From PDEs to Probability: Hamilton-Jacobi equations and Large Deviations

Della Corte, S.

DOI
10.4233/uuid:e0bd7657-3d84-445d-a850-7726d1f71bbe

Publication date
2025

Document Version
Final published version

Citation (APA)
Della Corte, S. (2025). *From PDEs to Probability: Hamilton-Jacobi equations and Large Deviations*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:e0bd7657-3d84-445d-a850-7726d1f71bbe

From PDEs to Probability: Hamilton-Jacobi Equations and Large Deviations

Serena Della Corte

FROM PDEs TO PROBABILITY: HAMILTON-JACOBI EQUATIONS AND LARGE DEVIATIONS

COLOPHON

ISBN 978-94-6518-111-0

Cover by: Inga Hamelmann

Printed by: Gildeprint

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

FROM PDEs TO PROBABILITY: HAMILTON-JACOBI EQUATIONS AND LARGE DEVIATIONS

DISSERTATION

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Tuesday, the 11th of November 2025 at 12:30 o'clock

by

Serena DELLA CORTE

Master of Science in Mathematics
University of Naples "Federico II", Italy

born in Naples, Italy

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. F.H.J. Redig,	Delft University of Technology, *promotor*
Dr. R.C. Kraaij,	Delft University of Technology, *copromotor*

Independent members:

Prof. dr. J.M.A.M. van Neerven,	
	Delft University of Technology
Dr. H. Yoldas,	Delft University of Technology
Dr. D. Tonon,	University of Padova
Dr. P. Nyquist,	Chalmers University of Technology
Prof. dr. A. Papapantoleon,	Delft University of Technology, reserve member

This work was supported by The Netherlands Organisation for Scientific Research (NWO), grant number 613.009.148.

We work in the dark - we do what
we can - we give what we have.
Our doubt is our passion, and our
passion is our task.
The rest is the madness of art.

Henry James, The Middle Years.

CONTENTS

Part I

INTRODUCTION AND THEORETICAL BACKGROUND

1 INTRODUCTION TO HAMILTON-JACOBI EQUATIONS

The study of Hamilton-Jacobi equations has garnered significant interest in recent years due to its wide-ranging applications across numerous fields. These equations emerge from various contexts, including control problems, mean–field games and large deviation theory.

In this thesis we focuses on well-posedness of Hamilton-Jacobi equations and their application to the large deviation theory. Specifically, we will explore the following questions:

- What are the existing limitations and gaps in Hamilton-Jacobi theory, and how does this thesis contribute to filling these gaps?
- Are there specific examples or cases that existing approaches do not adequately address? If so, how does this thesis provide a more comprehensive framework to cover them?
- How does the connection between Hamilton-Jacobi equations and *Large deviations* theory lead to new insights, and in what ways does this link extend the scope of both theories?

We start our discussion with a typical basic example arising from control theory. The first section offers a basic and informal introduction to the topic and may be skipped by expert readers already familiar with the subject.

1.1 A TOY EXAMPLE FROM CONTROL THEORY

Consider a scenario in which one wishes to fly from Delft to Naples. The objective is to reach the destination in an optimal manner, minimising the total cost of the flight. This cost may be a function of several factors, including fuel consumption, time, and passenger comfort. It is necessary to control the flight path, speed and altitude in order to ensure that the objective is achieved. In the field of control theory, this scenario can be framed as an optimal control problem, where the objective is to minimise a cost function.

In this context, the term $y(s)$ denotes the position of the aircraft at a given time point s, whereas $\dot{y}(s)$ represents its velocity. We now introduce the control variable, denoted by $\theta : [0, \infty) \to \Theta$, which maps from the interval $[0, \infty)$ to the set of parameters that influence the evolution of the system and are under our control. These parameters include factors such as engine power, steering, and acceleration, which together determine the aircraft's motion. The set of all control variables is denoted by $Ctrl$.

The motion of the aircraft is then described by the following ordinary differential equation (ODE):

$$\begin{cases} \dot{y}(s) = f(y(s), \theta(s)) & t < s < T, \\ y(t) = x, \end{cases} \tag{1.1.1}$$

where $f : \mathbb{R}^d \times \Theta \to \mathbb{R}^d$ represents the effective speed of the system, determined by the chosen control inputs.

We then denote by $y_{x,\theta}(s)$ the solution of the ODE (1.1.1) for a particular choice of starting point $x \in \mathbb{R}^d$ and control $\theta \in Ctrl$.

In this context, the controlled evolution is associated with a cost functional C, which represents the cost of flying from Delft to Naples. This functional is typically expressed as follows:

$$C(x, t, \theta) := g(y_{x,\theta}(T)) + \int_t^T r(y_{x,\theta}(s), \theta(s))\, ds.$$

- $g(y(T))$ is the *terminal cost* that reflects how well the final state $y(T)$ meets the desired target (Naples),
- $r(y(s), \theta(s))$ is the *running cost* (e.g., fuel consumption),
- t is the starting time when the plane leaves,
- T is the final time when the plane reaches the final destination.

We then ask the following question:

Can we control the system such that we fly from $x = y_{x,\theta^*}(t)$ to $y = y_{x,\theta^*}(T)$ minimizing the cost?

or, mathematically,

can we find a θ^* such that the following function

$$v(x, t) = C(x, t, \theta^*) = \inf_{\theta \in Ctrl} C(x, t, \theta)$$

exists? and is this function unique?

The function $v(x, t)$ is called *value function* and plays an important role in this thesis. We will indeed prove that the value function is **in some sense** a solution to a partial differential equation (PDE) called Hamilton-Jacobi-Bellman (HJB) equation. In this way, answering to the question above is equivalent to prove the existence and uniqueness of a solution (we will explain later in which sense) to the related Hamilton-Jacobi-Bellman equation.

1.1.1 *The Dynamic Programming principle and the related Hamilton-Jacobi-Bellman equation*

An essential property of the value function that is significant in optimal control theory is the *Dynamic Programming principle.* (DPP)

In this section, we present an informal overview of the principle and demonstrate how it can be used to derive a Hamilton-Jacobi equation. For a precise statement, we refer the reader to further sections.

In general, DPP is based on the idea that to solve a complex problem one can break it down into simpler sub-problems and solve them sequentially. Let us once more consider the journey from Delft to Naples. We start from a point x in the space $\mathbb{R}^d$ at a time $t \geq 0$, and we consider a first small trip in the small time interval $[t, t + \Delta t]$. Select any control parameter, denoted by $\omega \in \Theta$, and apply the constant control $\theta(s) \equiv \omega$ over the time interval $[t, t+\Delta t]$. The system evolves to the point $y_{x,\omega}(t+\Delta t)$. At this point, we switch to the optimal control strategy for the remaining time interval, $[t+\Delta t, T]$. We now calculate the total cost of the trip.

Over the interval $[t, t + \Delta t]$, the system evolves according to the dynamics:

$$\dot{y}(s) = f(y(s), \omega), \quad y(t) = x.$$

The cost for this period is:

$$\int_t^{t+\Delta t} r(y_{x,\omega}(s), \omega)\, ds.$$

The cost from time $t + \Delta t$ to T is the minimum cost starting from $y_{x,\omega}(t + \Delta t)$, i.e., $v(y_{x,\omega}(t + \Delta t), t + \Delta t)$. Thus, the total cost is:

$$\int_t^{t+\Delta t} r(y_{x,\omega}(s), \omega)\, ds + v(y_{x,\omega}(t + \Delta t), t + \Delta t), \tag{1.1.2}$$

optimized over the interval $[t, t + \Delta t]$. The Dynamic Programming principle roughly states that the value function $v(x, t)$ is the minimum of (1.1.2). To be more precise,

$$v(x, t) = \inf_{\omega \in Ctrl} \left\{ \int_t^{t+\Delta t} r(y_{x,\omega}(s), \omega)\, ds + v(y_{x,\omega}(t + \Delta t), t + \Delta t) \right\}. \tag{DPP}$$

The idea behind this principle is that, at any point t, the optimal cost to go from state x is the sum of:

- The cost accumulated from t to an intermediate time $t + \Delta t$,
- The optimal cost-to-go from $t + \Delta t$ onward.

