Well-posedness of Hamilton–Jacobi equations in population dynamics and applications to large deviations

Kraaij, Richard C.; Mahé, Louis

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Well-posedness of Hamilton-Jacobi equations in population dynamics and applications to large deviations

Richard C. Kraaij[*]   Louis Mahé [†]

March 28, 2020

Abstract

We prove Freidlin-Wentzell type large deviation principles for various rescaled models in populations dynamics that have immigration and possibly harvesting: birth-death processes, Galton-Watson trees, epidemic SI models, and prey-predator models.

The proofs are carried out using a general analytic approach based on the well-posedness of a class of associated Hamilton-Jacobi equations. The notable feature for these Hamilton-Jacobi equations is that the Hamiltonian can be discontinuous at the boundary. We prove a well-posedness result for a large class of Hamilton-Jacobi equations corresponding to one-dimensional models, and give partial results for the multi-dimensional setting.

Keywords: Large deviations; population dynamics; Hamilton-Jacobi equations; boundary conditions

MSC2010 classification: 49L25; 60F10; 60J80; 92D25

1 Introduction

The evolution of the population size of one or more interacting species can be modeled using Markov jump processes. The jumps correspond to the result of interactions between and amongst individuals of the population: predation, reproduction or correspond to global phenomena such as immigration, emigration, harvesting, or mutation (see e.g. [1, 19]).

A much studied question is the dynamics of the population in the large population scaling limit. Typical results include laws of large numbers to a solution of a first-order differential equation or weak convergence of the process to a solution of a stochastic differential equation as in [3, 6, 19] and [13, Chapter 11].

Additional work is required for the estimation of the exit time from a domain: time spent before extinction, time for the propagation of a disease among a population, time for a mutation to establish itself in a population. This motivates the study of path-space large deviations of Freidlin-Wentzell type [15]. In short, given a sequence of stochastic processes $X_n(t)$, one aims to establish

$$\mathbb{P}[\{X_n(t)\}_{t \geqslant 0} \approx \{\gamma(t)\}_{t \geqslant 0}] \approx e^{-nI(\gamma)},$$

---

[*]Department of Applied Mathematics, Delft University of Technology, The Netherlands. *E-mail address*: R.C.Kraaij@tudelft.nl

[†]CPHT, Ecole Polytechnique, CNRS, Université Paris-Saclay, Route de Saclay, 91128 Palaiseau, France *E-mail address*: louis.mahe@polytechnique.edu

where

$$(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) \mathrm{d}s & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise.} \end{cases}$$

Here $\mathcal{AC}$ is an appropriate space of absolutely continuous trajectories and where the *Lagrangian* $\mathcal{L}$ is some non-negative function.

A notable feature of the large deviations for models in population dynamics, as well as in the analysis of queuing systems, is that the systems have boundaries (arising e.g. from the condition that a population size is non-negative) and that the jump rates vanish at the boundary. Large deviations for such systems have been considered by [4, 17, 20] in the context of population dynamics and in [23] in the context of queuing systems. Indeed, it is the boundary where the rates vanish, that leads to issues in adapting the methods of [15].

To obtain results, systems are assumed to start in the interior, and the large deviation principle is proven by approximating trajectories that hit the boundary by trajectories that remain in the interior. In addition, the rate function evaluated in the approximating trajectories need to approximate the rate function evaluated in the limiting trajectory. This approach has two major drawbacks:

(a) the pre-limit processes are not allowed to start on the boundary,

(b) the approach does not allow for discontinuous rates at the boundary.

These problems were also identified by [12] in the context of queuing systems in which the rates are discontinuous for natural reasons: if there are customers in the queue, a server is active whereas it is inactive if the queue is empty. This leads to the analysis of a system that has homogeneous rates in the interior of the space, but with discontinuity at the boundary. To treat the large deviations of such systems the authors instead turn to characterizing the large deviations via solutions to a Hamilton-Jacobi equation and prove well-posedness for the Hamilton-Jacobi equation. See also [11] for a model with discontinuities in the interior instead of at the boundary.

In this paper, we study the large deviations for population dynamics, including the possibility of starting at the boundary as well as discontinuous behaviour at the boundary. This leads to a set of models in which we have inhomogeneous rates in the interior that might vanish at the boundary or have discontinuities at the boundary.

Discontinuities appear naturally e.g. in the context of population dynamics in the presence of harvesting or predation. A (short-sighted) external entity takes out individuals from the population of the system at a fixed rate, e.g. predators hunting prey possibly until extinction, or the influence of the fishing industry on the population of fish in a lake or sea. Clearly, no population can be taken out of the system when there is no population left.

Our aim is to prove the large deviation principle for jump processes that can treat discontinuities as well as vanishing rates. Our main result complements the established results mentioned above by covering the middle ground. As an added benefit, our approach partially extends [17, 20] by allowing for non-Lipschitz rates. In particular, this means we can consider models that do not have a deterministic limit (See e.g. [16, Example 2.5.3].) We do mention that we cannot allow for an as general state-space as in the recent paper [17]. Neither have we pushed our results for multi-dimensional models into the domain of [12] that were treated with non-quadratic test functions. We further discuss the possibilities and restrictions of our approach in Section 3.4.

Even though this is not the focus of our paper, the results can be used to study the exit time from a domain. In the context of biological processes starting at the

boundary, this allows e.g. to study the time that a new species needs to establish itself in a hostile new environment.

We follow [12] in studying large deviations via the theory of viscosity solutions to Hamilton-Jacobi equations. However, we rather follow the more recent framework introduced by [14], who essentially reduce proving the large deviation principle to a well-posedness question to a class of Hamilton-Jacobi equations. This method has been used recently in various papers, see e.g. [**?**,5,10] and references therein, in which various steps in this method have been explained in detail.

The main technical step in this paper is, therefore, the establishment of well-posedness for the associated Hamilton-Jacobi equations. To be able to treat discontinuities at the boundary in our specific context, we also introduce a framework to account for which types of behaviour are exhibited by the process close to the boundary. Indeed, our Lagrangian needs to take into account the various types of possible behaviour as in [12]. Finally, we argue that there are settings in which the large deviation principle as well as the well-posedness of the Hamilton-Jacobi equation fails to hold.

The rest of the paper is organized as follows. We start in Section 2 with general preliminaries. In addition, we introduce a framework to describe the exact structure of our space and its boundaries. We then introduce collections of Hamiltonians that take into account the various types of behaviour exhibited by the processes close to the boundary. Next, in Section 3 we proceed by applying our main abstract results to a collection of examples to illustrate the range of applications. We treat various one-dimesional models including birth-death processes, processes with births only like the process that models the size of a Galton-Watson tree and of some simple epidemic models. We then discuss the failure of well-posedness for a class of birth-death processes. We conclude the section with a treatment of a class of multi-dimensional processes in which there is immigration and harvesting for each species.

We proceed with the comparison principle for Hamilton-Jacobi equations on spaces with boundaries in Section 4 and to apply these results to the theory of large deviations in Section 5. We conclude our paper with Section 6 in which we apply the results of Sections 4 and 5 to the examples of Section 3.

2 Preliminaries

Before starting with our main results, we state some preliminaries needed for the formulation of our main results. In Section 2.1, we give some general mathematical preliminaries including the main definitions of large deviations theory. We proceed in Section 2.2 by introducing the class of spaces on which our results will be applicable, in addition we introduce a suitable framework to effectively study the behavior of our Hamiltonians on the boundaries of our spaces. Finally, in Section 2.3 we introduce a class of Hamiltonians and Lagrangians which are suited to for our spaces with boundary.

2.1 General preliminaries

Let E be a d-dimensional convex polyhedron with non-empty interior. A full definition follows on page 5 below. We denote by $\mathcal{P}(E)$ the space of Borel probability measures on E and we denote by $D_E(\mathbb{R}^+)$ the space of paths $\gamma : \mathbb{R}^+ \to E$ that are right continuous and have left limits. We endow $D_E(\mathbb{R}^+)$ with the Skorokhod topology, cf. [13, Section 3.5]. An important property is that under this topology $D_E(\mathbb{R}^+)$ is Polish if E is Polish.

We denote by $C_b(E), C_l(E)$ and $C_u(E)$ the spaces of continuous functions that are bounded, bounded from below and bounded from above respectively. We write $C_c(E)$ for the functions with compact support in E relative to the topology of E. Finally, we denote $C_c^k(E)$ for functions in $C_c(E)$ that are k times continuously differentiable.

Let $\overline{\mathbb{R}} = [-\infty, \infty]$ be the two-point compactification of $\mathbb{R}$. We denote by $\mathcal{M}(E, \overline{\mathbb{R}})$ the space of measurable functions from E into $\overline{\mathbb{R}}$. The spaces $\mathcal{M}_b(E, \overline{\mathbb{R}})$, $\mathcal{M}_l(E, \overline{\mathbb{R}})$ and $\mathcal{M}_u(E, \overline{\mathbb{R}})$ are the subspaces of $\mathcal{M}(E, \overline{\mathbb{R}})$ with functions that are bounded, bounded from above and below from respectively.

We say that a set $F \subseteq C_b(E)$ *separates points* in E if for all $x, y \in E$, $x \neq y$ there is $f \in F$ with $f(x) \neq f(y)$.

For a function $f : E \to \mathbb{R}$, we denote by f^* and f_* the upper semi-continuous and lower semi-continuous regularization respectively.

For a set $G \subseteq E$, we write $\mathrm{ch}G$ for the closed convex hull of G in E.

Definition 2.1. Let $\{X_n\}_{n \geqslant 1}$ be a sequence of random variables on a Polish space $\mathcal{X}$. Furthermore, consider a function $I : \mathcal{X} \to [0, \infty]$ and a sequence $\{r_n\}_{n \geqslant 1}$ of positive numbers such that $r_n \to \infty$. We say that

- I is a *good rate-function* if the set $\{x \mid I(x) \leqslant c\}$ is compact for every $c \geqslant 0$.

- the sequence $\{X_n\}_{n \geqslant 1}$ is *exponentially tight* at speed r_n if, for every $a \geqslant 0$, there exists a compact set $K_a \subseteq \mathcal{X}$ such that $\limsup_n r_n^{-1} \log \mathbb{P}[X_n \notin K_a] \leqslant -a$.

- the sequence $\{X_n\}_{n \geqslant 1}$ satisfies the *large deviation principle* with speed r_n and good rate-function I, denoted by

$$\mathbb{P}[X_n \approx a] \asymp e^{-r_n I(a)},$$

if, for every closed set $A \subseteq \mathcal{X}$, we have

$$\limsup_{n \to \infty} r_n^{-1} \log \mathbb{P}[X_n \in A] \leqslant - \inf_{x \in A} I(x),$$

and, for every open set $U \subseteq \mathcal{X}$,

$$\liminf_{n \to \infty} r_n^{-1} \log \mathbb{P}[X_n \in U] \geqslant - \inf_{x \in U} I(x).$$

We denote $\mathcal{AC}(E)$ for the set of absolutely continuous curves in E. For the sake of completeness, we recall the definition of absolute continuity.

Definition 2.2. A curve $\gamma : [0, T] \to E$ is absolutely continuous in E if there exists a function $g \in L^1([0, T], \mathbb{R}^d)$ such that for $t \in [0, T]$ we have $\gamma(t) = \gamma(0) + \int_0^t g(s)ds$. We write $g = \dot{\gamma}$. A curve $\gamma : \mathbb{R}^+ \to E$ is absolutely continuous in E if the restriction to $[0, T]$ is absolutely continuous for every $T \geqslant 0$.

Finally, we write $\mathbb{N} = \{0, 1, \dots\}$ for the set of natural numbers including 0.

2.2 Boundary structure of our state-spaces

For $x, v \in \mathbb{R}^d$, we write

$$\mathcal{B}_{x,v} := \left\{ x + w \mid w \in \mathbb{R}^d, \langle w, v \rangle \geqslant 0 \right\}, \quad \mathcal{B}_{x,v}^\circ := \left\{ x + w \mid w \in \mathbb{R}^d, \langle w, v \rangle > 0 \right\},$$

for the closed and open half-spaces based at x in the direction v.

Definition 2.3 (d dimensional convex Polyhedron)**.** We will say that E is a d *dimensional convex Polyhedron* if there are $x_1, \ldots, x_k, y_1, \ldots, y_l \in \mathbb{R}^d$ $v_1, \ldots, v_k, w_1, \ldots, w_l \in \mathbb{R}^d$, $v_i \neq v_j, v_i \neq w_j, w_i \neq w_j$ such that

$$E := \left(\bigcap_{i=1}^{k} \mathcal{B}_{x_i, v_i} \right) \cap \left(\bigcap_{j=1}^{l} \mathcal{B}_{y_j, w_j}^{\circ} \right). \tag{2.1}$$

Without loss of generality, we assume that the interior of E is non-empty. (Otherwise we perform a coordinate transform and work in a lower-dimensional space).

Definition 2.4 (Tangent spaces to a d dimensional convex Polyhedron)**.** Let E be a d dimensional convex Polyhedron.

(a) Let $\mathcal{J} = \mathcal{J}(E)$ be the set of subsets $J \subseteq \{1, \ldots, k\}$ such that

$$E_J := E \cap \bigcap_{i \in J} \left(\mathcal{B}_{x_i, v_i} \setminus \mathcal{B}_{x_i, v_i}^{\circ} \right) \neq \emptyset.$$

E_J is a $|J|$ dimensional face of E.

(b) For $J \in \mathcal{J}$ such that $J \neq \emptyset$ write $\Gamma_J := \bigcap_{i \in J} \mathcal{B}_{0, v_i}$, $\Gamma_\emptyset := \mathbb{R}^d$.

(c) For each $x \in E$ we define the set $T_x E := \bigcap_{J \in \mathcal{J} : x \in E_J} \Gamma_J$.

(d) We denote by $J^*(x)$ the set of largest cardinality J such that $x \in E_J$. Note that $T_x E = \Gamma_{J^*(x)}$.

(e) We write $TE = \left\{ (x, v) \in E \times \mathbb{R}^d \, \middle| \, v \in T_x E \right\}$. We equip TE with the subspace topology inherited from $E \times \mathbb{R}^d$.

For each point $x \in E$, the space $T_x E$ is the set of 'tangent vectors' at x: if $v \in T_x E$ then we have for small $\lambda > 0$ that $x + \lambda v \in E$.
We give four examples to clarify the definition.

Example 2.5 (Half line)**.** Let $E = [0, \infty)$. Then for $(x_1, v_1) = (0, 1)$, we have $E = B_{0,1}$. It follows that $\mathcal{J} = \{\emptyset, \{1\}\}$, $E_\emptyset = E$, $\Gamma_\emptyset = \mathbb{R}$, $E_{\{1\}} = \{0\}$, $E_{\{1\}} = \{0\}$ and $\Gamma_{\{1\}} = [0, \infty)$. The tangent spaces are given by $T_0 E = \Gamma_{\{1\}} = [0, \infty)$ and $T_x E = \Gamma_\emptyset = \mathbb{R}$.

Example 2.6 (Half open half closed interval)**.** Consider $E = [0, 1)$. Then for $(x_1, v_1) = (0, 1), (y_1, w_1) = (1, -1)$, we have $E = \mathcal{B}_{0,1} \cap \mathcal{B}_{1,-1}^{\circ}$. The set $\mathcal{J}$ equals $\mathcal{J} = \{\emptyset, \{1\}\}$ and $E_\emptyset = E$, $\Gamma_\emptyset = \mathbb{R}$ and $E_{\{1\}} = \{0\}$. The tangent spaces are given by $T_0 E = \Gamma_{\{1\}} = [0, \infty)$ and $T_x E = \mathbb{R}$ for $x \in (0, 1)$.

Example 2.7 (Quarter space)**.** Consider $E = [0, \infty)^2$. Then for $(x_1, v_1) = ((0,0), (1,0)), (x_2, v_2) = ((0,0), (0,1))$, we have $E = \mathcal{B}_{x_1, v_1} \cap \mathcal{B}_{x_2, v_2}$. The set $\mathcal{J}$ equals $\mathcal{J} = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$. Furthermore:

$$\begin{aligned}
E_\emptyset &= E, & \Gamma_\emptyset &= \mathbb{R}^2, \\
E_{\{1\}} &= \{(x, y) \mid x = 0, y \geqslant 0\}, & \Gamma_{\{1\}} &= \{(v_1, v_2) \mid v_1 \geqslant 0\}, \\
E_{\{2\}} &= \{(x, y) \mid x \geqslant 0, y = 0\}, & \Gamma_{\{2\}} &= \{(v_1, v_2) \mid v_2 \geqslant 0\}, \\
E_{\{1,2\}} &= \{(x, y) \mid x = 0, y = 0\}, & \Gamma_{\{1,2\}} &= \{(v_1, v_2) \mid v_1, v_2 \geqslant 0\}.
\end{aligned}$$

Thus, the tangent spaces are given by

$$\begin{aligned}
T_{(0,0)} E &= J_{\{1,2\}} = [0, \infty)^2, & \\
T_{(x,0)} E &= J_{\{1\}} = [0, \infty) \times \mathbb{R} & \text{if } x > 0, \\
T_{(0,y)} E &= J_{\{2\}} = \mathbb{R} \times [0, \infty) & \text{if } y > 0, \\
T_{(x,y)} E &= J_\emptyset = \mathbb{R}^2 & \text{if } x, y > 0.
\end{aligned}$$

Example 2.8 (Quarter space without corner)**.** Consider $E = [0, \infty)^2 \setminus \{(0,0)\}$. Then for $(x_1, v_1) = ((0,0), (1,0))$, $(x_2, v_2) = ((0,0), (0,1))$, $(y_1, w_1) = ((0,0), (1,1))$, we have $E = \mathcal{B}_{x_1, v_1} \cap \mathcal{B}_{x_2, v_2} \cap \mathcal{B}^\circ_{y_1, w_1}$. The set $\mathcal{J}$ equals $\mathcal{J} = \{\emptyset, \{1\}, \{2\}\}$. Furthermore:

$$E_\emptyset = E, \qquad\qquad\qquad \Gamma_\emptyset = \mathbb{R}^2,$$
$$E_{\{1\}} = \{(x,y) \,|\, x = 0, y < 0\}, \qquad \Gamma_{\{1\}} = \{(v_1, v_2) \,|\, v_1 \geqslant 0\},$$
$$E_{\{2\}} = \{(x,y) \,|\, x > 0, y = 0\}, \qquad \Gamma_{\{2\}} = \{(v_1, v_2) \,|\, v_2 \geqslant 0\}.$$

In this case, the tangent spaces are the same as in previous example, except for that the point $(0,0)$ is not included in the space E.

