1-32, 2012.
- [43] S. Mirrahimi, G. Barles, B. Perthame, P.E. Souganidis. A singular Hamilton-Jacobi equation modeling the tail problem. SIAM J. Math. Anal., 44 (6), 4297–4319, 2012.
- [44] B. Perthame and G. Barles. Dirac concentrations in Lotka-Volterra parabolic PDEs. Indiana Univ. Math. J. 57, 3275–3301 (2008).
- [45] B. Perthame and M. Gauduchon. Survival thresholds and mortality rates in adaptive dynamics: conciliating deterministic and stochastic simulations. Math. Med. Biol. 27(3), 195–210 (2010).
- [46] P.E. Protter. *Stochastic integration and differential equations*, second Edition. Stochastic Modelling and Applied Probability, 21, Springer, Berlin, 2004.
- [47] C. Desmarais, E. Schertzer, Z. Talygiás. K-Branching random walk with noisy selection: large population limits and phase transitions. arXiv:2509.26254.
- [48] D. Waxman and S. Gavrilets. 20 Questions on Adaptive Dynamics. Journal of Evolutionary Biology 18, 1139-1154 (2005).
- [49] O. Zeitouni. Branching random walks and Gaussian fields. Lecture notes, (2012).