We use now (DPP) to obtain the inequality

$$v(x, t) \leq \int_t^{t+\Delta t} r(y_{x,\omega}(s), \omega)\, ds + v(y(t + \Delta t), t + \Delta t).$$

To convert this inequality into a differential form, we rearrange the terms and divide by $\Delta t > 0$

$$\frac{v(y_{x,\omega}(t + \Delta t), t + \Delta t) - v(x, t)}{\Delta t} + \frac{1}{\Delta t} \int_t^{t+\Delta t} r(y_{x,\omega}(s), \omega)\, ds \geq 0.$$

Taking the limit as $\Delta t \to 0$, we obtain

$$v_t(x, t) + \nabla_x v(y_{x,\omega}(t), t) \cdot \dot{y}_{x,\omega}(t) + r(y_{x,\omega}(t), \omega) \geq 0. \tag{1.1.3}$$

Since $y_{x,\omega}(\cdot)$ satisfies the ODE

$$\dot{y}(s) = f(y(s), \omega), \quad t \le s \le t + \Delta t, \quad y(t) = x,$$

we substitute this into (1.1.3), leading to

$$v_t(x,t) + f(x,\omega) \cdot \nabla_x v(x,t) + r(x,\omega) \ge 0.$$

This inequality holds for any control parameter $\omega \in \Theta$, so we can write:

$$-v_t(x,t) + \sup_{\theta \in \Theta} \{-f(x,\theta) \cdot \nabla_x v(x,t) - r(x,\theta)\} \le 0.$$

We now prove that the expression above actually equals zero. Suppose that there exists a control $\omega^*(\cdot)$ optimizing (DPP), that is,

$$\int_t^{t+\Delta h} r(y_{x,\omega^*}(s), \omega^*(s))\, ds + v(y_{x,\omega^*}(t+\Delta t), t+\Delta t) = v(x,t).$$

Rearranging and dividing by Δt we obtain

$$\frac{v(y_{x,\omega^*}(t+\Delta t), t+\Delta t) - v(x,t)}{\Delta t} + \frac{1}{\Delta} \int_t^{t+\Delta t} r(y_{x,\omega^*}(s), \omega^*(s))\, ds = 0.$$

Taking the limit as $\Delta t \to 0$, and assuming $\omega^*(t) = \omega^* \in \Theta$, we have:

$$v_t(x,t) + \nabla_x v(x,t) \cdot \dot{y}_{x,\omega^*}(t) + r(x,\omega^*) = 0,$$

or equivalently,

$$-v_t(x,t) - f(x,\omega^*) \cdot \nabla_x v(x,t) - r(x,\omega^*) = 0,$$

for some $\omega^* \in \Theta$. This means that $v(x,t)$ is the solution of the Hamilton-Jacobi equation

$$-u_t(x,t) + Hu(\cdot,t)(x) = 0, \tag{1.1.4}$$

with the Hamiltonian H given by

$$Hu = \sup_{\theta \in \Theta} \{-f(x,\theta) \cdot \nabla_x u(x,t) - r(x,\theta)\}.$$

When the Hamiltonian has a representation as the supremum of a function, the equation is called Hamilton-Jacobi-Bellman equation.

The derivation of the Hamilton-Jacobi-Bellman equation that we have showed above relies on several assumptions. In particular, the proof sketched above is based on the assumption of the existence of optimal controls and the regularity of the value function, specifically its differentiability. Nevertheless, in numerous real-world problems, such assumptions are overly restrictive.

For instance, it is possible that the value function will not always be smooth, and that its differentiability may be compromised. Furthermore, it is not always the case that optimal controls can be identified. These challenges demonstrate that the classical approach, while valuable, is not a comprehensive solution for all control problems. It is therefore necessary to develop a more general framework that can accommodate non-smoothness and irregularities.

In order to overcome such limitations, we will use the concept of *viscosity solution*. Viscosity solutions represent a weak solution concept for Hamilton-Jacobi equations, offering particular utility in instances where the value function exhibits a lack of regularity. This approach does not necessitate the differentiability of the solution; instead, it focuses on the definition of the solution through the use of subsolutions and supersolutions.

Viscosity solutions will provide the necessary generalization to handle cases where classical solutions are inapplicable, ensuring the existence and uniqueness of solutions even in more complex scenarios.

1.1.2 *Viscosity solutions: a conceptual overview*

In this section, we provide a heuristic introduction to the concept of viscosity solutions and show that the value function is indeed a viscosity solution of the Hamilton-Jacobi-Bellman equation (1.1.4).

The concept of viscosity solutions, introduced by Crandall and Lions in the 80s [CL83], provides a powerful framework for analyzing Hamilton-Jacobi equations, for example in optimal control problems where the value function may lack smoothness. Viscosity solutions are defined through comparisons with test functions, allowing us to make sense of the solution even when it is not differentiable.

Consider the Hamilton-Jacobi equation

$$u(x) - Hu(x) = 0. \tag{1.1.5}$$

The concept of viscosity solutions is built upon the *maximum principle*. The latter essentially states that if we take two functions, u and f, in the domain of an operator H that satisfies the maximum principle and we find a point x_0 such that

$$(u - f)(x_0) = \sup_x (u - f)(x),$$

then,

$$Hu(x_0) \leq Hf(x_0).$$

Heuristically, if H satisfies the maximum principle, u is a classical solution of (1.1.5), ϕ is a test function in the domain of H and x_0 is a point such that

$$u(x_0) - \phi(x_0) = \sup_x (u - \phi)(x),$$

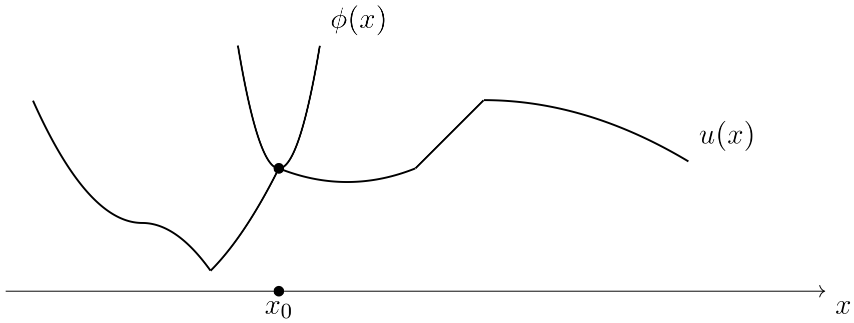

Figure 1: An illustration of ϕ touching from above u in x_0.

by the maximum principle we can then write

$$\begin{aligned} 0 &= u(x_0) - Hu(x_0) \\ &\geq u(x_0) - H\phi(x_0). \end{aligned}$$

This justifies the following informal definition. We will give the precise definition in Section 1.2.1.

Assume that at some x_0, a continuous function u can be "touched from above" by some smooth function ϕ at x_0. By this, we mean that the difference $u(x) - \phi(x)$ has a vanishing (local) maximum in a neighborhood of x_0 (see Figure 1).

Then, u is called a *viscosity subsolution* of the Hamilton-Jacobi equation (1.1.5) if, for every smooth test function ϕ that touches u from above at a point x_0, we have:

$$u(x_0) - H\phi(x_0) \leq 0.$$

Similarly, v is a viscosity supersolution if the inequality holds in the opposite direction when ϕ touches v from below. A viscosity solution is a function that is both a subsolution and a supersolution.

This definition allows us to extend the concept of solutions to a much broader class of functions, particularly in cases where classical solutions do not exist or are difficult to compute.

1.1.3 *The value function as a viscosity solution*

To conclude the discussion started in Section 1.1, we now provide an informal sketch of the proof that the value function $v(x,t)$, which represents the minimal cost of the control problem, is a viscosity solution to the equation (1.1.4). Here we will assume that the value function is bounded and continuous (but not that it is differentiable).

For a time-dependent Hamilton-Jacobi equation, as the one in (1.1.4), the subsolution property is translated in the following way:

If $v(x,t) - \phi(x,t)$ has a local maximum in (x_0, t_0), then

$$\phi_t(x_0, t_0) - H\phi(\cdot, t_0) \le 0.$$

Assume that (x_0, t_0) as above exists. To prove that v is a subsolution to (1.1.4) we need to show that

$$-\phi_t(x_0, t_0) + \sup_{\theta \in \Theta} \{-f(x_0, \theta) \cdot \nabla_x \phi(x_0, t_0) - r(x_0, \theta)\} \le 0.$$

Assume by contradiction that this is not true, i.e., assume that there exists a parameter $\theta \in \Theta$ and $a > 0$ such that

$$-\phi_t(x,t) - f(x,\theta) \cdot \nabla_x \phi(x,t) - r(x,\theta) \ge a > 0, \tag{1.1.6}$$

for all points (x,t) close enough to (x_0, t_0). Moreover, since (x_0, t_0) is a local maximum of $u - \phi$, we can also write

$$v(x,t) - \phi(x,t) \le v(x_0, t_0) - \phi(x_0, t_0), \tag{1.1.7}$$

for all points (x,t) close to (x_0, t_0).