2.3 A generating collection of Hamiltonians and an associated Lagrangian

Our goal is to prove a large deviation principle for the dynamics of stochastic systems, and write the rate function in Lagrangian form. The Lagrangian will as usual be given as the Legendre transform of a Hamiltonian. Due to the discontinuous nature of the dynamics at the boundary, this Hamiltonian will be constructed as the maximum over Hamiltonians that take into account all possible behaviours of the dynamics close to the boundary. We end this section with a simple example illustrating the construction.

Definition 2.9. Let E be a d dimensional convex Polyhedron. Suppose for each $J \in \mathcal{J}$, there is a map $\mathcal{H}_J : E_J \times \mathbb{R}^d \to \mathbb{R}$ such that

(a) for each $x \in E_J$ the map $p \mapsto \mathcal{H}_J(x, p)$ is convex;

(b) the map $\mathcal{H}_J$ is continuously differentiable in (x, p);

(c) the gradient $\partial_p \mathcal{H}_J(x, p) \in \Gamma_J$;

We say that $\mathbb{H} := \{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ is a *generating set of Hamiltonians*. For a generating set of Hamiltonians, denote

$$\mathcal{H}_\dagger(x, p) := \max_{J \in \mathcal{J} : x \in E_J} \mathcal{H}_J(x, p), \qquad \mathcal{H}_\ddagger(x, p) := \min_{J \in \mathcal{J} : x \in E_J} \mathcal{H}_J(x, p).$$

We additionally write $H_\dagger \subseteq C^2_c(E) \times C_b(E)$ and $H_\ddagger \subseteq C^2_c(E) \times C_b(E)$ for the operators defined by $H_\dagger f(x) := \mathcal{H}_\dagger(x, \nabla f(x))$ and $H_\ddagger f(x) := \mathcal{H}_\ddagger(x, \nabla f(x))$.

Definition 2.10. Let $\mathbb{H} = \{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians.

(a) For each $J \in \mathcal{J}(E)$, define $\mathcal{L}_J : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$ and $\mathcal{L} : E \times \mathbb{R}^d \to [0, \infty]$ by

$$\mathcal{L}_J(x, v) = \sup_{p \in \mathbb{R}^d} \{pv - \mathcal{H}_J(x, p)\}, \qquad \mathcal{L}(x, v) = \sup_{p \in \mathbb{R}^d} \{pv - \mathcal{H}_\dagger(x, p)\}.$$

(b) Denote by $\widehat{\mathcal{L}} : E \times \mathbb{R}^d \to [0, \infty]$ the map

$$\widehat{\mathcal{L}}(x, v) := \inf \left\{ \sum_{J : x \in E_J} \lambda_J \mathcal{L}_J(x, v_J) \,\middle|\, \sum_{J : x \in E_J} \lambda_J v_J = v, \ \sum_{J : x \in E_J} \lambda_J = 1, \lambda_J \geqslant 0 \right\}.$$

We start with a collection of preliminary properties of $\mathcal{L}, \widehat{\mathcal{L}}$ and their Legendre duals.

Lemma 2.11. *Let $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians.*

(a) *For each $x \in E$ the map $v \mapsto \widehat{\mathcal{L}}(x, v)$ is the convex hull of the Lagrangians $v \mapsto \mathcal{L}_J(x, v)$, i.e. the largest convex function that lies below all $\mathcal{L}_J$'s;*

(b) For each $x \in E$, the map $v \mapsto \mathcal{L}(x, v)$ is the lower semi-continuous regularization of $v \mapsto \widehat{\mathcal{L}}(x, v)$.

Proof. (a) follows from [22, Theorem 5.6]. (b) follows from [22, Theorem 16.5]. $\square$

The lemma shows that our Lagrangian turns out to be similar to the one in [12] in the sense that it it can be obtained as the lower semi-continuous hull of the Lagrangians that are obtained for the different types of boundary behaviour. Note that we do not re-define this convex hull to equal ∞ when the speed points outward. This issue will be taken care of by only considering trajectories that stay inside the state-space, i.e. in $\mathcal{AC}(E)$.

We end this section by working out the Hamiltonians and Lagrangians for a simple example arising from the large deviations of a rescaled a birth-death process on $\mathbb{N}$ with homogeneous jump rates 1 for $k \mapsto k+1$ and 1 for $k \mapsto k-1$ (if $k > 0$).

Example 2.12. Let $E = [0, \infty)$, so that $E = B_{0,1}$. See Example 2.5 for further notation. As $\mathcal{J} = \{\emptyset, \{1\}\}$ our generating set of Hamiltonians has two elements:

$$\mathcal{H}_{\emptyset}(x, p) = [e^p - 1] + \left[e^{-p} - 1\right],$$
$$\mathcal{H}_{\{1\}}(x, p) = [e^p - 1].$$

$\mathcal{H}_{\emptyset}$ captures the behaviour of the process in the interior and consists of two terms corresponding to jumps up and down. $\mathcal{H}_{\{1\}}$ captures the behaviour at the boundary point $\{0\}$ and therefore only consists of jumps going up. Note that indeed $\partial_p \mathcal{H}_{\{1\}}(0, p) \subseteq [0, \infty)$.

We obtain $\mathcal{H}_{\dagger}(x, p) = \mathcal{H}_{\ddagger}(x, p) = \mathcal{H}_{\emptyset}(x, p)$ for $x > 0$ and

$$\mathcal{H}_{\dagger}(0, p) = \begin{cases} [e^p - 1] + [e^{-p} - 1] & \text{if } p \leqslant 0, \\ [e^p - 1] & \text{if } p \geqslant 0, \end{cases}$$

$$\mathcal{H}_{\ddagger}(0, p) = \begin{cases} [e^p - 1] & \text{if } p \leqslant 0, \\ [e^p - 1] + [e^{-p} - 1] & \text{if } p \geqslant 0. \end{cases}$$

Performing Legendre transformation, we obtain

$$\mathcal{L}_{\emptyset}(x, v) = v \log \left(\frac{v + \sqrt{v^2 + 4}}{2} \right) - \sqrt{v^2 + 4} + 2,$$

$$\mathcal{L}_{\{1\}}(0, v) = v \log v - v + 1.$$

We thus find $\mathcal{L}(x, v) = \mathcal{L}_{\emptyset}(x, v)$ for $x > 0$ and

$$\mathcal{L}(0, v) = \begin{cases} \mathcal{L}_{\emptyset}(0, v) & \text{if } v \leqslant 0, \\ 0 & \text{if } v \in [0, 1], \\ \mathcal{L}_{\{1\}}(0, v) & \text{if } v \leqslant 1. \end{cases}$$

Clearly, the part $\mathcal{L}(0, v)$ for $v < 0$ will not play a role in the final large deviation principle as we will exclude trajectories that leave $[0, \infty)$.

3 Large deviations for a collection of examples

Before diving into the general results in Sections 4 and 5, we give a variety of contexts in which these general results can be applied. The examples are not meant to give an exhaustive list, but rather illustrate that the general results are applicable in a variety of contexts.

We begin with a collection of examples for the one-dimensional setting in Section 3.1 and proceed with large deviations for interacting species in Section 3.3.

3.1 One dimensional examples

We start with three examples in which there do not occur discontinuities in the Hamiltonian at the boundary, i.e. $H_\dagger = H_\ddagger$:

(1) Large deviations for birth and death processes with immigration;

(2) Large deviations for birth processes with jumps of arbitrary size.

(3) Large deviations for simple epidemic models.

We state these results to illustrate the extent of applicability of our method. We will discuss in each of these three cases why the result complements the literature [14, 17, 20].

We proceed with birth and death processes with immigration and harvesting. Harvesting introduces a discontinuity at the boundary: harvesting is not possible if there are no individuals.

As a final example, we discuss birth and death processes without any form of immigration or harvesting at the boundary. In this setting our results do not apply, and we argue why indeed a result of the type given in all other examples is not expected to hold.

Finally, we mention that the one-dimensional results for the dynamic Curie-Weiss model in [16] fall within the framework introduced in this paper and that the results can also be applied to the Wright-Fisher and Moran models with positive mutation rate or for one-dimensional models in the analysis of queuing sytems.

3.1.1 Birth and death processes with immigration

For $n \in \mathbb{N}$, let X_n be a Markov process on $\mathbb{N}$ with jump rate $\lambda_n(q)$ corresponding to the birth of an individual, $\mu_n(q)$ for the death of one individual and $\rho_n(q)$ for the arrival of one individual through immigration if the population has size $q \in \mathbb{N}$. The definition of λ_n and μ_n implies $\lambda_n(0) = \mu_n(0) = 0$.

Thus, the process has generator :

$$\mathcal{A}_n f(q) = (\lambda_n(q) + \rho_n(q)) (f(q+1) - f(q)) + \mu_n(q) (f(q-1) - f(q))$$

To make sure that the martingale problem, see Corollary 8.3.2 in [13], is well-posed, we impose for each n that

$$\sup_{q \in \mathbb{N}} \frac{\lambda_n(q) + \rho_n(q) + \mu_n(q)}{1 + q} < \infty.$$

Theorem 3.1. *Let* $E = [0, \infty)$. *Suppose that there are positive and continuous function* $\lambda, \mu,$ *and* ρ, *such that for every compact set* $K \subseteq E$:

$$\lim_{n \to \infty} \sup_{x \in K, xn \in \mathbb{N}} \left| \frac{1}{n} \lambda_n(nx) - \lambda(x) \right| + \left| \frac{1}{n} \mu_n(nx) - \mu(x) \right| + \left| \frac{1}{n} \rho_n(nx) - \rho(x) \right| = 0. \quad (3.1)$$

Moreover suppose that :

(a) $\lambda(0) = \mu(0) = 0$, $\rho(0) > 0$ *and* λ *and* μ *are strictly positive on* $(0, \infty)$;

(b) *the birth rate* λ *and the immigration rate* ρ *satisfy* $\lambda(x) + \rho(x) = \mathcal{O}(x \log(x))$ *at* ∞;

(c) $\exists\, \delta > 0$ *such that* μ *is increasing on* $[0, \delta]$.

Let X_n be solutions to the martingale problem or $\mathcal{A}_n$. Suppose that $\frac{1}{n} X_n(0)$ satisfies a large deviation principle on $\mathbb{R}^+$ with speed n and with good rate function I_0. Then the process $t \mapsto \frac{1}{n} X_n(t)$ satisfies a large deviation principle on $D_{\mathbb{R}^+}(\mathbb{R}^+)$ with speed n and with good rate function I:

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{b} if \ \gamma \in \mathcal{AC}(\mathbb{R}^+), \\ \infty & otherwise. \end{cases}$$

The Lagrangian $\mathcal{L}$ is the Legendre transform of

$$\mathcal{H}(x, p) = (\lambda(x) + \rho(x)) \left[e^p - 1 \right] + \mu(x) \left[e^{-p} - 1 \right]. \tag{3.2}$$

Remark 3.2. The result is a variant (without the diffusion part) of Theorem 10.22 of [14] for Lévy processes on $\mathbb{R}^d$ adapted to the context of $[0, \infty)$. The result is a complement to [17, 20] due to the non-compactness of the state-space and possibly non-Lipschitz rates.

3.1.2 Growing populations

As in Section 3.1.1, we consider an evolving population X_n. In this setting however, we assume that the population can only grow and starts with non-zero size.

Each individual gives k offspring at a rate that depends on the population size: the rate at which individuals give k offspring depends on n: $\nu_{k,n}(q)$. The generator of this process, if it exists, is given by

$$\mathcal{A}_n f(q) = \sum_{k=1}^\infty q \nu_{k,n}(q) \left[f(q + k) - f(q) \right].$$

As the rates of this jump process are potentially unbounded, we impose for each n the conditions

$$\sup_q \sum_{k=1}^\infty q \nu_{k,n}(q) < \infty, \qquad \sup_q \sum_{k=1}^\infty k \nu_{k,n}(q) < \infty. \tag{3.3}$$

These two bounds and Corollary 8.3.2 in [13] imply that the martingale problem for $\mathcal{A}_n$ has a unique solution.

Theorem 3.3. Let $E = (0, \infty)$. Suppose (3.3) is satisfied. Suppose in addition that there are continuous functions ν_k on E such that for each compact set $K \subseteq E$, we have

$$\lim_{n \to \infty} \sup_{x \in K, xn \in \mathbb{N}} |\nu_{k,n}(nx) - \nu_k(x)| = 0. \tag{3.4}$$

In addition, the functions ν_k satisfy

(a) $\sum_{k=1}^\infty \nu_k(x) > 0$ for all $x \in (0, \infty)$;

(b) the function $x \mapsto \sum_{k=1}^\infty k \nu_k(x)$ is bounded;

(c) there is some $\alpha > 0$ such that

$$\sup_x \sum_{k=1}^\infty x \nu_k(x) \frac{k^2}{(1+x)^2} e^{\alpha k/(1+x)} < \infty.$$

Let X_n be solutions to the martingale problem for $\mathcal{A}_n$. Suppose that $\frac{1}{n} X_n(0)$ satisfies the large deviation principle on E with speed n and with good rate function I_0. Then

9

the processes $t \mapsto \frac{1}{n}X_n(t)$ *satisfy the large deviation principle on* $D_E(\mathbb{R}^+)$ *with good rate function*

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s))ds & \textit{if } \gamma \in \mathcal{AC}(0, \infty), \\ \infty & \textit{otherwise,} \end{cases}$$

where $\mathcal{L}(x, v) = \sup_p pv - \mathcal{H}(x, p)$ *and where* $H : E \times \mathbb{R} \to \mathbb{R}$ *is defined as*

$$\mathcal{H}(x, p) = \sum_{k=1}^\infty x\nu_k(x)\left[e^{kp} - 1\right]. \tag{3.5}$$

Remark 3.4. As above the result is a variant (without the diffusion part) of Theorem 10.22 of [14] for Lévy processes on $\mathbb{R}^d$ adapted to the context of $(0, \infty)$. Note that due to the possibly infinite number of jump types, non-compact state-space and uni-directionality of the model the results of [17, 20] do not apply.

Remark 3.5. The growth condition in (c) arises when controlling the growth of the large deviations of the process. It is satisfied in any case in which ν has finite support uniformly in x.

In the case that ν has unbounded support, a similar expression without the exponent would appear when controlling the range of the process without looking at the large deviations. This is achieved by shifting the linear part into a drift term, and then controlling the second order Taylor expansion. This is reminiscent of the situation for the characteristics of a Lévy process. On the large deviations scale, we have to control the exponential function, which leads to the exponential factor.

Example 3.6 (Yule process)**.** The Yule process [2] is defined by taking rates independent of n and $\nu_1(q) = 1$ and $\nu_k(q) = 0$ for all $k > 1$. The Hamiltonian has the simple form $\mathcal{H}(x, p) = x\left[e^p - 1\right]$. Conditions (a) and (b) of Theorem 3.3 are trivially satisfies, whereas for (c) every $\alpha > 0$ works.

Example 3.7 (Poisson β number of children)**.** Consider the case in which each individual has an offspring that has Poisson size with parameter β. In this case the Hamiltonian has the form

$$\mathcal{H}(x, p) = \sum_{k=1}^\infty x\frac{\beta^k}{k!}e^{-\beta}\left[e^{kp} - 1\right].$$

Condition (a) of Theorem 3.3 is trivially satisfied, whereas (b) is satisfied as we are considering a Poisson random variable.

For Condition (c), choose $\alpha = \frac{1}{2}\beta$. Note that

$$\sum_{k=1}^\infty x\frac{\beta^k}{k!}e^{-\beta}\frac{k^2}{(1+x)^2}e^{\frac{1}{2}\beta/(1+x)} \leqslant \sum_{k=1}^\infty \frac{k^2\beta^k}{k!}e^{-\frac{1}{2}\beta} < \infty.$$

Thus also for this case the large deviation principle holds.

3.1.3 Spread of a disease: a model with susceptible and infected individuals

We consider a simple SI model (with Susceptible and Infected individuals) for the spread of a disease in a population [1]. The population has constant size n, and is divided in two groups: X_n susceptible, and Y_n infected, individuals. Because the total size of the population is fixed, we write $Y_n = n - X_n$, so that the whole system can be described in terms of X_n.

The susceptible individuals contract the disease upon contact with an infected individual at rate $\frac{\beta}{n}$, $\beta > 0$, which leads to a generator:

$$\mathcal{G}_n f(q) = \frac{\beta}{n} q(n-q)\left(f(q+1) - f(q)\right), \qquad q \in \{0, \dots, n\}.$$

In general, we can replace the rate $\frac{\beta}{n} q(n-q)$ by a rate $c_n(q)$ that satisfies $c_n \geqslant 0$ and $c_n(0) = 0$:

$$\mathcal{A}_n f(q) = c_n(q)\left(f(q+1) - f(q)\right), \qquad q \in \{0, \dots, n\}.$$

Theorem 3.8. *Let* $E = (0,1]$. *Suppose that there is a non-negative continuous function* $c \in C(0,1]$ *such that for every compact set* $K \subseteq E$:

$$\lim_{n \to \infty} \sup_{x \in K, xn \in \mathbb{N}} \left| \frac{1}{n} c_n(nx) - c(x) \right| = 0.$$

Moreover suppose that :

1. $c(x) > 0$ *for* $x \in (0,1)$ *and* $c(1) = 0$,

2. c *is decreasing in a neighbourhood of* 1.

Let X_n *be solutions to the martingale problem for* $\mathcal{A}_n$. *Suppose that* $\frac{1}{n}X_n(0)$ *satisfies the large deviation principle on* E *with speed* n *and with good rate function* I_0. *Then the processes* $\frac{1}{n}X_n$ *satisfy the large deviation principle on* $D_E(\mathbb{R}^+)$ *with good rate function*

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}(0,1], \\ \infty & \text{otherwise}, \end{cases}$$

where $\mathcal{L}(x,v) = \sup_p pv - \mathcal{H}(x,p)$ *and where* $H : E \times \mathbb{R} \to \mathbb{R}$ *is defined as*

$$\mathcal{H}(x,p) = c(x)\left[e^p - 1\right]$$

Remark 3.9. c does not need to be bounded in a neighbourhood of 0.

Remark 3.10. As above, this result complements [17] in the sense that the state-space is non-compact, the model uni-directional and the rates are possibly non-Lipschitz and non-bounded.

3.1.4 Birth and death processes with immigration and harvesting

We adjust the example in Section 3.1.1 by adding harvesting.

Again, for $n \in \mathbb{N}$, let X_n be a Markov process on $\mathbb{N}$ with birth-rate $\lambda_n(q)$, death-rate $\mu_n(q)$ for immigration-rate $\rho_n(q)$. In addition, a breeder kills a single individual at a rate $\beta_n(q)$ that depends possibly on the population size q. To make sure that the boundary 0 can be accessed as well as exited sufficiently fast, we assume that there exists a $\varepsilon > 0$ such that

$$\inf_n \inf_{q \in \{0, \dots, \lfloor \varepsilon n \rfloor\}} \rho_n(q) \geqslant \varepsilon, \qquad \inf_n \inf_{q \in \{1, \dots, \lfloor \varepsilon n \rfloor\}} \beta_n(q) \geqslant \varepsilon.$$

As before, the generator reads

$$\mathcal{A}_n f(q) = \begin{cases} (\lambda_n(q) + \rho_n(q))(f(q+1) - f(q)) + (\mu_n(q) + \beta_n(q))(f(q-1) - f(q)) & \text{if } q \in \mathbb{N} \setminus \{0\}, \\ n\rho_n(0)(f(1) - f(0)) & \text{if } q = 0. \end{cases}$$

As before, we also impose for each n that

$$\sup_{q \in \mathbb{N}} \frac{\lambda_n(q) + \rho_n(q) + \mu_n(q) + \beta_n(q)}{1 + q} < \infty.$$

to ensure well-posedness of the martingale problems, cf. Corollary 8.3.2 in [13].

Theorem 3.11. *Let* $\mathsf{E} = [0, \infty)$. *Suppose that there are positive and continuous function* λ, μ, ρ, *and* β, *such that for every compact set* $\mathsf{K} \subseteq \mathsf{E}$:

$$\lim_{n \to \infty} \sup_{x \in \mathsf{K}: nx \in \mathbb{N}} \left\{ \left| \frac{1}{n} \lambda_n(nx) - \lambda(x) \right| + \left| \frac{1}{n} \mu_n(nx) - \mu(x) \right| \right.$$
$$\left. + \left| \frac{1}{n} \rho_n(nx) - \rho(x) \right| + \left| \frac{1}{n} \beta_n(nx) - \beta(x) \right| \right\} = 0. \quad (3.6)$$

Moreover suppose that :

(a) $\lambda(0) = \mu(0) = 0$, $\rho(0), \beta(0) > 0$ *and* λ *and* μ *are strictly positive on* E°;

(b) *the birth rate* λ *and the immigration rate* ρ *satisfy* $\lambda(x) + \rho(x) = \mathcal{O}(x \log(x))$ *at* ∞.

Suppose that $(\frac{1}{n} X_n(0))$ *satisfies a large deviation principle with speed* n *and with good rate function* I_0. *Then the process* $t \mapsto \frac{1}{n} X_n(t)$ *satisfies a large deviation principle on* $D_E(\mathbb{R}^+)$ *with speed* n *and good rate function* I:

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}(\mathbb{R}^+), \\ \infty & \text{otherwise.} \end{cases}$$

The Lagrangian is given in Definition 2.10 constructed from generating set of Hamiltonians $\{H_{\partial_-}, H_\emptyset\}$, $(\partial_- = 0)$:

$$\mathcal{H}_\emptyset(x, p) = (\lambda(x) + \rho(x)) \left[e^p - 1 \right] + (\mu(x) + \beta(x)) \left[e^{-p} - 1 \right], \quad (3.7)$$
$$\mathcal{H}_{\partial_-}(p) = \rho(0) \left[e^p - 1 \right]. \quad (3.8)$$

Remark 3.12. In this example, we work in the context of a Hamiltonian depending on the location $x \in \mathsf{E}$. In this context $\mathsf{E} = [0, \infty) = \mathcal{B}_{0,1}$. Thus $\mathcal{J}(\mathsf{E}) = \{1\}$ as the left boundary point 0 is included in $\mathsf{E} = [0, \infty)$. For intuitive understanding, we write ∂_- for the set $\{1\}$ as it corresponds as the set $\mathsf{E}_{\{1\}}$ equals the left boundary point $\{0\}$. To fulfil Definition 2.9, we need to specify continuously differentiable $\mathcal{H}_\emptyset : \mathsf{E} \times \mathbb{R} \to \mathbb{R}$ and $\mathcal{H}_{\{1\}} = \mathcal{H}_{\partial_-} : \{0\} \times \mathbb{R} \to \mathbb{R}$ such that $\partial_p \mathcal{H}_\emptyset(x, p) \subseteq \mathbb{R}$ and $\partial_p \mathcal{H}_{\partial_-}(0, p) \subseteq [0, \infty)$. This is indeed satisfied as can be read of from (3.7) and (3.8).

3.2 A discussion on the failure of the large deviation principle for pure birth and death processes

We end our discussion of one-dimensional processes by showing that a large deviation principle of the type considered above is not to be expected for birth and death processes without immigration or harvesting, or more generally without a mechanism by which the boundary and the interior show similar behaviour. Our example is of similar nature to that of Example E in [23].
Consider the process $X_n(t)$ with generator

$$\mathcal{A}_n f(q) = aq \left(f(q + 1) - f(q) \right) + bq \left(f(q - 1) - f(q) \right)$$

where $a > 0, b \geqslant 0$. Note that $a = 1, b = 0$ corresponds to the Yule process of Example 3.6.

We argue first that the large deviation principle like the ones in the other examples cannot hold for the Yule process on $[0, \infty)$, unlike on $(0, \infty)$ as in Section 3.1.2. If this result would be true then our theory would indicate that the Hamiltonian $\mathcal{H}$ corresponding to the principle has the form

$$\mathcal{H}(x, p) = x \left[e^p - 1 \right],$$

12

so that the path-space Lagrangian becomes

$$\mathcal{L}(x, v) = v \log \frac{v}{x} - v + x.$$

The trajectory $t \mapsto t^2$ can be seen to have finite cost. This indicates that the large deviations cost to go from 0 to 1 in time 1 is finite. However, if $X_n(0) = 0$ for all n, the large deviations cost to be in 1 at time 1 is infinity.

In the general setting, with

$$\mathcal{H}(x, p) = ax \left[e^p - 1 \right] + bx \left[e^{-p} - 1 \right],$$

$a > 0, b \geqslant 0$, one can show that the large deviations contribution of having no jumps of size -1 is finite. As the cost to go from 0 to 1 in finite time, just using positive jumps is finite as for the Yule process, we find that also in this case a contradiction if $X_n(0) = 0$ for all n.

This show that care is needed when formulating a large deviation principle for the rescaled population size of a pure birth-death process when allowing 0 as a starting point. Even though we have not talked about the technical ingredients for our proofs, we want to indicate that this implies that Condition 4.8 and the corresponding Theorem 4.11 for one-dimensional processes leaves out the pure birth-death processes for good reason.

3.3 Large deviations for interacting species with immigration and harvesting

We proceed our discussion with a general multi-dimensional model. In this setting, we only offer one set of results that we think covers a range of interesting models. More models can be treated on a model by model basis by using the methods of Sections 4 and 5.

For each n, we consider k different interacting species that evolve in time. We denote their population sizes at time t by $(X^1(t), \ldots, X^k(t))$.

We will denote possible transitions of the Markov process by vectors $\Gamma = (\Gamma_1, \ldots, \Gamma_l) \in \mathbb{Z}^k$ indicating how many individuals of each type get born of die. We write $\delta_i \in \mathbb{Z}^k$ for the vector with a one in location i and zero elsewhere. We consider three kinds of transitions, denoted by sets $T_{int}, T_{imm}, T_{har}$ corresponding to interaction between or within species, immigration and harvesting.

Condition 3.13. We assume that the cardinality of T_{int} is finite. In addition, we assume

interaction For each $\Gamma \in T_{int}$, we have a rate $r_{\Gamma,n}(q)$ at which the transition from q to $q + \Gamma$ occurs. Clearly, if there are no individuals of species i, then the rate of a transition in which individuals of species i die is 0. In other words, if there is an i with $\Gamma_i < 0$ and $q_i = 0$, then $r_{\Gamma,n}(q) = 0$.

immigration $T_{imm} = \{\delta_i \,|\, i \in \{1, \ldots, k\}\}$. The corresponding transition rate $a_{i,n}(q)$ is assumed to be strictly positive.

harvesting $T_{har} = \{-\delta_i \,|\, i \in \{1, \ldots, k\}\}$. The corresponding transition rate $b_{i,n}(q)$ is assumed to be strictly positive.

The rates lead to a generator A_n that reads

$$A_n f(q) = \sum_{\Gamma \in T_{int}} r_{\Gamma,n}(q) \left[f(q + \Gamma) - f(q) \right] + \sum_i a_{i,n}(q) \left[f(q + \delta_i) - f(q) \right]$$

$$+ \sum_i \mathbb{1}_{\{q_i \neq 0\}} b_{i,n}(q) \left[f(q - \delta_i) - f(q) \right]. \quad (3.9)$$

For well-posedness of the martingale problems, we impose for each n that

$$\sup_{q \in \mathbb{N}^k} \frac{\sum_{\Gamma \in T_{int}} r_{\Gamma,n}(q) + \sum_i a_{i,n}(q) + \sum_i b_{i,n}(q)}{1 + \sum_i q_i} < \infty. \tag{3.10}$$

Well posedness is established by Theorem 8.3.1 in [13] by taking $\gamma(q) = \eta(q) = 1 + \sum_i q_i$.

We will consider the large deviations on the space $E := [0, \infty)^k$ after rescaling the process by n^{-1}. For a momentum $p \in \mathbb{R}^d$, denote by $\Gamma \cdot p$ the vector obtained by component-wise multiplication $(\Gamma_1 p_1, \ldots, \Gamma_k p_k)$. Note that $\delta_i \cdot p = p_i \delta_i$.

Theorem 3.14. *Let $E := [0, \infty)^k$ and let Condition 3.13 be satisfied. Suppose that there are positive and continuous function r_Γ, a_i, b_i with $\Gamma \in T_{int}$ and $i \in \{1, \ldots, k\}$ so that for each compact set $K \subseteq E$:*

$$\limsup_{n \to \infty} \sup_{x \in K} \left\{ \sum_{\Gamma \in T_{int}} \left| \frac{1}{n} r_{\Gamma,n}(nx) - r_\Gamma(x) \right| + \sum_i \left| \frac{1}{n} a_{i,n}(nx) - a_i(x) \right| \right.$$

$$\left. + \sum_i \left| \frac{1}{n} b_{i,n}(nx) - b_i(x) \right| \right\} = 0. \tag{3.11}$$

Moreover suppose that a_i, b_i, $i \in \{1, \ldots, k\}$, are strictly positive and r_γ, $\Gamma \in T_{int}$, is non-negative.

Finally assume that for any $\Gamma \in T_{int}$ with $\sum_i \Gamma_i > 0$ we have $r_\Gamma(x) = \mathcal{O}(s(x) \log(s(x)))$ with $s(x) = \sum x_i$ at infinity. In addition, assume $a_i(x) + b_i(x) = \mathcal{O}(s(x) \log s(x))$ at infinity.

Let $(X_n^1, \ldots, X_n^k)$ be a solution to the martingale problem for A_n as in (3.9). Suppose that $(\frac{1}{n} X_n^1(0), \ldots, \frac{1}{n} X_n^k(0))$ satisfies a large deviation principle with speed n and with good rate function I_0.

Then the process $t \mapsto (\frac{1}{n} X_n^1(t), \ldots, \frac{1}{n} X_n^k(t))$ satisfies a large deviation principle on $D_E(\mathbb{R}^+)$ with speed n and good rate function I:

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}(\mathbb{R}^+, \mathbb{R}^+), \\ \infty & \text{otherwise.} \end{cases}$$

The Lagrangian is given in Definition 2.10 constructed from generating set of Hamiltonians $\mathbb{H}$ defined below. Write

$$\mathcal{H}_0(x, p) = \sum_{\Gamma \in T_{int}} r_\Gamma(x) \left[e^{\Gamma \cdot p} - 1 \right]$$

and define $\mathbb{H}$ as

$$\mathcal{H}_\emptyset(x, p) = \mathcal{H}_0(x, p) + \sum_{i=1}^d a_i(x)(e^{p_i} - 1) + b_i(x)(e^{-p_i} - 1),$$

and for $J \in \mathcal{J}(E)$ and $x \in E_J$

$$\mathcal{H}_J(x, p) = \mathcal{H}_0(x, p) + \sum_{i \in J} a_i(x)(e^{p_i} - 1) + \sum_{i \notin J} a_i(x)(e^{p_i} - 1) + b_i(x)(e^{-p_i} - 1).$$

Remark 3.15. As before, we specify the boundaries. As $E = [0, \infty)^k = \bigcap_{i=1}^k \mathcal{B}_{0,e_i}$, where e_i is the unit-vector in the i-th direction, the set $\mathcal{J}(E)$ equals the set of all subsets of $\{1, \ldots, k\}$. For $J \subseteq \{1, \ldots, k\}$ we have $E_J = \{x \in [0, \infty)^k \,|\, x_i = 0 \text{ if } i \in J\}$ and $\Gamma_J := \{p \in \mathbb{R}^k \,|\, p_i \geqslant 0 \text{ if } i \in J\}$.

14

To conclude this section, we describe a prey-predator system and a SI model. To apply the main result of this section, we assume in both cases that there is immigration and harvesting. Because interactions between different species bring rates that are more than linear in the coordinates (e.g two individuals of different species meeting will induce a quadratic rate in the coordinates), the martingale problem might not be well-posed (see 3.10). Therefore we introduce a "carrying capacity" κ that describe the maximal population of a given species the environment can sustain [21]. Once that population has grown to this carrying capacity, competition for resources will prevent reproduction. If that population grows above κ through immigration, the other species will still interact as if there was only κ individuals, the environment being saturated.

Example 3.16 (Prey and predator system)**.** We consider two populations of individuals whose sizes are given by the vector $\left(X_n^1(t), X_n^2(t)\right)$. The individuals of the second species hunt that of the first species. On top of some harvesting and immigration rates, the different interaction rates given the current population sizes $\left(X_n^1(t), X_n^2(t)\right) = (q_1, q_2)$ are given by:

Birth of a prey: The prey population grows with reproduction but is limited by competition. $\Gamma_1 = (1,0)$ with rate $r_{1,n}(q) = q_1(n\kappa - q_1)$ if $q_1 \leqslant n\kappa$ and $r_{1,n}(q) = 0$ otherwise.

Death of a prey: A prey dies upon meeting a predator at rate α. $\Gamma_1 = (-1,0)$ with rate $r_{2,n}(q) = \frac{\alpha}{n}q_1 q_2$ if $q_1 \leqslant n\kappa$ and $r_{2,n}(q) = \alpha\kappa q_2$ otherwise.

Birth of a predator: The predator population grows when a predator kills and eats prey at rate β. $\Gamma_3 = (0,1))$ with rate $r_{3,n}(q) = \frac{\beta}{n}q_1 q_2$ if $q_1 \leqslant n\kappa$ and $r_{3,n}(q) = \beta\kappa q_2$ otherwise. The cut-off corresponds to the maximum carrying capacity of the population of prey. That is: the population of prey is large enough so that the amount of food for the population of predators is essentially unlimited, implying that their reproduction rate only depends on their own population size.