Consider now the constant control variable $\theta(t) \equiv \theta$ for $t_0 \le s \le T$. If we choose Δt small such that the dynamics $y_{x_0,\theta}(s)$ is close to the starting point x_0 in the interval $[t_0, t_0 + \Delta t]$, we can write

$$-\phi_t(y_{x_0,\theta}(s), s) - f(y_{x_0,\theta}(s), \theta) \cdot \nabla_x \phi(y_{x_0,\theta}(s), s) - r(y_{x_0,\theta}(s), \theta) > 0.$$

Using (1.1.7), we find

$$\begin{aligned} v(y_{x_0,\theta}(t_0 + \Delta t), t_0 + \Delta t) - v(x_0, t_0) &\le \phi(y_{x_0,\theta}(t_0 + \Delta t), t_0 + \Delta t) - \phi(x_0, t_0) \\ &= \int_{t_0}^{t_0+\Delta t} \frac{d}{dt} \phi(y_{x_0,\theta}(s), s)\, ds \\ &= \int_{t_0}^{t_0+\Delta t} \phi_t(y_{x_0,\theta}(s), s) + \nabla_x \phi(y_{x_0,\theta}(s), s) \cdot \dot{y}_{x_0,\theta}(s)\, ds \\ &= \int_{t_0}^{t_0+\Delta t} \phi_t(y_{x_0,\theta}(s), s) + f(y_{x_0,\theta}(s), \theta) \cdot \nabla_x \phi(y_{x_0,\theta}(s), s)\, ds. \end{aligned} \tag{1.1.8}$$

In addition, (DPP) provides the inequality

$$v(x_0, t_0) \le \int_{t_0}^{t_0+\Delta t} r(y_{x_0,\theta}(s), \theta)\, ds + v(y_{x_0,\theta}(t_0 + \Delta t), t_0 + \Delta t). \tag{1.1.9}$$

Combining (1.1.8) and (1.1.9), we find

$$\int_{t_0}^{t_0+\Delta t} -\phi_t(y_{x_0,\theta}(s), s) - f(y_{x_0,\theta}(s), \theta) \cdot \nabla_x \phi(y_{x_0,\theta}(s), s) - r(y_{x_0,\theta}(s), \theta)\, ds \le 0,$$

that contradicts (1.1.6). This prove that v is a subsolution to the equation

$$-u_t + \sup_{\theta\in\Theta} \{-f(x_0,\theta)\nabla_x \phi(x_0,t_0) - r(x_0,\theta)\} = 0.$$

The proof that v is also a supersolution follows similarly.

As illustrated above, the concept of viscosity solutions permits one to consider the case of a non-differentiable value function. It is important to note, however, that the example provided above in the context of control theory is just one of numerous instances where Hamilton-Jacobi equations are encountered. Indeed, these equations can be observed in a variety of contexts, including stochastic control, financial mathematics, and game theory. In these scenarios, the Hamiltonian can exhibit significantly greater complexity than that considered previously.

We can then ask the following question:

Can we prove the existence and uniqueness of viscosity solutions for a general Hamilton-Jacobi equation? Under which condition on the Hamiltonian and the value function?

Furthermore, it is worthwhile to consider the journey from Delft to Naples once more. The flight path is not merely a straight line in Euclidean space; rather, it is subject to the influence of the Earth's curvature. It is therefore reasonable to ask whether it is possible to set up a framework on a general Riemannian manifold. This leads us to a broader investigation into the properties of Hamilton-Jacobi equations in more complex settings. When dealing with irregular Hamiltonians or setting the problem on a general Riemannian manifold, it is necessary to adapt the mathematical framework in order to account for these challenges. This thesis aims to explore these extensions by providing insights into the conditions needed for the existence and uniqueness of viscosity solutions in more general cases.

1.2 INTRODUCTION TO HAMILTON-JACOBI EQUATIONS: CONCEPTS AND CONTEXT

In this thesis we study three type of Hamilton-Jacobi equations:

First-order elliptic equations:

$$f(x) - \lambda\mathcal{H}(x,\nabla f(x)) = h(x) \qquad \lambda > 0; \tag{1.2.1}$$

First-order parabolic equations:

$$\partial_t f(x,t) + \lambda f(x,t) - \mathcal{H}(x,\nabla_x f(x,t)) = 0 \qquad \lambda \geq 0; \tag{1.2.2}$$

and **Second-order elliptic equations:**

$$f(x) - \lambda\mathcal{H}(x, f(x), \nabla f(x), \nabla^2 f(x)) = h(x) \qquad \lambda > 0. \tag{1.2.3}$$

We will specify later the type of Hamiltonians that we consider and their domains and the sets in which we consider the above equations. The aim of this work is to study well-posedness for *viscosity solutions* of equations of the type as in (1.2.1), (1.2.2) and (1.2.3).

It is then natural to start introducing the notions of *viscosity solutions* and *comparison principle*. This is done in the following Section 1.2.1. Later, in Sections 1.2.2 and 1.2.4 we will discuss in details first and second-order Hamilton-Jacobi equations respectively.

1.2.1 *Viscosity solutions and comparison principle*

Consider the following boundary value problem:

$$|u'(x)| = 1 \quad \text{in} \quad (-1,1), \quad u(-1) = u(1) = 0.$$

The above problem is known as the *"one dimensional Eikonal problem"*. Three observations can be done (see also Figure 2):

1. There is no classical solution, i.e., a C^1- function, to the above problem. This is not hard to show and follows by Rolle's theorem.
2. There are infinitely many weak solutions, i.e. continuous functions which satisfy the equation at almost every point (Figure 2 shows three of them).
3. We need a criterion that allows us to choose a single solution in order to ensure the well-posedness of the problem. For example, if u is a solution almost everywhere, then $-u$ is also a solution. Therefore, to select a unique solution, we need to avoid the symmetry, for instance.

This is not an atypical situation. Consequently, it is important to develop a theory that permits merely continuous functions to serve as solutions to the Hamilton-Jacobi equation, while also offering a method to identify the appropriate solution from among the weak solutions to the problem.

The theory of *viscosity solutions* was introduced by M.G. Crandall and P.L. Lions in the 80s in [CL83]. The notion of a viscosity solution is derived from the concepts of *subsolutions* and *supersolutions*. The idea is to "replace" the differential $Du(x)$ at a point x where it does not exist with the differential $D\phi(x)$ of a smooth function ϕ touching the graph of u, from above for the subsolution condition and from below for the supersolution one, at the point x. We give below the definition. For simplicity, we first consider the case where $E \subset \mathbb{R}^d$ is a compact set and the elliptic case. Let $H \subseteq C(E) \times C(E)$ (we identify H with its graph), $\lambda > 0$ and $h \in C_b(E)$ and consider the following Hamilton-Jacobi equation

$$f(x) - \lambda H f(x) = h(x). \tag{1.2.4}$$

Definition 1.2.1. A bounded, upper semi-continuous function u is a *viscosity subsolution* of (1.2.4) if for every $\phi \in C^1(E)$ and $x_0 \in E$ such that $u(x_0) - \phi(x_0) = \sup_x u(x) - \phi(x)$, we have

$$u(x_0) - \lambda H\phi(x_0) \leq h(x_0).$$

A bounded,lower semi-continuous function v is a *viscosity supersolution* of (1.2.4) if for every $\phi \in C^1(E)$ and $x_0 \in E$ such that $u(x_0) - \phi(x_0) = \inf_x u(x) - \phi(x)$, we have

$$u(x_0) - \lambda H\phi(x_0) \geq h(x_0).$$

A bounded, continuous function u is a *viscosity solution* of (1.2.4) if it is both a subsolution and a supersolution.

An similar definition for the time-dependent version of equation (1.2.4) can be written similarly. Later, at the end of this section, we will present an equivalent definition of viscosity solution involving generalized differentials, which we postponed for now to maintain the clarity of the discussion.

Remark 1.2.2. It is evident that any solution in the classical sense is also a viscosity solution. Conversely, if a viscosity solution u is differentiable at a point x, it satisfies the equation in the classical sense at that point. Thus, the concept of a viscosity solution includes that of a classical solution.

We now revisit the Eikonal problem and show that only $u_1(x) = 1 - |x|$ (in red in Fig. 2) is a viscosity solution. The notion of viscosity solutions thus proves to be an effective framework for selecting a unique solution among many weak solutions. We only study what happens in the origin, since this is the critical point of this equation. Let $\phi \in C^1(-1,1)$ be a test function such that

$$u_1(0) - \phi(0) = \sup_x \big(u_1(x) - \phi(x)\big).$$

In this case, $\phi'(0) \in [-1,1]$, which implies that u_1 satisfies the conditions of a viscosity subsolution. Furthermore, u_1 is also a viscosity supersolution because no function $\phi \in C^1(-1,1)$ exists such that

$$u_1(0) - \phi(0) = \inf_x \big(u_1(x) - \phi(x)\big).$$

For weak solutions with a downward edge, such as the blue and green solutions in Fig. 2, these functions fail to qualify as supersolutions. Specifically, at the point x where the edge occurs, for any $\phi \in C^1(-1,1)$ such that

$$u(x) - \phi(x) = \inf_x \big(u(x) - \phi(x)\big),$$

we find that $\phi'(x) \in (-1,1)$, which violates the supersolution condition.

Thus, we conclude that u_1 is the unique viscosity solution.

The uniqueness of viscosity solutions is often given in terms of the *comparison principle*.

Definition 1.2.3. We say that a Hamilton-Jacobi equation satisfies the *comparison principle* if for a subsolution u and a supersolution v we have $u \le v$.

Remark 1.2.4. The comparison principle implies the uniqueness of viscosity solutions. Indeed, suppose that u and v are two solutions of the same equation satisfying the comparison principle. Then, by using that u is a subsolution and v is a supersolution, we get $u \le v$ from the comparison principle. By interchanging the role of u and v, we get the opposite inequality.

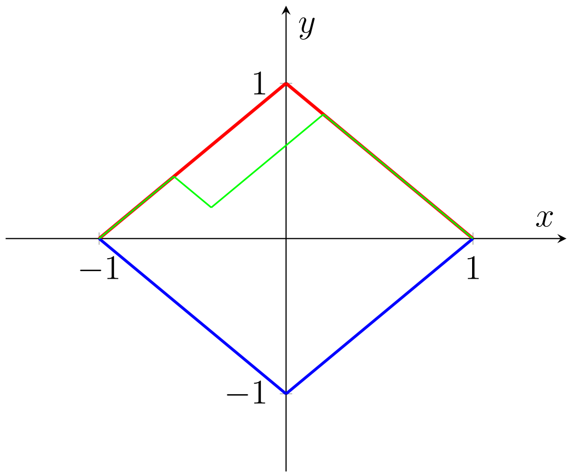

Figure 2: Visualization of the Eikonal problem. The red curve corresponds to the unique viscosity solution $u_1(x) = 1 - |x|$, while the blue and green curves represent weak solutions with downward edges.

When the set E is not compact, we need to introduce another definition of viscosity sub-supersolution. This is because in a non-compact setting the point $x_0 \in E$ such that $u(x_0) - \phi(x_0) = \sup_x u(x) - \phi(x)$ (or $u(x_0) - \phi(x_0) = \inf_x u(x) - \phi(x)$) might not exist. We solve this issue by considering sequences of optimizers.

Definition 1.2.5. A bounded, upper semi-continuous function $u\colon E \to \mathbb{R}$ is called a *(viscosity) subsolution* to (1.2.4) if, for all $(f, g) \in H$, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq E$ such that

$$\lim_{n\to\infty} u(x_n) - f(x_n) = \sup_{x\in E} u(x) - f(x),$$
$$\limsup_{n\to\infty} u(x_n) - \lambda g(x_n) - h(x_n) \leq 0.$$

A bounded, lower semicontinuous function $v\colon E \to \mathbb{R}$ is called a *(viscosity) supersolution* to (1.2.4) if, for all $(f, g) \in H$, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subseteq E$ such that

$$\lim_{n\to\infty} v(x_n) - f(x_n) = \inf_{x\in E} v(x) - f(x),$$
$$\liminf_{n\to\infty} v(x_n) - \lambda g(x_n) - h_2(x_n) \geq 0.$$

In the case where the test functions $f \in \mathcal{D}(H)$ have compact sub-superlevel sets, the above definition can be simplified by considering the limit point x_0 of the sequences x_n. This is shown in the following lemma.

Lemma 1.2.6. *Consider the Hamilton-Jacobi equation* (1.2.4).

(a) *Assume that every $f \in \mathcal{D}(H)$ has compact sublevel sets, and let $u\colon E \to \mathbb{R}$ be bounded and upper semicontinuous. Then, u a viscosity subsolution to (1.2.4) if and only if, for all $(f,g) \in H$, there exists some $x_0 \in E$ with*

$$\begin{aligned} u(x_0) - f(x_0) &= \sup_{x\in E} u(x) - f(x), \\ u(x_0) - \lambda g(x_0) &\le h(x_0). \end{aligned}$$

(b) *Assume that every $f \in \mathcal{D}(H)$ has compact superlevel sets, and let $v\colon E \to \mathbb{R}$ be bounded and lower semicontinuous. Then, v is a viscosity supersolution to (1.2.4) if and only if, for all $(f,g) \in H$, there exists some $x_0 \in E$ with*

$$\begin{aligned} v(x_0) - f(x_0) &= \inf_{x\in E} v(x) - f(x), \\ v(x_0) - \lambda g(x_0) &\ge h(x_0). \end{aligned}$$

Proof. We only show Part (a). Part (b) follows analogously.

First, assume that u is a viscosity subsolution to (1.2.4) and let $(f,g) \in H$. Then, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset E$ such that

$$\begin{aligned} &\lim_{n\uparrow\infty} u(x_n) - f(x_n) = \sup u(x) - f(x) =: C, \\ &\limsup_{n\uparrow\infty} u(x_n) - \lambda g(x_n) - h(x_n) \le 0. \end{aligned}$$

In particular, there exists some $n_0 \in \mathbb{N}$ with

$$u(x_n) - f(x_n) \ge C - 1 \quad \text{for all } n \in \mathbb{N} \text{ with } n \ge n_0.$$

Since u is bounded, it follows that

$$f(x_n) \le 1 - C + \|u\|_\infty \quad \text{for all } n \in \mathbb{N} \text{ with } n \ge n_0.$$

As f has compact sublevel sets, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $x_{n_k} \to x_0 \in E$. Since $u(x_n) - f(x_n) \to C$ as $n \to \infty$, it follows that

$$C = \lim_{k\to\infty} u(x_{n_k}) - f(x_{n_k}) \le u(x_0) - f(x_0),$$

where in the last step, we used that u is upper semicontinuous and f is continuous. Therefore, we have shown that

$$u(x_0) - f(x_0) = C = \lim_{k\to\infty} u(x_{n_k}) - f(x_{n_k}),$$

which, since f is continuous, implies that

$$u(x_0) = \lim_{k\to\infty} u(x_{n_k}).$$

Now, since g and h are continuous, we find that

$$\begin{aligned} u(x_0) - \lambda g(x_0) - h(x_0) &= \lim_{k\to\infty} u(x_{n_k}) - \lambda g(x_{n_k}) - h(x_{n_k}) \\ &\le \limsup_{n\uparrow\infty} u(x_n) - \lambda g(x_n) - h(x_n) \le 0. \end{aligned}$$

For the other implication, one chooses $x_n = x_0$ for all $n \in \mathbb{N}$. □

The following discussion, which stems from the above lemma, will play a key role in our strategy and will appear frequently throughout this thesis.

Suppose we want to prove the comparison principle for

$$f(x) - \lambda H f(x) = h(x), \tag{1.2.5}$$

with $\mathcal{D}(H) = C_b(E)$, $\lambda > 0$ and $h \in C_b(E)$. We then need to prove that for every subsolution u and supersolution v to (1.2.5), $u \leq v$. Of course, we want to use the unique information that we have about u and v, that is, u is a subsolution and v is a supersolution. We hence need to consider test functions. Note that the test functions are in the domain of the Hamiltonian H and keep in mind Lemma 1.2.6. We now perform a trick:

Suppose we can find two Hamiltonians H_1, H_2 with domains $\mathcal{D}(H_1)$ and $\mathcal{D}(H_2)$ respectively, and such that

1. Every viscosity subsolution of (1.2.5) is also a viscosity subsolution of $f(x) - \lambda H_1 f(x) = h(x)$;
2. Every viscosity supersolution of (1.2.5) is also a viscosity supersolution of $f(x) - \lambda H_2 f(x) = h(x)$.