Death of a predator: The predators live for an exponential time with mean μ. That is: $\Gamma_4 = (0,-1))$ and $r_{4,n}(q) = \mu q_2$.

The model satisfies the conditions of Theorem 3.14 with $\mathcal{H}_0$:

$$\mathcal{H}_0(x,p) = x_1(\kappa - (x_1 \wedge \kappa))(e^{p_1} - 1) + \alpha(x_1 \wedge \kappa)x_2\left(e^{-p_1} - 1\right)$$
$$+ \beta(x_1 \wedge \kappa)x_2\left(e^{p_2} - 1\right) + \mu y\left(e^{-p_2} - 1\right).$$

Example 3.17 (SI model with recovery and population dynamics)**.** Consider a population where individuals are divided into three groups: the susceptible, the infected, and the recovered immune individuals. We model the sizes of these three groups by the process (X_n^1, X_n^2, X_n^3). Denote $S(x) = x_1 + x_2 + x_3$. Consider the following interaction transitions for the process given $\left(X_n^1(t), X_n^2(t), X_n^3(t)\right) = (q_1, q_2, q_3)$:

Infection: A susceptible individual has a chance to fall sick upon meeting an infected individual . The transition vector reads $\Gamma_1 = (-1,1,0)$ and the transition rate is $r_{1,n}(q) = \frac{\beta}{n}(q_1 \wedge \kappa)q_2$.

Recovery: An infected individual recovers from the disease and gains immunity. $\Gamma_2 = (0,-1,1)$ and the transition rate is $r_{2,n}(q) = \alpha q_2$.

Birth of an individual: Suppose that individual are born susceptible, $\Gamma_3 = (1,0,0)$ and $r_{3,n}(q) = \frac{S(q)}{n}(n\kappa - S(q))$ if $s(q) \leqslant n\kappa$ and $r_{3,n}(q) = 0$ otherwise.

Death of an individual: for $k \in \{1,2,3\}$, we have transitions $\Gamma_{4,k} = -\delta_k$ with $r_{4,k,n}(q) = \mu_k q_k$.

The conditions for Theorem 3.14 hold with $\mathcal{H}_0$:

$$\mathcal{H}_0(x,p) = \beta(x_1 \wedge \kappa)x_2 \left(e^{p_2 - p_1} - 1\right) + \alpha x_2 \left(e^{p_3 - p_2} - 1\right)$$

$$+ S(x)(\kappa - (S(x) \wedge \kappa)) \left(e^{p_1} - 1\right) + \sum_{k=1}^{3} \mu_k x_k \left(e^{-p_k} - 1\right).$$

3.4 Discussion on extendability of the methods

In the beginning of Section 3 we stated that our results serve as an illustration of the scope of the general framework in the paper. We discuss shortly the extent of applicability of our methods.

For the one dimensional setting, one can consider models that combine features of the examples of Section 4.3. A quick scan of the proofs of these results shows that the three main conditions that need to be verified (notation and definitions will be introduced below) are:

- There needs to be a good containment function Υ for $\mathbb{H}$ (See Definition 4.2 below). A containment function is a type of Lyapunov function that implies that the process remains in compact sets on finite time intervals with large probability.

- We need to verify Conditions 4.7 and 4.8 for $\mathbb{H}$. These conditions can simply be read of from the generating set of Hamiltonians.

It should be noted that Condition 4.7 is not satisfied for the generating set of Hamiltonians that arises from the scaling limit of diffusions with reflecting boundary, see e.g. Sections 9.3 and 10.5 in [14]. A relaxation of this condition that covers both this case, as well as reflected diffusions, would be of interest.

For the multi-dimensional setting, we have only considered the case $E = [0,\infty)^d$ in Theorem 3.14 which in turn is immediately derived from Theorem 4.16.

Theorem 4.16 can be extended in a straightforward way by adapting Step 2 of its proof. The extension is restricted, however, to the context where the transitions that are discontinuous at the boundary are perpendicular to the boundary.

The only work known to the authors that includes a comparison principle for Hamiltonians with discontinuous rates for non-perpendicular jumps is [12]. In this setting, the rates, however, are spatially homogeneous. A comparison principle for a Hamiltonian that features a combination of inhomogeneous rates and discontinuity for non-perpendicular jumps would be of interest as well. What is necessary for such a substantial extension is unclear to the authors.

4 A general framework for Hamilton-Jacobi equations with a boundary

In the approach to large deviations by [14] and [12], also applied more recently in e.g. [?, 5, 10], the distributional information of the process at finite n related to the large deviations is encoded in the solutions f_n of a class of Hamilton-Jacobi equation $f - \lambda H_n f = h$, $\lambda > 0, h \in C_b(E_n)$. Given that the Hamiltonians H_n have a natural limiting upper bound $H_\dagger$ and lower bound $H_\ddagger$, semi-relaxed limits $\overline{f}$ and $\underline{f}$ of f_n, see (5.1) below, give a sub-solution and super-solution to

$$f - \lambda H_\dagger f = h, \qquad f - \lambda H_\ddagger f = h$$

respectively. To show that the sub- and super-solution coincide and give a proper 'solution' to this combination of equations, we need the comparison principle. This we carry out in the sections below.

Throughout the section, we assume that E is a convex polyhedron with non-empty interior.

4.1 Viscosity solutions to Hamilton-Jacobi equations

Definition 4.1 (Viscosity solutions). Let $H_\dagger \subseteq C_l(E) \times \mathcal{M}_u(E, \overline{\mathbb{R}})$ and $H_\ddagger \subseteq C_u(E) \times \mathcal{M}_l(E, \overline{\mathbb{R}})$ and let $\lambda > 0$ and $h \in C_b(E)$. Consider the Hamilton-Jacobi equations

$$f - \lambda H_\dagger f = h, \tag{4.1}$$
$$f - \lambda H_\ddagger f = h. \tag{4.2}$$

We say that u is a *(viscosity) subsolution* of equation (4.1) if u is bounded, upper semi-continuous and if, for every $f \in \mathcal{D}(H_\dagger)$ there exists a sequence $x_n \in \mathbb{R}$ such that

$$\lim_{n \uparrow \infty} u(x_n) - f(x_n) = \sup_x u(x) - f(x),$$
$$\lim_{n \uparrow \infty} u(x_n) - \lambda H_\dagger f(x_n) - h(x_n) \leqslant 0.$$

We say that v is a *(viscosity) supersolution* of equation (4.2) if v is bounded, lower semi-continuous and if, for every $f \in \mathcal{D}(H_\ddagger)$ there exists a sequence $x_n \in \mathbb{R}$ such that

$$\lim_{n \uparrow \infty} v(x_n) - f(x_n) = \inf_x v(x) - f(x),$$
$$\lim_{n \uparrow \infty} v(x_n) - \lambda H_\ddagger f(x_n) - h(x_n) \geqslant 0.$$

We say that u is a *(viscosity) solution* of equations (4.1) and (4.2) if it is both a subsolution to (4.1) and a supersolution to (4.2).

We say that (4.1) and (4.2) satisfy the *comparison principle* if for every subsolution u to (4.1) and supersolution v to (4.2), we have $u \leqslant v$.

We will study the Hamilton-Jacobi equations for the operators $H_\dagger, H_\ddagger$ that have been constructed from a generating set of Hamiltonians $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ as in Definition 2.9.

4.2 A general method to verify the comparison principle

In this section, we give the main technical results used in the text above that can be used to verify comparison principles. These methods are fairly standard, do not use any structure of our particular setting in a crucial way, and follow those in [5, 7, 14, 16]. The proofs of Proposition 4.5 and Lemma 4.6 do require minor adjustments of these methods and have therefore been given in Appendix A below. The results are based on a good containment function and a good penalization function.

Good containment functions play the role of a Lyapunov function and allow our analysis to be restricted to compact regions in E. The penalization functions are used in a distance like way.

Definition 4.2. Let $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians. We say that $\Upsilon : E \to \mathbb{R}$ is a *good containment function* (for $\mathbb{H} = \{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$) if

(Υa) $\Upsilon \geqslant 0$ and there exists a point x_0 such that $\Upsilon(x_0) = 0$,

(Υb) Υ is twice continuously differentiable,

17

(Υc) Υ has compact sub-level sets: for every $c \geqslant 0$, the set $\{x \in E \,|\, \Upsilon(x) \leqslant c\}$ is compact,

(Υd) we have $\max_{J \in \mathcal{J}(E)} \sup_{z \in E_J} \mathcal{H}_J(z, \nabla \Upsilon(z)) < \infty$.

Definition 4.3. We say that $\Psi : E^2 \to \mathbb{R}^+$ is a *good penalization functions* if

(Ψa) $\Psi(x, y) = 0$ if and only if $x = y$;

(Ψb) Ψ is twice continuously differentiable in both coordinates;

(Ψc) $(\nabla \Psi(\cdot, y))(x) = -(\nabla \Psi(x, \cdot))(y)$.

The first result can be found as Proposition 3.7 of [7] or Lemma 9.2 in [14].

Lemma 4.4. *Let u be bounded and upper semi-continuous, let v be bounded and lower semi-continuous, let $\Psi : E^2 \to \mathbb{R}^+$ be a good penalization function and let $\Upsilon : E \to \mathbb{R}^+$ have compact sub-level sets. Fix $\varepsilon > 0$. For every $\alpha > 0$ there exist points $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$, such that*

$$\frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \alpha\Psi(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\varepsilon})$$
$$= \sup_{x,y \in E} \left\{ \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \alpha\Psi(x,y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) \right\}.$$

Let $\{x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}\}_{\alpha,\varepsilon}$ be such a collection of points. For every $\varepsilon > 0$ we have

(a) *The set $\{x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \,|\, \alpha > 0\}$ is relatively compact in E.*

(b) *All limit points of $\{(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})\}_{\alpha>0}$ are of the form (z, z) and for these limit points we have $u(z) - v(z) = \sup_{x \in E}\{u(x) - v(x)\}$.*

(c) *We have*
$$\lim_{\alpha \to \infty} \alpha\Psi(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = 0.$$

The following result gives us the explicit condition that can be used to verify the comparison principle.

Proposition 4.5. *Let $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians. Let Υ be a good containment function for $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ and let Ψ be a good penalization function. Fix $\lambda > 0$, $h \in C_b(E)$ and consider u a subsolution to $f - \lambda H_\dagger f = h$ and v a supersolution to $f - \lambda H_\ddagger f = h$.*
For every $\alpha, \varepsilon > 0$ let $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$ be such that

$$\frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\varepsilon})$$
$$= \sup_{x,y \in E} \left\{ \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \Psi_\alpha(x,y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) \right\}. \quad (4.3)$$

Suppose that

$$\liminf_{\varepsilon \to 0} \liminf_{\alpha \to \infty} \mathcal{H}_\dagger\left(x_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right)$$
$$- \mathcal{H}_\ddagger\left(y_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right) \leqslant 0, \quad (4.4)$$

then $u \leqslant v$. In other words: the comparison principle holds for subsolutions to $f_\lambda H_\dagger f = h$ and supersolutions to $f - \lambda H_\ddagger f = h$.

We will establish (4.4) in Sections 4.3 and 4.4 below for a collection of Hamiltonians that appear in our applications.

The next lemma aids the verification of (4.4) by giving control on the sequences $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$. The result is an adaptation of Lemma 9.3 in [14]. For a slightly less involved variant, see Lemma 5 in [16].

Lemma 4.6. *Let $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians. Let $h \in C_b(E)$ and $\lambda > 0$ and let ν be a supersolution to $f - \lambda_{\ddagger} H = h$. Let Ψ be a good penalization function and Υ be a good containment function for $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$. Moreover, for every $\alpha, \varepsilon > 0$ let $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$ be as in (4.3). Then we have that*

$$\sup_{\varepsilon,\alpha} \mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha(\nabla\Psi(\cdot, y_{\alpha,\varepsilon}))(x_{\alpha,\varepsilon})) < \infty. \tag{4.5}$$

4.3 The comparison principle for one dimensional systems

We proceed by applying the general result for the verification of the comparison principle in the one-dimensional setting. For the verification, we need additional assumptions on the generating Hamiltonians. We start by interpreting Definitions 2.3, 2.4 and 2.9 in this one-dimensional setting.

Consider $E \subseteq \mathbb{R}$ be an interval and let ∂_-, ∂_+ be the boundaries of E in the set $\mathbb{R} \cup \{-\infty, \infty\}$. We will consider the setting that E is open, half open or closed. Clearly, E is of the form as in Definitions 2.3 and 2.4. We will use $-, +$ as indices instead of $1, \ldots, k$ for the sets that generate E.

In particular, if $\partial_- \in E$, then $E_- := \{\partial_-\}$ and $\Gamma_{\{-\}} = [0, \infty)$ and if $\partial_+ \in E$, then $E_+ := \{\partial_+\}$ and $\Gamma_{\{+\}} = (-\infty, 0]$.

In this setting, a collection of generating Hamiltonians takes a particularly simple form, there is a base Hamiltonian $\mathcal{H}_\emptyset$, as well as an Hamiltonian for each boundary point $z \in E \cap \partial E$. I.e. we have a generating set of Hamiltonians $\{\mathcal{H}_\emptyset\} \cup \{\mathcal{H}_z\}_{z \in \{-,+\}}$. Finally, condition (c) of Definition 2.9 translates into

$$\partial_- \in E \cap \partial E \qquad \Longrightarrow \qquad \partial_p \mathcal{H}_-(\partial_-, p) \geqslant 0,$$
$$\partial_+ \in E \cap \partial E \qquad \Longrightarrow \qquad \partial_p \mathcal{H}_+(\partial_+, p) \leqslant 0.$$

To prove the comparison principle, we will use the following conditions on the Hamiltonian.

Condition 4.7. Let $\{\mathcal{H}_\emptyset\} \cup \{\mathcal{H}_z\}_{z \in \{-,+\}}$ be a generating set of Hamiltonians.

Behaviour in the interior For each direction $d \in \{-, +\}$ and all compact $K \subseteq E^\circ$ we have either (1) or (2)

(1)
$$\lim_{p \to d\infty} \inf_{x \in K} \mathcal{H}_\emptyset(x, p) = \infty;$$

(2) there is a continuous function $h^d : K \to \mathbb{R}$ such that

$$\lim_{p \to d\infty} \sup_{x \in K} \left| \mathcal{H}_\emptyset(x, p) - h^d(x) \right| = 0.$$

Relating the boundary to the interior For each boundary $\partial_s \in E \cap \partial E$ it holds

$$-sp \geqslant 0 \qquad \Longrightarrow \qquad \mathcal{H}_s(\partial_s, p) \geqslant \mathcal{H}_\emptyset(\partial_s, p);$$
$$-sp \leqslant 0 \qquad \Longrightarrow \qquad \mathcal{H}_s(\partial_s, p) \leqslant \mathcal{H}_\emptyset(\partial_s, p).$$

The main step when verifying the comparison principle for points at the boundary is bounding the Hamiltonians $\mathcal{H}_\dagger$ and $\mathcal{H}_\ddagger$ by $\mathcal{H}_\emptyset$. This step allows us to proceed as if we were in the interior. We thus pose additional conditions for the behaviour of $H_\emptyset$ close to the boundary. We distinguish between a weak and a strong condition. The weak condition is sufficient for the verification of the comparison principle, whereas the strong condition plays an important role when we take sums of Hamiltonians, see Lemma 4.13 below.

Condition 4.8 (Conditions for the boundary). Let $\{\mathcal{H}_\emptyset\} \cup \{\mathcal{H}_z\}_{z \in \{-,+\}}$ be a generating set of Hamiltonians. For each boundary $s \in \{-,+\}$ such that $\partial_s \in \mathbb{R} \cap E$ and each direction $d \in \{-,+\}$ either the weak or the strong condition holds:

Strong condition There is a closed interval of positive length $K \subseteq E$ and $\partial_s \in K$ such that
$$\lim_{p \to d\infty} \inf_{x \in K} \mathcal{H}_\emptyset(x, p) = \infty.$$

Weak condition We have either (1) or (2):

(1) There is a closed interval of positive length $K \subseteq E$ and $\partial_s \in K$ and a continuous function h_s^d on K such that

$$\lim_{p \to d\infty} \sup_{x \in K} \left| \mathcal{H}_\emptyset(x, p) - h_s^d(x) \right| = 0.$$

(2) For $x_\alpha, y_\alpha \to \partial_s$ and $p_\alpha = \alpha(x_\alpha - y_\alpha) \to d\infty$, we have

$$\liminf_{\alpha \to \infty} \mathcal{H}_\emptyset(x_\alpha, p_\alpha) - \mathcal{H}_\emptyset(y_\alpha, p_\alpha) \leqslant 0.$$

Remark 4.9. Note that weak condition (1) implies weak condition (2), but is often easier to check. For the Hamiltonian $\mathcal{H}(x, p) = x\left[e^{-p} - 1\right]$ on the space $[0, \infty)$ and $s = -, d = -$, condition (2) holds as

$$\mathcal{H}_\emptyset(x_\alpha, \alpha(x_\alpha - y_\alpha)) - \mathcal{H}_\emptyset(y_\alpha, \alpha(x_\alpha - y_\alpha)) = [x_\alpha - y_\alpha]\left[e^{-\alpha(x_\alpha - y_\alpha)} - 1\right] \leqslant 0$$

due to the anti-symmetry of the occurrence of the difference $x_\alpha - y_\alpha$. The strong condition, nor the weak condition (1) holds.