If we then prove that for every subsolution u of $f - \lambda H_1 f = h$ and every supersolution v of $f - \lambda H_2 f = h$,

$$\sup_x u - v \leq 0,$$

we also automatically proved the comparison principle for $f - \lambda H f = h$. The advantage here is that we can choose the domain of H_1, H_2 (so we can choose the test functions) as soon as we prove 1 and 2.

For this reason, we will often work with a pair of Hamilton-Jacobi equations

$$f - \lambda H_1 f = h_1, \tag{1.2.6}$$

$$f - \lambda H_2 f = h_2, \tag{1.2.7}$$

with $\lambda > 0$, $h_1 \in C_l(E)$ and $h_2 \in C_u(E)$ and $H_1 \subseteq C_l(E) \times C(E)$ and $H_2 \subseteq C_u(E) \times C(E)$.

In this scenario, we consider the following definitions.

Definition 1.2.7.

- We say that a bounded, upper semi-continuous function u is a viscosity subsolution of the system (1.2.6)-(1.2.7) if it is a viscosity subsolution of (1.2.6) as in Definition 1.2.5.
- We say that a bounded, lower semi-continuous function v is a viscosity supersolution of the system (1.2.6)-(1.2.7) if it is a viscosity supersolution of (1.2.7) as in Definition 1.2.5.
- We say that the system (1.2.6)-(1.2.7) satisfies the comparison principle if for all u subsolution of (1.2.6) and v supersolution of (1.2.7), $\sup_x u - v \leq \sup h_1 - h_2$.

Essentially, we will consider the scenario of Figure 3.

We conclude this section with a different definition of viscosity solutions that involves generalized differentials. Even if we will always use the definition via test function, this

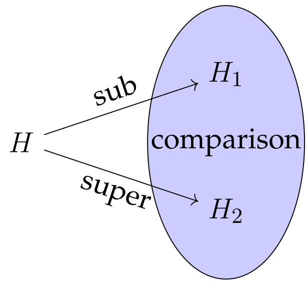

Figure 3: In this diagram, an arrow connecting an operator A with operator B with subscript 'sub' means that viscosity subsolutions of $f - \lambda A f = h$ (or $\partial_t f - Af = 0$) are also viscosity subsolutions of $f - \lambda B f = h$ (or $\partial_t f - Bf = 0$). Similarly for arrows with a subscript 'super'.

version will be part of our discussion about second-order Hamilton-Jacobi equations in Section 1.2.4. We start with the definition of a notion of derivative for non-smooth functions that generalize the classical notion of derivative by using the Taylor expansion at a point.

Definition 1.2.8 (Generalized Derivatives)**.** Let $S(d)$ be the set of $d \times d$ symmetric matrices. For $u \in USC(E)$ and $x \in E$, we define the set

$$J^{2,+}u(x) := \left\{ (p,X) \in \mathbb{R}^n \times S(d) \;\middle|\; \begin{array}{c} u(z+x) \leq u(x) + p \cdot z + \frac{1}{2}\langle X \cdot z, z\rangle \\ + o(|z|^2) \text{ as } z \to 0 \end{array} \right\},$$

and for $v \in LSC(E)$ and $x \in E$

$$J^{2,-}v(x) := \left\{ (p,X) \in \mathbb{R}^n \times S(d) \;\middle|\; \begin{array}{c} v(z+x) \geq v(x) + p \cdot z + \frac{1}{2}\langle X \cdot z, z\rangle \\ + o(|z|^2) \text{ as } z \to 0 \end{array} \right\}.$$

We call $J^{2,+}u(x)$ the *second order Super-Jet* of u at x and $J^{2,-}v(x)$ the *second order Sub-Jet* of v at x.

Consider again equation (1.2.4). Recall that the Hamiltonian H can be a function of $x, f(x)$ and its first and second derivative (we only consider first and second-order equations). To better illustrate the upcoming concept, let us consider an operator $\mathbf{H}$ such that $Hf(x) = \mathbf{H}(x, f(x), Df(x), D^2 f(x))$.

Definition 1.2.9 (Viscosity solutions via generalized derivatives)**.** A bounded, upper semi-continuous function u is a viscosity subsolution of (1.2.4) if

$$u(x) - \lambda \mathbf{H}(x, u(x), p, X) \leq h(x) \qquad \text{for all } x \in E \text{ and for all } (p,X) \in J^{2,+}u(x).$$

A bounded, lower semi-continuous function v is a viscosity supersolution of (1.2.4) if

$$v(x) - \lambda \mathbf{H}(x, v(x), p, X) \geq h(x) \qquad \text{for all } x \in E \text{ and for all } (p,X) \in J^{2,-}v(x).$$

As we said above, the notion of viscosity solutions via generalized gradient will be part of our discussion in Section 1.2.4 but not of our strategy in the next chapters that will be, instead, only based on the definition via test functions for both first and second-order equations. For this reason, we do not give here the proof of the equivalence between the two notions of viscosity solutions. This equivalence is not difficult to prove and it is based on the fact that the super-jet and the sub-jet can be rewritten as

$$J^{2,+}u(x) = \left\{(D\phi(x), D^2\phi(x)) \,|\, \phi \in C^2(E), u - \phi \,\text{has a local maximum at}\, x\right\};$$

$$J^{2,-}u(x) = \left\{(D\phi(x), D^2\phi(x)) \,|\, \phi \in C^2(E), u - \phi \,\text{has a local minimum at}\, x\right\}.$$

1.2.2 *First-order Hamilton-Jacobi equations*

In this section, we introduce the first-order Hamilton-Jacobi (and Hamilton-Jacobi-Bellman and Isaacs) equations. Throughout the section we will consider equations of the type

$$u(x) - \lambda\mathcal{H}(x, \nabla u(x)) = h(x), \tag{1.2.8}$$

where λ is a positive constant and h is a continuous bounded function, and its time–dependent version

$$\begin{cases} \partial_t u(x,t) - \mathcal{H}(x, \nabla_x u(t,x)) = 0, & \text{if } t > 0, \\ u(0,x) = u_0(x) & \text{if } t = 0. \end{cases}$$

We set our equations on an open set $E \subseteq \mathbb{R}^d$.

As mentioned in Section 1.2.1, the theory of viscosity solutions for equations of the above form was introduced in the 1980s by M.G. Crandall and P.L. Lions in their works [CL81; CL83] continued with L.C. Evans in [CEL84]. Since then, the question of the well-posedness of Hamilton-Jacobi equations in the viscosity sense has been a topic of extensive investigation. Over time, the research has been conducted with the clear objective of expanding the theory to encompass increasingly general classes.

The following section provides an overview of the typical strategy encountered in the literature when proving the comparison principle for first-order Hamilton-Jacobi equations. Later, in the beginning of Chapter 2, we briefly discuss our main motivation to improve this strategy.

1.2.3 *Typical strategy for proving the comparison principle*

Consider the static equation (1.2.8). Let u be a subsolution and v a supersolution of (1.2.8). The aim is to prove the comparison principle, i.e.,

$$\sup_x u(x) - v(x) \leq 0.$$

The typical strategy is based on the *"doubling variables procedure"*. The trick consists in considering the following auxiliary function

$$\Phi_\varepsilon(x,y) = u(x) - v(y) - \frac{|x-y|^2}{\varepsilon^2} - \varepsilon\big(|x|^2 + |y|^2\big).$$

Since u and v are bounded and USC and LSC respectively, Φ_ε is bounded from above, upper semicontinuous and tends to $-\infty$ if $|x| \to \infty$ or $|y| \to \infty$. Hence, it has a global maximum $(x_\varepsilon, y_\varepsilon)$. As the doubling variables procedure is a long and technical method, we break the discussion into simple steps:

Step 1: We want to prove that x_ε and y_ε converge for $\varepsilon \to 0$ to the same point. To this aim, note that $\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \geq \Phi_\varepsilon(0,0)$, thus

$$u(x_\varepsilon) - v(y_\varepsilon) \geq u(0) - v(0) + \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} + \varepsilon\big(|x_\varepsilon|^2 + |y_\varepsilon|^2\big).$$

We then obtain that

$$2(\|u\|_{L^\infty(\mathbb{R}^n)} + \|v\|_{L^\infty(\mathbb{R}^n)}) \geq \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} + \varepsilon\big(|x_\varepsilon|^2 + |y_\varepsilon|^2\big).$$

This implies that $(x_\varepsilon - y_\varepsilon) \to 0$ as $\varepsilon \to 0$, and $|x_\varepsilon| + |y_\varepsilon| \leq \frac{C}{\sqrt{\varepsilon}}$.

Step 2: We prove that

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} \to 0 \quad \text{as } \varepsilon \to 0.$$

This follows observing that $\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \geq \Phi_\varepsilon(x_\varepsilon, x_\varepsilon)$ which implies

$$\begin{aligned}\frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2} &\leq v(x_\varepsilon) - v(y_\varepsilon) + \varepsilon(|x_\varepsilon|^2 - |y_\varepsilon|^2) \\ &\leq v(x_\varepsilon) - v(y_\varepsilon) + C\varepsilon^{3/2},\end{aligned}$$

and using that v is lower semicontinuous.

Step 3: We want to use now the subsolution and supersolution properties. For that, we need to find the right test functions. We know that $x \mapsto \Phi_\varepsilon(x, y_\varepsilon)$ has a maximum at x_ε, which means

$$x \mapsto u(x) - \left(\frac{|x - y_\varepsilon|^2}{\varepsilon^2} + \varepsilon|x|^2 + (x - x_\varepsilon)^2\right)$$

has a *unique* (because of the term $(x - x_\varepsilon)^2$) maximum at x_ε. As u is a viscosity subsolution of (1.2.8), using $\phi_1(x) = \frac{|x-y_\varepsilon|^2}{\varepsilon^2} + \varepsilon|x|^2 + (x - x_\varepsilon)^2$ as test function, we have

$$u(x_\varepsilon) - \lambda H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) \leq h(x_\varepsilon).$$

Similarly, $y \mapsto \Phi_\varepsilon(x_\varepsilon, y)$ has a maximum at y_ε, which yields

$$y \mapsto v(y) - \left(-\frac{|x_\varepsilon - y|^2}{\varepsilon^2} - \varepsilon|y|^2 - (y - y_\varepsilon)^2 \right)$$

has a *unique* (because of the term $-(y - y_\varepsilon)^2$) minimum at y_ε. Since v is a viscosity supersolution of (1.2.8), using the test function $\phi_2(y) = -\frac{|x_\varepsilon - y|^2}{\varepsilon^2} - \varepsilon|y|^2 - (y - y_\varepsilon)^2$, we obtain

$$v(y_\varepsilon) - \lambda H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) \geq h(y_\varepsilon).$$

Step 4: This step shows the reason why this method is called the "doubling variables procedure". The title is indeed justified by the first of the following series of equalities and inequalities:

$$\begin{aligned}
\sup_x (u - v) &\leq \sup_{x,y} u(x) - v(y) - \frac{|x-y|^2}{\varepsilon^2} \\
&\leq \liminf_{\varepsilon \to 0} \sup_{x,y} u(x) - v(y) - \frac{|x-y|^2}{\varepsilon^2} - \varepsilon\big(|x|^2 + |y|^2\big) \\
&= \liminf_{\varepsilon \to 0} u(x_\varepsilon) - v(y_\varepsilon) \\
&\leq \liminf_{\varepsilon \to 0} h(x_\varepsilon) - h(y_\varepsilon) \\
&\qquad + \lambda \left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \\
&\leq \liminf_{\varepsilon \to 0} \lambda \left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right].
\end{aligned}$$

Step 5: Suppose that

$$\liminf_{\varepsilon \to 0} \left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \leq 0,$$

then, the comparison principle follows.

We used the steps above to move the goal from trying to bound the difference between a subsolution and a supersolution

$$\sup_x u(x) - v(x)$$

to trying to bound the difference

$$\liminf_{\varepsilon \to 0} \left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \leq 0. \tag{1.2.9}$$

This is typically achieved by leveraging the regularity properties of the Hamiltonian. Over the past 40 years, the primary goal has been to extend this approach to increasingly broader classes of Hamiltonians. In the following, we give the typical assumptions on the Hamiltonian that have been used in the literature to prove (1.2.9).

Theorem 1.2.10. *Consider the Hamilton-Jacobi equation* (1.2.8). *Suppose the Hamiltonian is continuous and it satisfies one of the following properties:*

(I) $H(x,p) = H(x)$ *for every* $x \in E, p \in \mathbb{R}^d$.

(II) $p \mapsto H(x,p)$ *is uniformly coercive, that is,*

$$\sup_{x\in K} \lim_{|p|\to\infty} H(x,p) = \infty \quad \textit{for every compact } K.$$

(III) There exists a constant $C > 0$ *such that*

$$|H(x,p) - H(y,p)| \le C(1+|p|)|x-y|,$$
$$|H(x,p) - H(x,q)| \le C|p-q|.$$

Then,

$$\liminf_{\varepsilon\to 0}\left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \le 0. \tag{1.2.10}$$

Proof. 1. Assume (I). This means that the limit in (1.2.10) is equal to

$$\liminf_{\varepsilon\to 0} H(x_\varepsilon) - H(y_\varepsilon)$$

that vanishes by the continuity of H and the fact that x_ε and y_ε converge to the same point (by Step 1 above).

2. Assume (II). By the supersolution inequality for v in Step 3, we get

$$H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) \le \frac{1}{\lambda}\sup_x (v-h) < \infty.$$

Then, by coercivity, the sequence $p_\varepsilon = \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}$ is bounded, allowing us to extract a converging subsequence p_{ε_k}. Then, again by continuity of H and Step 1, we can conclude that

$$\begin{aligned}
&\liminf_{\varepsilon\to 0}\left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \\
&\le \liminf_{k\to\infty}\left[H\left(x_{\varepsilon_k}, \frac{2(x_{\varepsilon_k} - y_{\varepsilon_k})}{{\varepsilon_k}^2} + 2\varepsilon_k x_{\varepsilon_k}\right) - H\left(y_{\varepsilon_k}, \frac{2(x_{\varepsilon_k} - y_{\varepsilon_k})}{{\varepsilon_k}^2} - 2\varepsilon_k y_{\varepsilon_k}\right)\right] \\
&\le 0.
\end{aligned}$$

3. Assume (III). Then we can write

$$\begin{aligned}
&\liminf_{\varepsilon\to 0}\left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \\
&\le \liminf_{\varepsilon\to 0} H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \\
&\quad + H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \\
&\quad + H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)
\end{aligned}$$

By the Lipschitz assumption, we have

$$\begin{aligned}
&H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \leq 2C\varepsilon|x_\varepsilon|, \\
&H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) \leq C|x_\varepsilon - y_\varepsilon|\left(1 + 2\frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon^2}\right), \\
&H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2}\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right) \leq 2C\varepsilon|y_\varepsilon|.
\end{aligned}$$

We can then conclude

$$\begin{aligned}
\liminf_{\varepsilon\to 0} &\left[H\left(x_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} + 2\varepsilon x_\varepsilon\right) - H\left(y_\varepsilon, \frac{2(x_\varepsilon - y_\varepsilon)}{\varepsilon^2} - 2\varepsilon y_\varepsilon\right)\right] \\
&\leq 2C\left(\varepsilon(|x_\varepsilon| + |y_\varepsilon|) + \frac{|x_\varepsilon - y_\varepsilon|}{2} + \frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon^2}\right) = 0,
\end{aligned}$$

where the last equality follows from Step 1 and 2 above.

□

1.2.4 *Second-order Hamilton-Jacobi equations*

The discussion of Section 1.2.3 can be applied to a general equation

$$u(x) - \lambda H u(x) = h(x). \tag{1.2.11}$$

Then, we saw that the comparison principle $\sup u - v \leq 0$ for (1.2.11) follows if a inequality of the type

$$\liminf_{\alpha\to\infty} H\left(\frac{\alpha}{2}d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha, \cdot)\right)(y_\alpha) \leq 0 \tag{1.2.12}$$

holds. The estimate (1.2.12), then translates into explicit conditions on H (see Proposition 1.2.10, with $\alpha \to \infty$ instead of $\varepsilon \to 0$). For second order equations, however, this strategy fails. Consider the basic example of the Laplacian operator $Hf(x) = \frac{1}{2}\Delta f(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$. The estimate (1.2.12) translates to

$$H\left(\frac{\alpha}{2}d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2}d^2(x_\alpha, \cdot)\right)(y_\alpha) = 2\alpha,$$

which diverges as $\alpha \to \infty$. Starting with the works of [Jen88; JLS88] it was realized that whereas the first-order viscosity solution method explores the optimizers (x_α, y_α) only in a one-dimensional sense (fix y_α and vary x for the subsolution part, and fix x_α and vary y for the supersolution part), for second-order equations, one needs to explore the neighborhood of the optimizing points in a two-dimensional sense.