Remark 4.10. Consider the Hamiltonian corresponding to the Yule process: $\mathcal{H}(x, p) = x\left[e^p - 1\right]$ on the space $[0, \infty)$. It follows that Condition 4.8 fails for $d = +$, corresponding with our discussion in Section 3.2

Theorem 4.11. *Let* $\mathbb{H} := \{\mathcal{H}_\emptyset\} \cup \{\mathcal{H}_z\}_{z \in \{-,+\}}$ *be a generating set of Hamiltonians. Suppose that* $\mathbb{H}$ *has a good containment function* Υ *and satisfies Conditions 4.7 and 4.8. Let* $h \in C_b(E)$ *and* $\lambda > 0$.
Then the comparison principle holds for subsolutions to $f - \lambda H_\dagger f = h$ *and supersolutions to* $f - \lambda H_\ddagger f = h$.

Proof of Theorem 4.11. We verify the conditions for Proposition 4.5 using the penalization function $\Psi(x, y) = \frac{1}{2}(x - y)^2$.
Thus, for $\alpha, \varepsilon > 0$ let $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$ be such that

$$\frac{u(x_{\alpha,\varepsilon})}{1 - \varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1 + \varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1 - \varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1 + \varepsilon}\Upsilon(y_{\alpha,\varepsilon})$$
$$= \sup_{x,y \in E} \left\{ \frac{u(x)}{1 - \varepsilon} - \frac{v(y)}{1 + \varepsilon} - \Psi_\alpha(x, y) - \frac{\varepsilon}{1 - \varepsilon}\Upsilon(x) - \frac{\varepsilon}{1 + \varepsilon}\Upsilon(y) \right\}.$$

We fix $\varepsilon > 0$ and prove that

$$\liminf_{\alpha \to \infty} \mathcal{H}_\dagger\left(x_{\alpha,\varepsilon}, \alpha \nabla \Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right) - \mathcal{H}_\ddagger\left(y_{\alpha,\varepsilon}, \alpha \nabla \Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right) \leqslant 0. \qquad (4.6)$$

This implies that (4.4) is satisfied and the comparison principle holds.

From this point onward, we drop $\varepsilon > 0$ from our notation. In addition, we write $p_\alpha := \alpha \nabla \Psi(\cdot, y_\alpha)(x_\alpha) = \alpha(x_\alpha - y_\alpha)$ to shorten formula's below. By Lemma 4.4, the sequences $\{x_\alpha, y_\alpha\}_{\alpha > 0}$ are contained in some compact set and all limit points are of the form (z, z). Thus, without loss of generality, we can assume that $(x_\alpha, y_\alpha) \to (z, z)$. The proof proceeds depending on whether z is in the interior or is a boundary point.

Step 1: Suppose that $z \in E^\circ$. Then there is some compact set $K \subseteq E^\circ$ with $z \in K$. Without loss of generality, we can assume that (x_α, y_α) contains a subsequence such that $p_\alpha \geqslant 0$. We will denote the subsequence also by α. We argue using Condition 4.7 holds for $d = +$ (In the case we have a subsequence with $p_\alpha \leqslant 0$, we would argue with $d = -$).

Suppose that Condition 4.7 (1) holds. Then it follows by Lemma 4.6 that $\sup_\alpha |p_\alpha| < \infty$. In this case, we can extract a converging subsequence with index $\hat{\alpha}$, which implies, using the continuity of $\mathcal{H}_\emptyset$ that

$$\lim_{\hat{\alpha} \to \infty} \mathcal{H}_\dagger\left(x_{\hat{\alpha}}, p_{\hat{\alpha}}\right) - \mathcal{H}_\dagger\left(y_{\hat{\alpha}}, p_{\hat{\alpha}}\right) = \lim_{\hat{\alpha} \to \infty} \mathcal{H}_\emptyset\left(x_{\hat{\alpha}}, p_{\hat{\alpha}}\right) - \mathcal{H}_\emptyset\left(y_{\hat{\alpha}}, p_{\hat{\alpha}}\right) = 0$$

implying (4.6).

Suppose that Condition 4.7 (2) holds. If as above, the sequence p_α contains a bounded subsequence, then (4.6) holds. If, however, $p_\alpha \to \infty$, then

$$\lim_{\hat{\alpha} \to \infty} \mathcal{H}_\dagger\left(x_{\hat{\alpha}}, p_{\hat{\alpha}}\right) - \mathcal{H}_\dagger\left(y_{\hat{\alpha}}, p_{\hat{\alpha}}\right) = \lim_{\hat{\alpha} \to \infty} \mathcal{H}_\emptyset\left(x_{\hat{\alpha}}, p_{\hat{\alpha}}\right) - \mathcal{H}_\emptyset\left(y_{\hat{\alpha}}, p_{\hat{\alpha}}\right) = h^+(z) - h^+(z) = 0$$

implying (4.6).

Step 2: Suppose that $z = \partial_-$. The case of $z = \partial_+$ is similar. We show that we can argue using the operator $\mathcal{H}_\emptyset$ instead of $\mathcal{H}_\dagger$ and $\mathcal{H}_\ddagger$. By Condition 4.7, we find that $p_\alpha \geqslant 0$ implies $\mathcal{H}_-(\partial_-, p_\alpha) \geqslant \mathcal{H}_\emptyset(\partial_-, p_\alpha)$, so that $\mathcal{H}_\ddagger(y_\alpha, p_\alpha) = \mathcal{H}_\emptyset(y_\alpha, p_\alpha)$. In addition, if $p_\alpha = 0$, then $\mathcal{H}_\dagger(x_\alpha, p_\alpha) = \mathcal{H}_\emptyset(x_\alpha, p_\alpha)$, whereas if $p_\alpha > 0$, then $x_\alpha > y_\alpha \geqslant \partial_-$, so that also $\mathcal{H}_\dagger(x_\alpha, p_\alpha) = \mathcal{H}_\emptyset(x_\alpha, p_\alpha)$. Similar arguments treat the case $p_\alpha \leqslant 0$. We conclude that for all $\alpha > 0$ we have

$$\mathcal{H}_\dagger(x_\alpha, p_\alpha) = H_\emptyset(x_\alpha, p_\alpha), \qquad \mathcal{H}_\ddagger(y_\alpha, p_\alpha) = \mathcal{H}_\emptyset(y_\alpha, p_\alpha).$$

This implies we can argue by using $\mathcal{H}_\emptyset$ instead of $\mathcal{H}_\dagger, \mathcal{H}_\ddagger$, which is analogous as to the proof in step 1. $\qquad \square$

To conclude this section on the comparison principle for one-dimensional systems, we show that satisfying Conditions 4.7 and 4.8 is stable under the formation of finite sums. This makes checking the conditions easy: if it holds for easily identifiable building blocks of a Hamiltonian, then it holds for the sum of these Hamiltonians.

Definition 4.12. Let $\mathbb{H}_i := \{\mathcal{H}_{J,i}\}_{J \in \mathcal{J}(E)}$, $i \in \{1, \ldots, k\}$ be generating sets of Hamiltonians. We define $\mathbb{H} := \sum_{i=1}^k \mathbb{H}_i$ to be the set of Hamiltonians $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$, where $\mathcal{H}_J = \sum_{i=1}^k \mathcal{H}_{J,k}$. Note that $\mathbb{H}$ is also a generating set of Hamiltonians.

Lemma 4.13. *Let* $\mathbb{H}_i := \{\mathcal{H}_{\emptyset,i}\} \cup \{\mathcal{H}_{z,i}\}_{z \in \{-,+\}}$, $i \in \{1, \ldots, k\}$ *be a generating sets of Hamiltonians that satisfy Condition 4.7. Suppose that for each* $s \in \{-,+\}$ *with* $\partial_s \in E \cap \partial E$ *we have (a) or (b):*

  (a) $\mathcal{H}_{\emptyset,1}, \ldots, \mathcal{H}_{\emptyset,k}$ *satisfy Condition 4.8.*

  (b) *There is* $i \in \{1, \ldots, k\}$ *such that* $\mathcal{H}_{\emptyset,i}$ *satisfies strong Condition 4.8.*

Then $\mathbb{H} := \mathbb{H}_1 + \cdots + \mathbb{H}_k$ *satisfies Conditions 4.7 and 4.8.*

Proof. All claims are immediate. $\qquad \square$

4.4 The comparison principle for a class of multi-dimensional systems

We proceed with giving easily verifiable conditions for a class of Hamiltonians on $E := [0, \infty)^d$ that corresponds to the large deviations of d interacting species. Regarding the boundary structure that we introduced in Section 2.2, we write $E_i = \{x \in E, x_i = 0\}$ for all i. For $J \subseteq \{1, \ldots, d\}$, we write $E_J := \bigcap_{i \in J} E_i$.

The dynamics that we will consider consists of two parts. The first part arises from interaction between individuals. This leads to a base Hamiltonian $\mathcal{H}_0$ that vanishes on the boundary. We will consider a second part that consists of immigration of individuals, and a second part that consists of harvesting. This final part introduces a discontinuity of the dynamics at the boundary.

Condition 4.14. We have a continuous map $\mathcal{H}_0(x, p)$ such that $p \mapsto \mathcal{H}_0(x, p)$ is convex for each $x \in E$. In addition, we have for each $i \in \{1, \ldots, d\}$ positive continuous functions $a_i, b_i : E \to \mathbb{R}$.
$\mathcal{H}_0$ and a_i, b_i satisfy either (a) or (b):

(a) (i) For each compact $K \subseteq E$ there is a $M_K \geqslant 0$ such that

$$\mathcal{H}_0(x, p) \geqslant -|p| M_K.$$

   (ii) For each $i \in \{1, \ldots, d\}$ the functions a_i, b_i are strictly positive continuous.

(b) (i) For each compact set $K \subseteq E^\circ$ there is some $M_K \geqslant 0$ such that

$$\lim_{|p| \to \infty} \inf_{x \in K} \frac{\mathcal{H}(x, p)}{|p|} \geqslant M_K.$$

   (ii) The functions a_i, b_i are strictly positive on a neighbourhood of ∂E.

The difference in conditions (a) and (b) is as follows: (b) requires at least linear growth of $\mathcal{H}_0$ but puts no conditions on a_i, b_i on the interior of E, whereas (a) relaxes the growth of $\mathcal{H}_0$ and puts conditions on a_i, b_i on the interior instead.

Remark 4.15. Note that a map $\mathcal{H}_0$ that is continuously differentiable immediately satisfies the lower bound in (a).(i) as $\mathcal{H}_0(x, p) \geqslant |p| |\partial_p \mathcal{H}_0(x, 0)|$.

We then consider the generating set of Hamiltonians $\mathbb{H}$ defined as

$$\mathcal{H}_\emptyset(x, p) = \mathcal{H}_0(x, p) + \sum_{k=1}^d a_k(x)(e^{p_k} - 1) + b_k(x)(e^{-p_k} - 1)$$

and for $J \in \mathcal{J}(E)$ and $x \in E_J$

$$\mathcal{H}_J(x, p) = \mathcal{H}_0(x, p) + \sum_{k \in J} a_k(x)(e^{p_k} - 1) + \sum_{l \notin J} a_l(x)(e^{p_l} - 1) + b_l(x)(e^{-p_l} - 1).$$

Theorem 4.16. *Let Condition 4.14 be satisfied for $\mathbb{H}$ and suppose there is a good containment function Υ for $\mathbb{H}$. Then the comparison principle holds for subsolutions to $f - \lambda H_\dagger f = h$ and supersolutions to $f - \lambda H_\ddagger f = h$.*

Proof. Before starting, note that Condition 4.14 implies that for any compact set $K \subseteq E$, $\forall\, 1 \leqslant j \leqslant d$,

$$\lim_{|p| \to \infty} \inf_{x \in K} \mathcal{H}_\emptyset(x, p) = \infty. \tag{4.7}$$

To establish the comparison principle, we use Proposition 4.5 with penalization function $\Upsilon(x,y) = \frac{1}{2}|x-y|^2$. Thus, for $\alpha, \varepsilon > 0$ let $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$ be such that

$$\frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\varepsilon})$$
$$= \sup_{x,y \in E} \left\{ \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \Psi_\alpha(x,y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) \right\}.$$

We fix $\varepsilon > 0$ and prove that

$$\liminf_{\alpha \to \infty} \mathcal{H}_\dagger\left(x_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right) - \mathcal{H}_\ddagger\left(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})\right) \leqslant 0. \qquad (4.8)$$

Henceforth, we drop the ε from the notation. By Lemma 4.4 the sequences (x_α, y_α) are contained in some compact set with limit points of the form (z,z). Up to a subsequence we can suppose $\{x_\alpha, y_\alpha\} \to (z,z)$. Write

$$p_\alpha = (p_{\alpha,1}, p_{\alpha,2}, ..., p_{\alpha,d}) := \alpha\nabla\Psi(\cdot, y_\alpha)(x_\alpha).$$

The rest of the proof depends on whether z belongs to the boundary or not.

Step 1: Suppose that $z \in E^\circ$. Then let $K \subseteq E^\circ$ a compact set with $z \in K$. Lemma 4.6 combined with (4.7) ensures that $\forall\, 1 \leqslant j \leqslant d,\, |p_{\alpha,j}| < \infty$. Thus going to a subsequence such that p_α converges, continuity of $\mathcal{H}_\emptyset$ implies that (4.8) holds.

Step 2: Suppose that $z \in \partial E$. Note that as $p_{\alpha,j} = \alpha(x_{\alpha,j} - y_{\alpha,j})$ for each j:

$$\{1 \leqslant j \leqslant d \,|\, x_{\alpha,j} = 0\} \subseteq \{1 \leqslant j \leqslant d \,|\, p_{\alpha,j} \leqslant 0\}.$$

Next, let $J \in \mathcal{J}(E)$ such that $x_\alpha \in E_J$. We find that $p_{\alpha,j} \leqslant 0$ and, hence,

$$\mathcal{H}_\emptyset(x_\alpha, p_\alpha) - \mathcal{H}_J(x_\alpha, p_\alpha) = \sum_{j \in J} b_j(x_\alpha)\left(e^{-p_{\alpha,j}} - 1\right) \geqslant 0.$$

As this holds for all J with $x_\alpha \in E_J$, we find $\mathcal{H}_\dagger(x_\alpha, p_\alpha) = \max_{J \in \mathcal{J}(E): x_\alpha \in E_J} \mathcal{H}_J(x_\alpha, p_\alpha) = \mathcal{H}_\emptyset(x_\alpha, p_\alpha)$. Similarly, one establishes $\mathcal{H}_\ddagger(y_\alpha, p_\alpha) = \mathcal{H}_\emptyset(y_\alpha, p_\alpha)$.

Thus (4.8) follows as in step 1. $\qquad \square$

5 Application to large deviations

In the extensive monograph, [14], Feng and Kurtz introduced a general method to prove large deviations with speed r_n for a sequence of Markov processes $\{X_n\}_{n \geqslant 1}$, having generators $\{A_n\}_{n \geqslant 1}$, via the well-posedness of associated Hamilton-Jacobi equations. See [?, ?] for a recent new proof of this result. The Feng-Kurtz strategy is based on three observations.

(1) By an extended variant of the projective limit theorem, the large deviation principle for processes follows from exponential tightness combined with large deviation principles for the finite dimensional distributions.

(2) The large deviation principle for finite dimensional distributions is established by a variant of Bryc's theorem, using the Markov property of the processes. It suffices to assume large deviations for time 0 and that the semigroup of conditional log-Laplace transforms

$$V_n(t)f(x) = \frac{1}{r_n}\log \mathbb{E}\left[e^{r_n f(X_n(t))} \,\Big|\, X_n(0) = x\right]$$

converge to a limiting semigroup $V(t)$.

(3) Using control theory, the limiting semigroup can be rewritten as a variational semigroup, which allows to rewrite the rate-function in Lagrangian form.

The convergence of semigroups for (2) uses non-linear semigroup theory. We discuss the connection of the convergence of semigroups to the comparison principles that we studied in Section 4.

For each n, we distinguish three objects of importance:

- The semigroup $\{V_n(t)\}_{t \geqslant 0}$;

- The formal generator $H_n = \frac{d}{dt}|_{t=0} V_n(t)f$ of the semigroup $\{V_n(t)\}_{t \geqslant 0}$ is given by $H_n f = r_n^{-1} e^{-nf} A_n e^{nf}$;

- The resolvents $\{R(\lambda)\}_{\lambda > 0}$, formally given by $R(\lambda) = (\mathbb{1} - \lambda H_n)^{-1}$.

The main step in the Trotter-Kato-Kurtz proof of the convergence of semigroups is that the convergence of resolvents implies the convergence of semigroups. This result is usually extended with the fact that the convergence of generators, and a limiting generator that is sufficiently big, implies the convergence of resolvents.

In our setting, it is not clear how to directly construct a limiting generator as the coefficients of the generator can be discontinuous at the boundary. As a consequence, it is also not clear how to establish the convergence of resolvents directly.

The theory of viscosity solutions provides us with an alternative. [14] introduce two operators $H_\dagger \subseteq ex - \mathrm{subLIM}_n H_n$, $H_\ddagger \subseteq ex - \mathrm{superLIM}_n H_n$ that serve as a limiting 'upper' and 'lower' bound for the operators H_n, see Definition 5.3 below. These upper and lower bounds can be used to obtain upper and lower bounds for the limiting resolvent: let h_n converge boundedly and uniformly on compacts to a function h, then one can show that the functions $\bar{f}, \underline{f}$ given by

$$
\begin{aligned}
\bar{f}(x) &:= \sup \left\{ \limsup_{n \to \infty} R_n(\lambda) h_n(x_n) \,\middle|\, x_n \to x \right\}, \\
\underline{f}(x) &:= \inf \left\{ \liminf_{n \to \infty} R_n(\lambda) h_n(x_n) \,\middle|\, x_n \to x \right\},
\end{aligned}
\tag{5.1}
$$

are a viscosity subsolution to $f - \lambda H_\dagger f = h$ and a viscosity supersolution to $f - \lambda H_\ddagger f = h$ respectively. Thus, if the comparison principle is satisfied, there is a unique limiting function $\bar{f} = \underline{f}$, which we will denote by $R(\lambda)h$. By the Crandall-Liggett theorem, [8], these resolvents generate a semigroup

$$
V(t)f = \lim_{n \to \infty} R\left(\frac{t}{n}, H\right)^n f,
$$

and by the Trotter-Kurtz approximation theorem we have $V_n(t) \to V(t)$.

To conclude this discussion: the comparison principle is sufficient to establish (2), which, combined with a standard verification of exponential tightness (1), is sufficient to establish path-space large deviations of a sequence of Markov processes.

Before discussing this results in Section 5.2, we introduce control theory in Section 5.1, so that (3) leads us to a Lagrangian form of the rate function. Using control theory, we follow [14] in introducing a variational semigroup $\{\mathbf{V}(t)\}_{t \geqslant 0}$, as well as a variational resolvent $\{\mathbf{R}(\lambda)\}_{\lambda > 0}$. It can be shown, cf. [14, Theorem 8.27] that the variational resolvent also yields subsolutions to $f - \lambda H_\dagger f = h$ and supersolutions to $f - \lambda H_\ddagger f = h$. Thus, the comparison principle establishes that the variational resolvent must equal the resolvent $R(\lambda)$. This in turn establishes the equality of semigroups, which leads to a Lagrangian form of the rate function.

5.1 Control theory

Let E be a d dimensional convex Polyhedron, let $\mathbb{H} = \{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians and let $\mathcal{L}_J : E_J \times \mathbb{R}^d \to [0, \infty]$ and $\mathcal{L} : E \times \mathbb{R}^d \to [0, \infty]$ the corresponding Lagrangians as defined in Definition 2.10. Using $\mathcal{L}$ we introduce a variational semigroup and resolvent:

$$\mathbf{V}(t)f(x) := \sup_{\substack{\gamma \in \mathcal{AC}, \\ \gamma(0) = x}} \left\{ f(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s)) d \right\},$$

$$\mathbf{R}(\lambda)f(x) := \limsup_{t \to \infty} \sup_{\substack{\gamma \in \mathcal{AC} \\ \gamma(0) = x}} \left\{ \int_0^t \lambda^{-1} e^{-\lambda^{-1} t} \left(f(\gamma(t)) - \int_0^s \mathcal{L}(\gamma(r), \dot{\gamma}(r)) dr \right) ds \right\}.$$

In the next two propositions, we establish the conditions that are needed for the application of the control theory component of [14, Theorem 8.27]. The first result can be used to establish the path-space compactness of the set of trajectories that start in a compact set and have uniformly bounded Lagrangian cost. The compactness of this set can be used to establish various properties of $\mathbf{V}$ and $\mathbf{R}$. The second result is crucial in establishing that the lower semi-continuous regularization of $\mathbf{R}(\lambda)h$ is a viscosity supersolution to the Hamilton-Jacobi equation $f - \lambda H_\ddagger f = h$. These properties are proven in [14], where the results of the propositions below are taken as an input for the theory. The outcomes of these are used to establish the variational expression of the rate-function in Theorem 5.8.

Proposition 5.1. *Let* $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ *be a generating set of Hamiltonians. Suppose there is a good containment function* Υ *for* $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$. *Then we have*

(a) $\mathcal{L} : E \times \mathbb{R}^d \to [0, \infty]$ *is lower semi-continuous and for each compact set* $K \subseteq E$ *and* $c \in \mathbb{R}$ *the set*
$$\left\{ (x, v) \in K \times \mathbb{R}^d \,\middle|\, \mathcal{L}(x, v) \leqslant c \right\}$$
is compact in $E \times \mathbb{R}^d$

(b) *For each compact* $K \subseteq E$, $T > 0$ *and* $0 \leqslant M < \infty$, *there exists a compact set* $K' = K'(K, T, M) \subseteq E$ *such that* $\gamma \in \mathcal{AC}$ *and* $\gamma(0) \in K$ *and*

$$\int_0^T \mathcal{L}(\gamma(s), \dot{\gamma}(s)) \, ds \leqslant M$$

implies $\gamma(t) \in K'$ *for all* $0 \leqslant t \leqslant T$.

(c) *For each* $f \in C_b^2(E)$ *and compact* $K \subseteq E$, *there exists a right-continuous non-decreasing function* $\psi_{f,K} : \mathbb{R}^+ \to \mathbb{R}^+$ *such that* $\lim_{r \to \infty} r^{-1} \psi_{f,K}(r) = 0$ *and*

$$|\langle \nabla f(x), v \rangle| \leqslant \psi_{f,K}(\mathcal{L}(x, v)), \qquad \forall (x, v) \in E \times \mathbb{R}^d, x \in K.$$

Proof. (a) $\mathcal{L} : E \times \mathbb{R}^d \to [0, \infty]$ is lower semi-continuous as it is the Legendre transform of $H_\ddagger$. The compactness of the level sets follows as in the proof of Lemma 2 in [16].

(b) The proof is a standard Lyapunov function proof. We follow [5]. Let γ be absolutely continuous with Lagrangian cost bounded by M and $\gamma(0) \in K$. Let $t \leqslant T$,

then

$$\Upsilon(\gamma(t)) = \Upsilon(\gamma(0)) + \int_0^t \langle \nabla\Upsilon(\gamma(s)), \dot\gamma(s)\rangle ds$$

$$\leqslant \Upsilon(\gamma(0)) + \int_0^t \mathcal{L}(\gamma(s), \dot\gamma(s)) + \mathcal{H}(\gamma(s), \nabla\Upsilon(\gamma(s))) ds$$

$$\leqslant \sup_{y\in K} \Upsilon(y) + M + \int_0^T \sup_z \sup_{J:z\in E_J} \mathcal{H}_J(z, \nabla\Upsilon(z)) ds$$

$$=: C < \infty.$$

Thus, we can take $K' = \{z \in E \,|\, \Upsilon(z) \leqslant C\}$.
(c) follows as in the proof of Lemma 10.21 in [14]. Note that as the H_J's are continuous, we have

$$\overline{\mathcal{H}}_K(c) = \sup_{|p|\leqslant c}\sup_{x\in K} \mathcal{H}_\dagger(x, p) \leqslant \sup_{|p|\leqslant c}\sup_{x\in K}\max_{J:x\in E_J}\mathcal{H}_J(x, p) < \infty,$$

which is an essential ingredient for the proof in [14]. $\qquad\square$

Proposition 5.2. *Let $\{\mathcal{H}_J\}_{J\in\mathcal{J}(E)}$ be a generating set of Hamiltonians. Suppose there is a good containment function Υ for $\{\mathcal{H}_J\}_{J\in\mathcal{J}(E)}$.*
Then there exists for each $x \in E$ and $f \in C_c^2(E)$ a $\gamma \in \mathcal{AC}$ with $\gamma(0) = x$ such that for all $t_1 < t_2$:

$$\int_{t_1}^{t_2} \mathcal{H}_\ddagger f(\gamma(s)) ds \leqslant \int_{t_1}^{t_2} \langle df(\gamma(s)), \dot\gamma(s)\rangle - \mathcal{L}(\gamma(s), \dot\gamma(s)) \, ds.$$

The proof of the proposition uses the theory of differential inclusions. In Appendix D, we give the relevant basic definitions and a result that establishes the existence of a solution to 'well behaved' differential inclusions.

Proof. Pick $f \in C_c^2(E)$ and $x \in E$.
Step 1: Suppose we have a solution γ_f to the differential inclusion

$$\dot\gamma_f(s) \in F_f(x) := \text{ch}\bigcup_{J:x\in E_J} \partial_p\mathcal{H}_J(\gamma_f(s), \nabla f(\gamma_f(s))), \qquad \gamma_f(0) = x. \qquad (5.2)$$

We show that

$$\int_{t_1}^{t_2} \mathcal{H}_\ddagger f(\gamma_f(s)) ds \leqslant \int_{t_1}^{t_2} \langle \nabla f(\gamma_f(s)), \dot\gamma_f(s)\rangle - \mathcal{L}(\gamma_f(s), \dot\gamma_f(s)) \, ds. \qquad (5.3)$$

It suffices to show that for almost every s, we have $\nabla f(\gamma_f(s))\dot\gamma_f(s) - \mathcal{L}(\gamma_f(s), \dot\gamma_f(s)) \geqslant \mathcal{H}_\ddagger f(\gamma_f(s))$. In particular, it suffices to do this for all times s at which $\dot\gamma_f(s) \in F_f(\gamma_f(s))$.
Thus, suppose $v \in F_f(y)$. Then there are λ_J and $v_J = \partial_p\mathcal{H}_J(y, \nabla f(y))$ such that $v = \sum \lambda_J v_J$. As $\mathcal{L}(y, v) \leqslant \widehat{\mathcal{L}}(y, v) \leqslant \sum \lambda_J \mathcal{L}_J(x, v_J)$:

$$\langle \nabla f(y), v\rangle - \mathcal{L}(y, v) \geqslant \sum \lambda_J \left(\langle \nabla f(y), v_J\rangle - \mathcal{L}_J(y, v_J)\right)$$

$$\geqslant \sum \lambda_J \mathcal{H}_J(y, \nabla f(y))$$

$$\geqslant \mathcal{H}_\ddagger(y, \nabla f(y)),$$

establishing (5.3).
Step 2: We construct a solution to the differential inclusion (5.2), for which we use Lemma D.4. Note that the Lemma assumes growth bounds on the size of F_f, that are not necessarily satisfied.

26

Step 2a: Therefore, we start with *a priori* control on the range of a solution. Let, as in step 1, γ_f be a solution to (5.2). (5.3) implies

$$\int_0^T \mathcal{L}(\gamma_f(s), \dot{\gamma}_f(s))ds \leqslant \int_0^T \langle \nabla f(\gamma_f(s)), \dot{\gamma}_f(s) \rangle - \mathcal{H}_{\ddagger} f(\gamma_f(s)) \, ds.$$

Note that as f has compact support in E, the map $x \mapsto \mathcal{H}_{\ddagger}(x, \nabla f(x))$ is bounded from below as all $\mathcal{H}_J$ are continuous, and F_f is bounded on compact sets, there is some $M \geqslant 0$ such that

$$\int_0^T \mathcal{L}(\gamma_f(s), \dot{\gamma}_f(s))ds \leqslant M.$$

We conclude by Proposition 5.1 (b) that the trajectory γ_f remains in some compact set $K' \subseteq E$.

Step 2b: By step 2a, it suffices to construct a solution to $\dot{\gamma} \in \hat{F}_f(\gamma)$, where $\hat{F}_f$ equals F_f on K', and is equal to 0 outside a neighbourhood of K'. For example, we can smoothly multiply F_f by a cut-off function.

Thus, we construct a solution using Lemma D.4. We will verify the conditions for F_f. Note that the modification above can be made such to preserve the conditions for this lemma.

We check the conditions (a)-(d). (a) is clear by definition. For (b), fix $x \in E$ and a neighbourhood $\mathcal{U}$ of $F(x)$. We write $J^* := J^*(x)$. As for each J the set E_J is closed in E by Definition 2.4, we can choose a neighbourhood $\mathcal{V}$ of x that is contained in the E interior of E_{J^*}. This implies that for $y \in \mathcal{V}_0$ the set $F_f(y)$ is the convex hull of $\partial_p \mathcal{H}_J(y, \nabla f(y))$ with $J \subseteq J^*$. Because $\partial_p \mathcal{H}_J$ is continuous for each such J by Definition 2.9 (b), we can find a smaller neighbourhood $\mathcal{V} \subseteq \mathcal{V}_0$ with $x \in \mathcal{V}$ so that $F_f(y) \subseteq \mathcal{U}$ for all $y \in \mathcal{V}$, establishing that F_f is upper semi-continuous.

For (c), first note that $T_E(x) = T_x E$. Let $J^* := J^*(x)$. By Definition 2.9 (c), we have $\partial_p \mathcal{H}_{J^*}(x, \nabla f(x)) \subseteq T_x E$ implying that $F_f(x) \cap T_x E \neq \emptyset$.

Finally, (d) follows from Definition 2.9 (d).

Thus, by Lemma D.4 there is a solution γ_f to the differential inclusion (5.2). $\qquad \square$

5.2 Synthesis: a general large deviation principle

To connect the Hamilton-Jacobi equation to the large deviation principle, we introduce some additional concepts. For $E \subseteq \mathbb{R}^d$ and $E_n \subseteq E$ for all n, we write $E = \lim_{n \to \infty} E_n$ if for every $x \in E$ there exist $x_n \in E_n$ such that $x_n \to x$.

For $f_n \in C_b(E_n)$ and $f \in C_b(E)$, we write $\text{LIM} f_n = f$ if $\sup_n |f_n| < \infty$ and if for each compact $K \subseteq E$, we have

$$\lim_{n \to \infty} \sup_{x \in K \cap E_n} |f_n(x) - f(x)| = 0.$$

Definition 5.3 (Condition 7.11 in [14]). Suppose that for each n we have an operator $H_n \subseteq C_b(E_n) \times C_b(E_n)$. Let $(v_n)_{n \geqslant 0}$ be a sequence of real numbers such that $v_n \uparrow \infty$.

The *extended sub-limit*, denoted by $ex - \text{subLIM}_n H_n$, is defined by the collection $(f, g) \in C_l(E) \times \mathcal{M}_u(E, \overline{\mathbb{R}})$ for which there exist $(f_n, g_n) \in H_n$ such that

$$\text{LIM} f_n \wedge c = f \wedge c, \qquad \forall c \in \mathbb{R}, \tag{5.4}$$

$$\sup_n \frac{1}{v_n} \log |g_n| < \infty, \qquad \sup_n \sup_{x \in E_n} g_n(x) < \infty, \tag{5.5}$$

and that, for every compact set $K \subseteq E$ and every sequence $z_n \in K \cap E_n$ satisfying $\lim_n z_n = z$ and $\lim_n f_n(z_n) = f(z) < \infty$, we have

$$\limsup_{n \uparrow \infty} g_n(z_n) \leqslant g^*(z). \tag{5.6}$$

The *extended super-limit*, denoted by $ex-\mathrm{superLIM}_n\, H_n$, is defined by the collection $(f,g) \in C_u(E) \times \mathcal{M}_l(E, \overline{\mathbb{R}})$ for which there exist $(f_n, g_n) \in H_n$ such that

$$\mathrm{LIM}\, f_n \vee c = f \vee c, \qquad \forall c \in \mathbb{R}, \tag{5.7}$$

$$\sup_n \frac{1}{\nu_n} \log |g_n| < \infty, \qquad \inf_n \inf_{x \in E_n} g_n(x) > -\infty, \tag{5.8}$$

and that, for every compact set $K \subseteq E$ and every sequence $z_n \in K \cap E_n$ satisfying $\lim_n z_n = z$ and $\lim_n f_n(z_n) = f(z) > -\infty$, we have

$$\liminf_{n\uparrow\infty} g_n(z_n) \geqslant g_*(z). \tag{5.9}$$

The *extended limit*, denoted by $ex - \mathrm{LIM}_n\, H_n$ is defined as $ex - \mathrm{subLIM}_n\, H_n \cap ex - \mathrm{superLIM}_n\, H_n$.

Remark 5.4. Suppose that $H \subseteq C_b(E) \times C_b(E)$ such that for all $(f,g) \in H$ there are $(f_n, g_n) \in H_n$ with $f = \mathrm{LIM}\, f_n$ and $g = \mathrm{LIM}\, g_n$, then $H \subseteq ex - \mathrm{LIM}_n\, H_n$.

All our large deviations statements will be based on the following assumption.

Assumption 5.5. Let $E_n \subseteq E$ be Polish subsets satisfying $E = \lim_{n\to\infty} E_n$.
Assume that for each $n \geqslant 1$, we have $A_n \subseteq C_b(E_n) \times C_b(E_n)$ and existence and uniqueness holds for the $D_{E_n}(\mathbb{R}^+)$ martingale problem for (A_n, μ) for each initial distribution $\mu \in \mathcal{P}(E_n)$. Letting $\mathbb{P}_y^n \in \mathcal{P}(D_{E_n}(\mathbb{R}^+))$ be the solution to (A_n, δ_y), the mapping $y \mapsto \mathbb{P}_y^n$ is measurable for the weak topology on $\mathcal{P}(D_{E_n}(\mathbb{R}^+))$. Let X_n be the solution to the martingale problem for A_n and set

$$H_n f = \frac{1}{r_n} e^{-r_n f} A_n e^{r_n f} \qquad e^{r_n f} \in \mathcal{D}(A_n),$$

for some sequence of speeds $\{r_n\}_{n\geqslant 1}$, with $\lim_{n\to\infty} r_n = \infty$. Let $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$ be a generating set of Hamiltonians as in Definition 2.9 and suppose that $H_\dagger \subseteq ex - \mathrm{subLIM}_n\, H_n$ and $H_\ddagger \subseteq ex - \mathrm{superLIM}_n\, H_n$. Note: by definition $C_c^2(E) \subseteq \mathcal{D}(H_\dagger) \cap \mathcal{D}(H_\ddagger)$.