We state the key result, Theorem 3.2, of the user's guide [CIL92], which is formulated in terms of the second-order sub- and super-jets defined in Definition 1.2.8.

Theorem 1.2.11 (A simplified version of Theorem 3.2 of [CIL92]). *Let $u : \mathbb{R}^d \to \mathbb{R}$ be upper semi-continuous and $v : \mathbb{R}^d \to \mathbb{R}$ be lower semi-continuous. Suppose (x_α, y_α) is such that*

$$u(x_\alpha) - v(y_\alpha) - \frac{\alpha}{2} d^2(x_\alpha, y_\alpha) = \sup_{x,y} \left\{ u(x) - v(y) - \frac{\alpha}{2} d^2(x,y) \right\}.$$

Then for any $\varepsilon > 0$ there exist $X_\alpha, Y_\alpha \in S(d)$, such that

$$(\alpha(x_\alpha - y_\alpha), X_\alpha) \in \overline{J}^{2,+} u(x_\alpha), \qquad (\alpha(x_\alpha - y_\alpha), Y_\alpha) \in \overline{J}^{2,-} v(y_\alpha)$$

with, $\overline{J}^{2,+}u(x)$ and $\overline{J}^{2,-}v(y)$ the closures of $J^{2,+}u(x)$ and $J^{2,-}v(y)$, and

$$-\left(\frac{1}{\varepsilon} + 2\alpha\right) \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \leq \begin{pmatrix} X_\alpha & 0 \\ 0 & -Y_\alpha \end{pmatrix} \leq \alpha \left(1 + 2\varepsilon\right) \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}. \tag{1.2.13}$$

Using now the definition of sub- and supersolution via generalized differentials (Definition 1.2.9) instead of the definition via test functions (Definition 1.2.5), leads for $Hf(x) = \frac{1}{2}\Delta f(x) = \frac{1}{2} \operatorname{Tr}\left(D^2 f(x)\right)$ to having to establish the estimate

$$\liminf_{\alpha \to \infty} \frac{1}{2} \operatorname{Tr}(X_\alpha) - \frac{1}{2} \operatorname{Tr}(Y_\alpha) \leq 0. \tag{1.2.14}$$

We now perform a "trick". Conjugate the two right-hand matrices (1.2.13) with the matrix

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix},$$

one then obtains

$$\frac{1}{2} \begin{pmatrix} X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \\ X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \end{pmatrix} \leq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

and one observes that (1.2.14) can be estimated by

$$\frac{1}{2} \operatorname{Tr}(X_\alpha) - \frac{1}{2} \operatorname{Tr}(Y_\alpha) = \frac{1}{4} \operatorname{Tr} \begin{pmatrix} X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \\ X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \end{pmatrix} \leq 0$$

leading to the desired estimate.

A natural question then is: if we can treat $Hf(x) = \frac{1}{2}\Delta f(x)$, can we also treat its discrete variant [1]

$$Hf(x) = \frac{1}{2} \left[f(x+1) - f(x) \right] + \frac{1}{2} \left[f(x-1) - f(x) \right]?$$

1 In probabilistic terms, if we can get well-posedness for the generator of Brownian motion, can we can well-posedness for the generator of the, in some sense, simpler random walk.

Clearly, this operator suffers from the same divergence as the Laplacian as in (1.2.14), but Theorem 1.2.11 is not directly of use either.

We will see in Chapter 6 that we can interpret Theorem 1.2.11 as the construction of two test functions $\phi, \psi \in C^2$ that are squeezed in between u and v on one-hand and $\frac{\alpha}{2}d^2$ on the other. To be more precise:

$$u(x_\alpha) - \phi(x_\alpha) = \sup u - \phi \quad \text{and} \quad v(y_\alpha) - \psi(y_\alpha) = \inf v - \psi,$$

and

$$\phi(x_\alpha) - \psi(y_\alpha) - \frac{\alpha}{2}d^2(x_\alpha, y_\alpha) = \sup_{x,y}\left\{u(x) - v(y) - \frac{\alpha}{2}d^2(x,y)\right\}. \tag{1.2.15}$$

As before, the comparison principle now follows from the estimate

$$\liminf_{\alpha\to\infty} H\phi(x_\alpha) - H\psi(y_\alpha) \leq 0.$$

We will show a new strategy to prove (1.2.15), demonstrating its applicability to different types of operators, including the two considered above.

2 OUR MOTIVATION: LARGE DEVIATIONS THEORY

In recent years, Hamilton-Jacobi (HJ) equations have been shown to be versatile in a wide range of applications, including control theory and mean-field games. However, our primary motivation for this research stems from the field of *large deviations theory* (LD).

In our quotidian experiences, we frequently encounter predictable patterns: for instance, the outcome of a die roll tends to average 3.5 over repeated trials, or a weather station records temperatures within a typical range. However, it is the rare and extreme deviations from these patterns, such as rolling a series of sixes or an unprecedented heat wave, that can have significant impacts on systems ranging from games to global infrastructure. The analysis and quantification of these rare events are the focus of a mathematical framework known as large deviations theory, which aims to explore the underlying mechanisms giving rise to such phenomena.

"Any large deviation is done in the least unlikely of all the unlikely ways." - Frank den Hollander, Large Deviations, ([Hol00]).

Essentially, even though a rare event may seem highly unlikely, it does not happen in a random or chaotic manner; instead, it follows a particular pattern that is, indeed, the least improbable of the rare possibilities. This is an important concept in large deviations theory because we are not just looking at the possibility of rare events but also how they actually happen. Instead of just saying that a rare event might occur, we want to understand how it happens in the most probable or typical way given that it is rare. This is crucial for predicting or modeling rare phenomena, especially when systems are complex, as it helps us focus on the most likely paths or trajectories that lead to these rare outcomes.

The first complete framework for large deviation theory was created by Srinivasa Varadhan in his famous paper from 1966 [Var66] (for which he won the Abel's prize). Since then, large deviation theory has grown a lot, with lots of studies linking it to other areas of mathematics and lots of different uses in the natural sciences. We mention a few of the most important ones here. In 1979, Freidlin and Wentzell, [FW79], introduced pathwise large deviations for stochastic processes, which offered a new way of looking at the subject. Ellis, in his work from 1985 [Ell85], looked at the link between large deviations and statistical mechanics, which led to applications in physics. Dembo and Zeitouni's 1998 book introduced lots of new ideas that are now used a lot in large deviation theory. Den Hollander's lectures from 2000 give a detailed overview of large deviations, with lots of examples and applications.

The connection between pathwise Large deviations for stochastic processes and Hamilton - Jacobi equations was first developed by Fleming [Fle77], who characterized the large deviation convergence for exit probabilities as convergence of solutions for a sequence of Hamilton–Jacobi equations. Later, Evans and Ishii [EI85], and Fleming and Souganidis [FS86] applied the theory of viscosity solutions to this context, enabling the approach to cover a wider variety of examples. In 2006, Jin Feng and Thomas G. Kurtz developed a complete and applicable theory to prove large deviations in a metric space in their monograph [FK06]. The idea is that the large deviations behaviour of a sequence of processes X_ε is encoded in the solution of a Hamilton-Jacobi equation with the Hamiltonian H_ε that is expressible in terms of the *linear generator* of the process. The Large deviations statement can hence be derived by proving the convergence (in terms of graph convergence) of H_ε and uniqueness of viscosity solutions (in terms of *comparison principle*) for the limiting Hamilton-Jacobi equation. We refer to Section 2.2 for the detailed discussion about the connection between the two theories.

While the connection between large deviations and Hamilton-Jacobi equations is well-documented in some specific cases, many intriguing examples emerging from the LD framework remain unexplored in the HJ existing literature. In particular, examples arising from biology, biochemistry, and systems with multi-scale dynamics often involve Hamiltonians that contradict the assumptions typically employed in standard approaches. These include irregular or non-differentiable Hamiltonians, non-coercive settings, or scenarios where viscosity solutions lack the usual smoothness, i.e., many examples fall outside the setting of Theorem 1.2.10. Our aim is to bridge this gap by developing methods that extend the well-posedness theory for Hamilton-Jacobi equations to encompass these cases.