Definition 5.6. Let $r_n > 0$ be a sequence of real numbers such that $r_n \to \infty$. We say that a sequence of processes $\{X_n\}_{n\geqslant 1}$ satisfies the *exponential compact containment condition* at speed $\{r_n\}_{n\geqslant 1}$ if for every $T > 0$ and $a \geqslant 0$, there exists a compact set $K_{a,T} \subseteq E$ such that

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbb{P}\left[X_n(t) \notin K_{a,T} \text{ for some } t \leqslant T\right] \leqslant -a.$$

The following result follows along the lines of Proposition A.15 in [5], whose proof is based on Lemma 4.22 in [14].

Proposition 5.7. *Consider the setting of Assumption 5.5. Suppose that $\{X_n(0)\}_{n\geqslant 1}$ is exponentially tight with speed $\{r_n\}_{n\geqslant 1}$.*
Suppose that Υ is a good containment function for $\{\mathcal{H}_J\}_{J\in\mathcal{J}(E)}$. Then the processes $\{X_n\}_{n\geqslant 1}$ satisfy the exponential compact containment condition at speed $\{r_n\}_{n\geqslant 1}$

Theorem 5.8. *Consider the setting of Assumption 5.5. Suppose that $X_n(0)$ satisfies a large deviation principle with speed $\{r_n\}_{n\geqslant 1}$ and good rate function I_0.*

(a) *Suppose that Υ is a good containment function for $\{\mathcal{H}_J\}_{J\in\mathcal{J}(E)}$. Then the processes $\{X_n\}_{n\geqslant 1}$ are exponentially tight with speed r_n in $D_E(\mathbb{R}^+)$.*

(b) *In addition to the assumption in (a), suppose that for each $\lambda > 0$ and $h \in C_b(E)$ the comparison principle is satisfied for subsolutions to $f - \lambda H_\dagger f = h$ and supersolutions to $f - \lambda H_\ddagger f = h$. Then the large deviation principle is satisfied with speed r_n for the processes X_n with good rate function I*

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) \, ds & \text{if } \gamma \in \mathcal{AC}(E), \\ \infty & \text{otherwise.} \end{cases}$$

Proof of Theorem 5.8. (a) follows from Proposition 5.7 and Corollary 4.19 in [14]. For the verification of condition(c) for the latter result, we use $F = C_c^2(E)$, whereas (d) is satisfied as $C_c^2(E) \subseteq \mathcal{D}(H_\dagger)$ by Assumption 5.5.

(b) follows from Theorem 8.27 and Corollary 8.28 in [14]. For the application of this result, we use $\mathbf{H}_\dagger = H_\dagger$, $\mathbf{H}_\ddagger = H_\ddagger$ and $\mathcal{A}f(x, \nu) = \langle \nabla f(x), \nu \rangle$. The conditions for the control theory part in [14] have been verified in Propositions 5.1 and 5.2. Note that the rate function in [14] still involves an infimum over control measures. As our Lagrangian is convex in the speed variable, Jensen's inequality gives the final form. $\qquad\square$

6 Proofs for the examples

We proceed by giving proofs for the results given in Section 3. These proofs rely on Theorem 5.8, which means that for each set of examples we have to provide a containment function and verify the comparison principle for the corresponding set of Hamilton-Jacobi equations.

6.1 One dimensional examples

For our one-dimensional examples, the construction of a containment function will be case dependent. This will also hold for the verification of the comparison principle, although all these verifications will be carried out using Theorem 4.11.

6.1.1 Birth and Death process with immigration, proof of Theorem 3.1

Remark 6.1. Using the framework of Section 4, we have $E = \mathcal{B}_{0,1}$, $\mathcal{J} = \{\partial_-, \emptyset\}$ and $E_\emptyset = E = [0, \infty)$, $E_{\partial_-} = \{\partial_-\}$ with $\partial_- = 0$. For $x > 0$, $T_x E = \mathbb{R}$ and $T_{\partial_-} E = \mathbb{R}^+$.

Lemma 6.2. *Consider the setting of Theorem 3.1. Then there is a good containment function Υ for $\mathcal{H}$.*

Proof. Let $\Upsilon : \mathbb{R}^+ \to \mathbb{R}^+$ be twice continuously differentiable with $\Upsilon(x) = \log(\log(x))$ for $x \geq e^e$ and Υ is zero somewhere in $[0, e^e)$. Plugging Υ' into $\mathcal{H}$ for $x > e^e$:

$$\mathcal{H}(x, \Upsilon'(x)) \leq (\lambda(x) + \rho(x)) \left(e^{\frac{1}{x \log(x)}} - 1 \right)$$

$$= \frac{\lambda(x) + \rho(x)}{x \log(x)} \left(\sum_{k=1}^\infty \frac{1}{k!} \frac{1}{(x \log(x))^{k-1}} \right) < \infty.$$

$\qquad\square$

Proof of Theorem 3.1. We apply Theorem 5.8. We start with the verification of Assumption 5.5. The space rescaled operators A_n are given by

$$A_n f(x)(x) = (\lambda_n(nx) + \rho_n(nx)) \left[f\left(x + \frac{1}{n}\right) - f(x) \right] + \mu_n(nx) \left[f\left(x - \frac{1}{n}\right) - f(x) \right].$$

As a consequence the operators $H_n = \frac{1}{n} e^{-nf} A_n e^{nf}$ have the form

$$H_n f(x) = \frac{1}{n} (\lambda_n(xn) + \rho_n(xn)) \left[e^{n\left(f\left(x + \frac{1}{n}\right) - f(x)\right)} - 1 \right] + \frac{1}{n} \mu_n(xn) \left[e^{n\left(f\left(x - \frac{1}{n}\right) - f(x)\right)} - 1 \right].$$

Taylor's theorem yields $H \subseteq ex - LIM_n H_n$ with H given by $Hf(x) = \mathcal{H}(x, f'(x))$ as in (3.2) and domain $\mathcal{D}(H) = C_c^2(\mathbb{R}^+)$. In particular, we have $H_\dagger = H_\ddagger = H$.

The condition in (a) of Theorem 5.8 follows by Lemma 6.2. Thus, we are left to verify the condition for (b): the comparison principle. For this, we use Theorem 4.11, for which we need to verify Conditions 4.7 and 4.8.

We start with the first one. Let $K \subseteq E^\circ$ be a compact set in the interior. Because both $(e^p - 1)$ and $(e^{-p} - 1)$ are bounded from below, we have $\lim_{p \to \pm\infty} \inf_{x \in K} \mathcal{H}(x, p) = \infty$. This yields Condition 4.7.

We proceed with Condition 4.8. Let $K \subseteq E$ a compact set with $\partial_- = 0 \in K$. Because $\rho(0) > 0$, we have $\lim_{p \to +\infty} \inf_{x \in K} \mathcal{H}(x, p) = \infty$. For $d = -$, we consider the weak condition (2). Let's pick x_α and y_α in E both converging to 0 such that $p_\alpha = \alpha(x_\alpha - y_\alpha) \to -\infty$. Then :

$$\mathcal{H}(x_\alpha, p_\alpha) - \mathcal{H}(y_\alpha, p_\alpha) \leqslant (\mu(x_\alpha) - \mu(y_\alpha))(e^{-p_\alpha} - 1) \leqslant 0.$$

Hence Condition 4.8 is satisfied and the comparison principle holds for the Hamilton-Jacobi equations for H. We conclude that the large deviation principle holds with Lagrangian rate function. $\qquad\square$

6.1.2 Growing populations, proof of Theorem 3.3

We first give a compact containment function.

Lemma 6.3. *Consider the setting of Theorem 3.3. Then there is a good containment function Υ for $\mathcal{H}$.*

Proof. We construct a containment function for the Hamiltonian by piecing together two parts that control the behaviour at 0 and at infinity respectively.
Control at 0: Pick $\Upsilon_0(x) = -\log x$. For $x \leqslant 1$, we have

$$\mathcal{H}(x, \Upsilon_0'(x)) = \sum_{k=1}^\infty x \nu_k(x) \left[e^{-kx^{-1}} - 1 \right] \leqslant 0.$$

Control at ∞: To control the behaviour at infinity, we consider Proposition B.1. For this we rewrite the Hamiltonian as

$$\mathcal{H}(x, p) = \sum_{k=1}^\infty x \nu_k(x) \left[e^{kp} - 1 \right] = p \sum_{k=1}^\infty k \nu_k(x) x + \sum_{k=1}^\infty x \nu_k(x) \left[e^{kp} - kp - 1 \right]$$

so that $b(x) = \sum_{k=1}^\infty k \nu_k(x) x$ and $\nu(x, \cdot) = \sum_{k=1}^\infty \nu_k(x) x \delta_{\{k\}}$. Thus, the bounds on b and ν in Proposition B.1 follow from Assumptions (b) and (c) on the functions ν_k. We obtain that there is a function Υ_∞ such that $\sup_x H(x, \Upsilon_\infty'(x)) < \infty$.
Combining the bounds: Let $\Upsilon : (0, \infty) \to \mathbb{R}^+$ be a twice continuously differentiable function that is equal to Υ_0 on the interval $(0, \frac{1}{2})$ and equal to Υ_∞ on $(1, \infty)$ and is equal to 0 somewhere in $(\frac{1}{2}, 1)$.
As Υ is twice continuously differentiable, $\mathcal{H}$ is continuous and

$$\sup_{x \in (0, \frac{1}{2}) \cup (1, \infty)} \mathcal{H}(x, \Upsilon'(x)) < \infty$$

we find $\sup_x \mathcal{H}(x, \Upsilon'(x)) < \infty$. $\qquad\square$

Proof of Theorem 3.3. We verify the conditions for Theorem 5.8. We start by checking Assumption 5.5. First of all, the generator of the process $X_n(t)$ is given by

$$A_n f(x) = n \sum_{k=1}^{\infty} x v_k (nx) \left[f \left(x + \frac{k}{n} \right) - f(x) \right].$$

Thus,

$$H_n f(x) = \frac{1}{n} e^{-nf(x)} \left(A_n e^{nf} \right)(x) = \sum_{k=1}^{\infty} x v_k (nx) \left[e^{n \left(f \left(x + \frac{k}{n} \right) - f(x) \right)} - 1 \right].$$

and we obtain by (3.4) that $H \subseteq \mathrm{LIM}\, H_n$, where H is the operator $\mathcal{D}(H) = C_c^2(E)$, $Hf(x) = \mathcal{H}(x, f'(x))$. We thus consider the setting that $H_\dagger = H_\ddagger = H$.

The condition in (a) of Theorem 5.8 follows from Lemma 6.3. We thus proceed with the verification of the condition in (b).

As before, we use Theorem 4.11. In this setting there is no boundary, thus, it suffices to verify Condition 4.7 for the interior. Fix a compact set $K \subseteq (0, \infty)$. First consider $d = +$. Then assumption (a) on the functions v_k, we find $\lim_{p \to \infty} \inf_{x \in K} \mathcal{H}(x, p) = \infty$, i.e. (1) of Condition 4.7 is satisfied. For $d = -$, (2) of Condition 4.7 is satisfied with $h^-(x) = -\sum_{k=1}^{\infty} x v_k(x)$.

Thus, the comparison principle holds by Theorem 4.11 and the large deviation principle follows as a result. $\qquad \square$

6.1.3 SI model, proof of Theorem 3.8

Lemma 6.4. *Consider the setting of Theorem 3.8. Then there is a good containment function Υ for $\mathcal{H}$.*

Proof. The only non-compactness issue arising for the SI model comes from the open boundary at 0. Therefore the construction that was used in the proof Lemma 6.3 to control the behaviour at 0 can be used here as well. $\qquad \square$

Proof of Theorem 3.8. The proof is similar to that of Theorem 3.3. The difference comes from the new boundary at $x = 1$. We verify Condition 4.8 for the right boundary for the Hamiltonian $\mathcal{H}(x, p) = c(x) [e^p - 1]$. For $d = +$, i.e. momenta away from the boundary, we have for $K = [\frac{1}{2}, 1]$ that

$$\lim_{p \to -\infty} \sup_{x \in K} |\mathcal{H}(x, p) + c(x)| = 0.$$

For $d = -$, i.e. momenta towards the boundary, note that if $x_\alpha, y_\alpha \to 1$ and $p_\alpha := \alpha(x_\alpha - y_\alpha) \to \infty$, then $x_\alpha > y_\alpha$ for large α. This implies that

$$\mathcal{H}(x_\alpha, p_\alpha) - \mathcal{H}(y_\alpha, p_\alpha) = [c(x_\alpha) - c(y_\alpha)] [e^{p_\alpha} - 1] \leqslant 0$$

as c is decreasing in a neighbourhood of 0. $\qquad \square$

6.1.4 Harvesting, proof of Theorem 3.11

Lemma 6.5. *Consider the setting of Theorem 3.11. Then there is a good containment function Υ for $\mathcal{H}$.*

Proof. The containment function introduced in the proof of Lemma 6.2 works also in this setting. $\qquad \square$

Proof of Theorem 3.11. We verify the conditions for Theorem 5.8. We start with Assumption 5.5. The tilted operator H_n for the process $t \mapsto \frac{1}{n} X_n(t)$ is given by:

$$H_n f(x) = \begin{cases} \frac{1}{n}(\lambda_n(x) + \rho_n(x)) \left[e^{f(x+\frac{1}{n})-f(x)} - 1 \right] + \frac{1}{n}(\mu_n(x) + \beta_n(x)) \left[e^{f(x-\frac{1}{n})-f(x)} - 1 \right] & \text{if } x > 0, \\ \frac{1}{n}\mu_n(x) \left[e^{f(0+\frac{1}{n})-f(0)} - 1 \right] & \text{if } x = 0. \end{cases}$$

Taking point-wise limits, this yields two limiting Hamiltonians: $\mathcal{H}_\emptyset$ and $\mathcal{H}_{\partial_-}$ given by 3.7 and 3.8 respectively. The operators $H_\dagger$ and $H_\ddagger$ defined in Definition 2.9 with domain $C_c^2(\mathbb{R}^+)$ are then

$$\mathcal{H}_\dagger(x, p) = \begin{cases} \mathcal{H}_\emptyset(x, p) & \text{if } x > 0 \\ \rho(0) \left[e^p - 1 \right] + (\beta(0) \left[e^{-p} - 1 \right] \vee 0) & \text{if } x = 0 \end{cases} \tag{6.1}$$

$$\mathcal{H}_\ddagger(x, p) = \begin{cases} \mathcal{H}_\emptyset(x, p) & \text{if } x > 0 \\ \rho(0) \left[e^p - 1 \right] + (\beta(0) \left[e^{-p} - 1 \right] \wedge 0) & \text{if } x = 0 \end{cases} \tag{6.2}$$

Next, we check that point-wise convergence can be strengthened to $H_\dagger \subseteq ex - \text{subLIM}_n H_n$. Pick $f \in C_c^2(\mathbb{R}^+)$ and write $f_n := f$ and $g_n := H_n f_n$. By our choice of f_n, g_n (5.5) and (5.4) are immediate. We show that for $z_n \to z$ that $\limsup_n g_n(z_n) \leqslant H_\dagger f(z)$.
Clearly, if $z \neq 0$ then $g_n(z_n) \to H_\emptyset f(z) = H_\dagger f(z)$. If $z = 0$, then the values $g_n(z_n)$ depend on whether $z_n = 0$ or not. In either case, a limiting upper bound is given by the maximum of the individual limits: $H_\emptyset f(0) \vee H_{\partial_-} f(0) = H_\dagger f(0)$. We conclude that indeed $\limsup_n g_n(z_n) \leqslant H_\dagger f(z)$.
As f was arbitrary, we conclude $H_\dagger \subseteq ex - \text{subLIM}_n H_n$. Similarly one proves $H_\ddagger \subseteq ex - \text{superLIM}_n H_n$

All other steps of the proof follow that of Theorem 3.1. The verification of the comparison principle changes slightly, for this we again use Theorem 4.11 and verify Conditions 4.7 and 4.8.
For the first condition of 4.7 as well as 4.8, note that as $\lambda, \mu > 0$ on the interior of E, ρ and α are continuous and satisfy $\rho(0), \alpha(0) > 0$, we have for each compact set $K \subseteq E$ that
$$\lim_{|p| \to \infty} \inf_{x \in K} \mathcal{H}_\emptyset(x, p) = \infty.$$

For the second condition of 4.7, note that the boundary at 0 is a left boundary, i.e. $s = -$. By (3.7) and (3.8), we have

$$\mathcal{H}_\emptyset(0, p) - \mathcal{H}_{\partial_-}(0, p) = \beta(0) \left[e^{-p} - 1 \right].$$

As $\beta(0) > 0$, the claim follows. We conclude that indeed Conditions 4.7 and 4.8 are satisfied. $\qquad \square$

6.2 Proof of Theorem 3.14

Before proving the result, we establish that there is a good containment function.

Lemma 6.6. *Consider the setting of Theorem 3.14. Then there is a good containment function for $\mathbb{H}$.*

Proof. Let $s(x) = \sum x_i$ and let Υ be twice continuously differentiable and equal to $\log \log s(x)$ for x with $s(x) \geqslant e^e$. The result then follows like Lemma 6.2. $\qquad \square$

Proof of Theorem 3.14. The result follows as in the proof of Theorem 3.11, using the containment function of Lemma 6.6 and the comparison principle obtained in Theorem 4.16. $\qquad \square$

A Proof of Proposition 4.5 and Lemma 4.6

The proof of Proposition 4.5 is a variant of Proposition A.9 in [5], itself inspired a combination of Lemma 9.3 in [14] and Lemma 2.3 in [10]. We reprove this result here as the boundary conditions introduces some additional issues involving the containment function that can be taken care of using the following lemma.

Lemma A.1. *Let $a_1, \ldots, a_k$ and $\delta_1, \ldots, \delta_k$ be constants in $\mathbb{R}$. Then we have*

$$\min_i \{a_i + \delta_i\} \leqslant \min_i a_i + \max_i \delta_i$$

Proof. Let j, l be such that $a_j + \delta_j = \min_i \{a_i + \delta_i\}$ and $a_l = \min_i a_i$. Then, we have $\min_i \{a_i + \delta_i\} = a_j + \delta_j \leqslant a_l + \max_i \delta_i = \min_i a_i + \max_i \delta_i$. $\qquad\square$

Proof of Proposition 4.5. Following the proof of Lemma A.8 in [5], we can extend the operators $H_\dagger, H_\ddagger$ to contain functions of the type

$$x \mapsto (1 - \varepsilon)\alpha\Psi(x, y) + \varepsilon\Upsilon(x) + c, \qquad y \mapsto -(1 + \varepsilon)\alpha\Psi(x, y) - \varepsilon\Upsilon(x) + c,$$

with $c \in \mathbb{R}$ and $\alpha, \varepsilon > 0$, respectively. Let $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$ such that (4.3) is satisfied. Then, for all α we obtain that

$$
\begin{aligned}
&\sup_x u(x) - v(x) \\
&= \lim_{\varepsilon \to 0} \sup_x \frac{u(x)}{1 - \varepsilon} - \frac{v(x)}{1 + \varepsilon} \\
&\leqslant \liminf_{\varepsilon \to 0} \sup_{x,y} \frac{u(x)}{1 - \varepsilon} - \frac{v(y)}{1 + \varepsilon} - \alpha\Psi(x, y) - \frac{\varepsilon}{1 - \varepsilon}\Upsilon(x) - \frac{\varepsilon}{1 + \varepsilon}\Upsilon(y) \\
&= \liminf_{\varepsilon \to 0} \frac{u(x_{\alpha,\varepsilon})}{1 - \varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1 + \varepsilon} - \alpha\Psi(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1 - \varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1 + \varepsilon}\Upsilon(y_{\alpha,\varepsilon}) \\
&\leqslant \liminf_{\varepsilon \to 0} \frac{u(x_{\alpha,\varepsilon})}{1 - \varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1 + \varepsilon}, \quad (\text{A.1})
\end{aligned}
$$

as Υ and Ψ are non-negative functions. Since u is a sub-solution to $f - \lambda H_\dagger f = h$ and v is a super-solution to $f - \lambda H_\ddagger f = h$, we find by our particular choice of $x_{\alpha,\varepsilon}$ and $y_{\alpha,\varepsilon}$ that

$$u(x_{\alpha,\varepsilon}) - \lambda\mathcal{H}_\dagger\left(x_{\alpha,\varepsilon}, (1 - \varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) + \varepsilon\nabla\Upsilon(x_{\alpha,\varepsilon})\right) \leqslant h(x_{\alpha,\varepsilon}), \quad (\text{A.2})$$

$$v(y_{\alpha,\varepsilon}) - \lambda\mathcal{H}_\ddagger\left(y_{\alpha,\varepsilon}, -(1 + \varepsilon)\alpha\nabla\Psi(x_{\alpha,\varepsilon}, \cdot)(y_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})\right) \geqslant h(y_{\alpha,\varepsilon}). \quad (\text{A.3})$$

For all $z \in E$, the map $p \mapsto \mathcal{H}_\dagger(z, p)$, being the maximum of convex maps, is convex. Thus, (A.2) implies that

$$
\begin{aligned}
u(x_{\alpha,\varepsilon}) &\leqslant h(x_{\alpha,\varepsilon}) + (1 - \varepsilon)\lambda\mathcal{H}_\dagger(x_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) + \varepsilon\lambda\mathcal{H}_\dagger(x_{\alpha,\varepsilon}, \Upsilon'(x_{\alpha,\varepsilon})) \\
&\leqslant h(x_{\alpha,\varepsilon}) + (1 - \varepsilon)\lambda H_\dagger(x_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) + \varepsilon\lambda c,
\end{aligned}
$$
$$(\text{A.4})$$

where $c := \max_i \sup_x \mathcal{H}_i(x, \Upsilon'(x)) < \infty$ by the uniform bound (Υd).

As far as the second inequality, first note that because Ψ is a good penalization function, we have $-(\nabla\Psi(x_{\alpha,\varepsilon}, \cdot))(y_{\alpha,\varepsilon}) = \nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})$. Next, we need a more sophisticated bound using the convexity of the maps $\{\mathcal{H}_J\}_{J \in \mathcal{J}(E)}$. First of all, for each J the convexity of $\mathcal{H}_J$ yields

$$
\begin{aligned}
&\mathcal{H}_J(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) \\
&\leqslant \frac{1}{1 + \varepsilon}\mathcal{H}_J(y_{\alpha,\varepsilon}, (1 + \varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) + \frac{\varepsilon}{1 + \varepsilon}\mathcal{H}_J(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon})).
\end{aligned}
$$

This implies, using Lemma A.1, that

$$
\mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}))
$$
$$
\leqslant \min_{J: y_{\alpha,\varepsilon} \in E_J} \left\{ \frac{1}{1+\varepsilon} \mathcal{H}_J(y_{\alpha,\varepsilon}, (1+\varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) \right.
$$
$$
\left. + \frac{\varepsilon}{1+\varepsilon} \mathcal{H}_J(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon})) \right\}
$$
$$
\leqslant \frac{\varepsilon}{1+\varepsilon} \mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, (1+\varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) + \frac{\varepsilon}{1+\varepsilon} c.
$$

Thus, (A.3) gives

$$
v(y_{\alpha,\varepsilon}) \geqslant h(y_{\alpha,\varepsilon}) + \lambda(1+\varepsilon)\mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) - \varepsilon\lambda c. \qquad (A.5)
$$

By combining (A.1) with (A.4) and (A.5), we find

$$
\sup_x u(x) - v(x)
$$
$$
\leqslant \liminf_{\varepsilon \to 0} \liminf_{\alpha \to \infty} \left\{ \frac{h(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{h(y_{\alpha,\varepsilon})}{1+\varepsilon} \right. \qquad (A.6)
$$
$$
+ \lambda \left(\frac{\varepsilon}{1-\varepsilon} + \frac{\varepsilon}{1+\varepsilon} \right) c \qquad (A.7)
$$
$$
\left. + \lambda \left[\mathcal{H}_{\dagger}(x_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) - \mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) \right] \right\}. \qquad (A.8)
$$

The term (A.8) vanishes by assumption, whereas (A.7) vanishes as $\varepsilon \downarrow 0$. Now observe that, for fixed ε and varying α, the sequence $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ takes its values in a compact set and hence admits converging subsequences. All these subsequences converge to points of the form (z, z). Therefore, as $\alpha \to \infty$, we find

$$
\liminf_{\varepsilon \to 0} \liminf_{\alpha \to \infty} \frac{h(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{h(y_{\alpha,\varepsilon})}{1+\varepsilon} \leqslant \liminf_{\varepsilon \to 0} |h| \frac{2\varepsilon}{1-\varepsilon^2} = 0,
$$

giving that also the term in (A.6) converges to zero.

We conclude that the comparison principle holds for $f - \lambda Hf = h$. $\qquad \square$

Proof of Lemma 4.6. Using that v is a super-solution to $f - \lambda H_{\ddagger}f = h$, we find that it is a super solution to the equation $f - \lambda \hat{H}_{\ddagger}f = h$, where $\hat{H}_{\ddagger}$ is super-extension, as above, of $H_{\dagger}$ that includes functions of the type $y \mapsto (-(1+\varepsilon)\alpha\Psi(x, y) - \varepsilon\Upsilon(y))$ in its domain, cf. Lemma A.8 of [5]. It follows that for the points $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$, we have

$$
\mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, (1+\varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) \leqslant \frac{v(y_{\alpha,\varepsilon}) - h(y_{\alpha,\varepsilon})}{\lambda} \leqslant \frac{|v - h|}{\lambda}.
$$

Similar to the use of the convexity of the functions $\mathcal{H}_J$ in the proof of Proposition 4.5, we find

$$
\mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}))
$$
$$
\leqslant \frac{1}{1+\varepsilon} \mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, (1+\varepsilon)\alpha\nabla\Psi(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon}))
$$
$$
+ \frac{\varepsilon}{1+\varepsilon} \max_J \mathcal{H}_J(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon}))
$$

which implies

$$
\sup_\alpha \mathcal{H}_{\ddagger}(y_{\alpha,\varepsilon}, \alpha(\nabla\Psi(\cdot, y_{\alpha,\varepsilon}))(x_{\alpha,\varepsilon})) \leqslant \frac{1}{1+\varepsilon} \left(\frac{|v - h|}{\lambda} + \varepsilon \sup_{z,J} \mathcal{H}_J(z, \nabla\Upsilon(z)) \right) < \infty.
$$

$\qquad \square$

B   A general method to constructing containment functions for the boundary at infinity

The following result has been proven in Example 4.23 in [14].

Proposition B.1. *Let* $\mathsf{G} \subseteq \mathbb{R}$ *be some subset and let* $\mathcal{H} : \mathsf{G} \times \mathbb{R} \to \mathbb{R}$ *defined by*

$$\mathcal{H}(x, p) = \frac{1}{2} a(x) p^2 + b(x) p + \int \left[e^{kp} - kp - 1 \right] \nu(x, dk).$$

Suppose there exists $\alpha > 0$ *such that*

$$\sup_x \int \left| \frac{z}{1 + |x|} \right|^2 e^{\alpha |z| / (1 + |x|)} \nu(x, dz) < \infty \tag{B.1}$$

and some $C > 0$ *such that*

$$x b(x) \leqslant C(1 + |x|^2), \qquad |a(x)| \leqslant C(1 + |x|^2).$$

Then there is some $\delta > 0$ *such that if* $\Upsilon(x) := \delta \log(1 + x^2)$, *then*

$$\sup_{x \in \mathsf{G}} \mathcal{H}(x, \Upsilon'(x)) < \infty.$$

C   Convex analysis

Definition C.1. Let $\phi : \mathbb{R}^d \to (-\infty, \infty]$ be a convex function. We write

$$\partial \phi(p) := \left\{ v \in \mathbb{R}^d \,\middle|\, \phi(p') \geqslant \phi(p) + \langle p' - p, v \rangle \right\},$$

for the (set-valued) sub-differential of ϕ at p.

Lemma C.2 (Theorem 23.4 in [22]). *If* $\phi(p) = \infty$ *then* $\partial \phi(p) = \emptyset$.

Proposition C.3 (Corollary 23.5.1 in [22]). *Let* ϕ *be a lower semi-continuous function, and* ϕ^* *its Legendre transform. Then* $v \in \partial \phi(p)$ *if and only if* $p \in \partial \phi^*(v)$.

D   Differential inclusions

We follow [9, 18]. Let $D \subseteq \mathbb{R}^d$ be a non-empty set. A *multi-valued mapping* $\mathsf{F} : D \to 2^{\mathbb{R}^d} \setminus \{\emptyset\}$ is a map that assigns to every $x \in D$ a set $\mathsf{F}(x) \subseteq \mathbb{R}^d$, $\mathsf{F}(x) \neq \emptyset$.

Definition D.1. Let $I \subseteq \mathbb{R}$ be an interval with $0 \in I$, $D \subseteq \mathbb{R}^d$, $x \in D$ and $\mathsf{F} : D \to 2^{\mathbb{R}^d} \setminus \emptyset$ a multi-valued mapping. A function γ such that

(a) $\gamma : I \to D$ is absolutely continuous,

(b) $\gamma(0) = x$,

(c) $\dot{\gamma}(t) \in \mathsf{F}(\gamma(t))$ for almost every $t \in I$

is called a *solution of the differential inclusion* $\dot{\gamma} \in \mathsf{F}(\gamma)$ a.e., $\gamma(0) = x$.

If we assume sufficient regularity on the multi-valued mapping F, we can ensure the existence of a solution to differential inclusions that remain inside D.

Definition D.2. Let $D \subseteq \mathbb{R}^d$ be a non-empty set and let $\mathsf{F} : D \to 2^{\mathbb{R}^d} \setminus \{\emptyset\}$ be a multi-valued mapping.

(i) We say that F is *closed, compact or convex valued* if each set $F(x)$, $x \in D$ is closed, compact or convex, respectively.

(ii) We say that F is *upper semi-continuous at* $x \in D$ if for each neighbourhood $\mathcal{U}$ of $F(x)$, there is a neighbourhood $\mathcal{V}$ of x in D such that $F(\mathcal{V}) \subseteq \mathcal{U}$. We say that F is *upper semi-continuous* if it is upper semi-continuous at every point.

Definition D.3. Let $D \subseteq \mathbb{R}^d$ be a closed non-empty set. The tangent cone to D at x is
$$T_D(x) := \left\{ z \in \mathbb{R}^d \;\middle|\; \liminf_{\lambda \downarrow 0} \frac{d(y + \lambda z, D)}{\lambda} = 0 \right\}.$$

The set $T_D(x)$ is sometimes called the the *Bouligand cotingent cone.*

Lemma D.4 (Theorem 2.2.1 in [18], Lemma 5.1 in [9])**.** *Let $D \subseteq \mathbb{R}^d$ be closed and let $F : D \to 2^{\mathbb{R}^d} \setminus \{\emptyset\}$ satisfy*

(a) F has closed convex values and is upper semi-continuous;

(b) F is upper semi-continuous;

(c) for every x, we have $F(x) \cap T_D(x) \neq \emptyset$;

(d) F has bounded growth: there is some $c > 0$ such that $|F(x)| = \sup\{|z| \,|\, z \in F(x)\} \leqslant c(1 + |x|)$ for all $x \in D$.

Then the differential inclusion $\dot{\gamma} \in F(\gamma)$ has a solution on $\mathbb{R}^+$ for every starting point $x \in D$.

References

[1] L. J. Allen. *An introduction to stochastic processes with applications to biology. 2nd edition.* CRC Press, 2010.

[2] K. B. Athreya and P. E. Ney. *Branching processes.* Dover Publications, 2004.

[3] V. Bansaye and S. Méléard. *Stochastic Models for Structured Populations.* Springer International Publishing, 2015.

[4] T. Chan. Large deviations and quasi-stationarity for density-dependent birth-death processes. *The Journal of The Australian Mathematical Society. Series B. Applied Mathematics*, 40:238–256, 10 1998.

[5] F. Collet and R. C. Kraaij. Dynamical moderate deviations for the Curie-Weiss model. *Stochastic Processes and their Applications*, 127(9):2900 – 2925, 2017.

[6] C. Coron, M. Costa, H. Leman, and C. Smadi. A stochastic model for speciation by mating preferences. *Journal of Mathematical Biology*, 9 2017.

[7] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc., New Ser.*, 27(1):1–67, 1992.

[8] M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general banach spaces. *American Journal of Mathematics*, 93(2):pp. 265–298, 1971.

[9] K. Deimling. *Multivalued differential equations*, volume 1 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1992.

[10] X. Deng, J. Feng, and Y. Liu. A singular 1-D Hamilton-Jacobi equation, with application to large deviation of diffusions. *Communications in Mathematical Sciences*, 9(1), 2011.

[11] P. Dupuis and R. S. Ellis. Large deviations for Markov processes with discontinuous statistics, ii: random walks. *Probability Theory and Related Fields*, 91(2):153–194, 6 1992.

[12] P. Dupuis, H. Ishii, and H. M. Soner. A viscosity solution approach to the asymptotic analysis of queueing systems. *The Annals of Probability*, 18(1):pp. 226–255, 1990.

[13] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and Convergence.* Wiley, 1986.

[14] J. Feng and T. G. Kurtz. *Large Deviations for Stochastic Processes.* American Mathematical Society, 2006.

[15] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems.* Springer-Verlag, second edition, 1998.

[16] R. C. Kraaij. Large deviations for finite state Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton–Jacobi equation. *Journal of Statistical Physics*, 164(2):321–345, 2016.

[17] P. Kratz and E. Pardoux. Large deviations for infectious diseases models. *preprint; ArXiv:1602.02803*, 2016.

[18] M. Kunze. *Non-Smooth Dynamical Systems.* Number nr. 1744 in Lecture Notes in Mathematics. Springer, 2000.

[19] J. Müller and C. Kuttle. *Methods and Models in Mathematical Biology.* Springer-Verlag Berlin Heidelberg, 2015.

[20] E. Pardoux and B. Samegni-Kepgnou. Large deviation principle for epidemic models. *Journal of Applied Probability*, 54(3):905?920, 2017.

[21] E. Renshaw. *Stochastic Population Processes.* Oxford, University Press, 2011.

[22] R. T. Rockafellar. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[23] A. Shwartz and A. Weiss. Large deviations with diminishing rates. *Mathematics of Operations Research*, 30(2):281–310, 2005.