This chapter contains an introduction of the large deviations theory (Section 2.1) and a discussion about its connection with Hamilton-Jacobi equation (Section 2.2).

2.1 INTRODUCTION TO THE LARGE DEVIATIONS THEORY

We start this section by presenting two motivating examples that illustrate the need for large deviations theory, followed by the formal definition of the large deviations principle (LDP) and its associated rate function.

Example 1: Rolling a Die and Rare Events

Imagine rolling a standard six-sided die n times. Each roll X_i is independent and takes one of the values $\{1, 2, 3, 4, 5, 6\}$ with equal probability $1/6$. The average outcome of n rolls is:

$$\hat{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

while for a single roll, the expected value is:

$$\mathbb{E}[X_i] = \frac{1+2+3+4+5+6}{6} = 3.5.$$

As we roll the die more times, we might expect $\hat{X}_n$ to get closer to this expected value of 3.5 (we call this fact the *law of large numbers*). But what if we ask:

"how likely is an unusually high average?".

For example, consider the probability that the average result $\hat{X}_n$ exceeds 5, meaning that the rolls are consistently much higher than expected. Achieving such a high average requires a series of unusually high rolls, which seems improbable. As n increases, it becomes increasingly *rare* to observe averages that deviate significantly from 3.5.

Why Is This Important? The probability of observing such extreme deviations decreases rapidly as n increases. However, quantifying how quickly this probability decreases is not immediately obvious. This type of question cannot be answered simply by knowing the expected value or the typical behavior of the die rolls. Instead, large deviations theory provides the tools to quantify the probabilities of rare and extreme outcomes, which are critical in applications like risk analysis and decision-making under uncertainty.

Example 2: Empirical average

Now consider a more general setting: let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables, each with the same distribution and expected value $\mu = E[X_1]$. The empirical average of the first n observations is defined as:

$$\hat{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

The objective is to understand the behavior of this empirical average as n approaches ∞. The analysis will begin with two fundamental facts in probability theory, namely the *law of large numbers* and the *central limit theorem*.

Law of Large Numbers (LLN):

The Law of Large Numbers (LLN) states that as n becomes large, the empirical average $\hat{X}_n$ converges to the expected value μ with high probability. More formally, we have the following:

$$\hat{X}_n \to \mu \quad \text{almost surely as} \quad n \to \infty.$$

This result is intuitive: as we observe more and more random variables, their average becomes closer and closer to the expected value of the underlying distribution.

Central Limit Theorem (CLT):

The Central Limit Theorem (CLT) provides a more precise description of the fluctuations of the empirical average around its expected value for large n. It states that the distribution of the scaled deviation of $\hat{X}_n$ from the expected value μ converges to a normal distribution as n increases. Specifically, for large n, we have:

$$\frac{\hat{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} \mathcal{N}(0,1),$$

where σ^2 is the variance of X_i. This means that the empirical average $\hat{X}_n$ behaves like a normal random variable with mean μ and variance $\frac{\sigma^2}{n}$ for large n. Essentially, for large n, the probability $\mathbb{P}\left(\left|\hat{X}_n - \mu\right| \leq \frac{k}{\sqrt{n}}\right)$ (i.e., deviations of order $O(\frac{1}{\sqrt{n}})$ from the expected value μ) can be approximated using the cumulative distribution function (CDF) of the standard normal distribution.

We call deviations of order $O(\frac{1}{\sqrt{n}})$ *small deviations*.

We can summarize the two facts above with Figure 4.

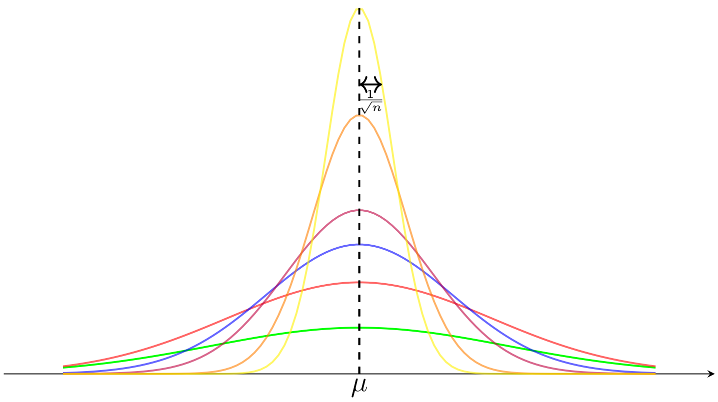

Figure 4: The Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) for the empirical average. As the sample size n increases, the distribution of the empirical average (colored curves) becomes more concentrated around the mean (μ), represented by the dashed line. The horizontal segment highlights the decreasing standard deviation of the empirical average, which is proportional to $\frac{1}{\sqrt{n}}$.

Large Deviations Principle (LDP):

Suppose now we want to calculate the probability of a large deviation for the empirical mean $\hat{X}_n$, specifically:

$$\mathbb{P}(\hat{X}_n \geq \mu + a),$$

where $a > 0$ is a fixed constant. This represents the probability that the empirical mean exceeds the expected value μ by an amount a, which is independent of the number of observations n. Since this deviation is of order $O(1)$, the fluctuations described by the central limit theorem, which are of order $O(1/\sqrt{n})$, cannot be applied here. Instead, we need tools from large deviations theory to quantify the probability of such rare events. We will compute this probability step by step. We begin by rewriting the probability:

$$\mathbb{P}(\hat{X}_n \geq \mu + a) = \mathbb{P}\left(S_n \geq n(\mu + a)\right),$$

where $S_n = \sum_{i=1}^{n} X_i$. By Markov's inequality, for any $\theta > 0$, we have:

$$\mathbb{P}(S_n \geq n(\mu + a)) = \mathbb{P}(e^{\theta S_n} \geq e^{\theta n(\mu+a)}) \leq \frac{\mathbb{E}[e^{\theta S_n}]}{e^{\theta n(\mu+a)}}.$$

The moment generating function of S_n is given by $\mathbb{E}[e^{\theta S_n}] = \left(\mathbb{E}[e^{\theta X_1}]\right)^n = e^{n\Lambda(\theta)}$, where $\Lambda(\theta) = \ln \mathbb{E}[e^{\theta X_1}]$ is the log-moment generating function of X_1. Substituting this, we obtain:

$$\mathbb{P}(S_n \geq n(\mu + a)) \leq e^{n(\Lambda(\theta) - \theta(\mu+a))}.$$

To tighten the bound, we minimize the exponent over $\theta > 0$. The optimal θ^* satisfies the equation:

$$\Lambda'(\theta^*) = \mu + a,$$

where $\Lambda'(\theta) = \frac{d}{d\theta}\Lambda(\theta)$ is the derivative of the log-moment generating function. Substituting θ^* back, we find:

$$\mathbb{P}(\hat{X}_n \geq \mu + a) \leq e^{-nI(\mu+a)},$$

where $I(\mu + a)$ is the *rate function*, defined as:

$$I(\mu + a) = \sup_{\theta > 0} \left\{ \theta(\mu + a) - \Lambda(\theta) \right\}.$$

Hence, for large n, the probability $\mathbb{P}(\hat{X}_n \geq \mu + a)$ behaves approximately as:

$$\mathbb{P}(\hat{X}_n \geq \mu + a) \approx e^{-nI(\mu+a)}.$$

More generally, we can write

$$\lim_{n\to\infty} \frac{1}{n} \ln \left(\mathbb{P} \left(\sum_{i}^{n} X_i \geq nx \right) \right) = -I(x).$$

We say that if $X_1, X_2, \ldots$ are i.i.d. random variables, the distributions of the averages $(\hat{X}_n)_n$ satisfy the large deviation principle with rate function I. This is known as Cramer's theorem (see [Cra38] or [CT22] for details). With the above examples in mind, we write now the formal definition of *rate function* and *large deviations principle*.

We consider $\mathcal{X}$ a *Polish space*, that is a complete separable metric space. Moreover, for a random variable X_n we call $P_n \in \mathcal{P}(\mathcal{X})$ the law of X_n, that is, $P_n(A) = \mathbb{P}(X_n \in A)$.

Definition 2.1.1 (Rate function)**.** A function $I : \mathcal{X} \to [0, \infty]$ is called a *rate function* if it is lower semicontinuous and not identically $+\infty$.

A rate function $I(x)$ is said to be *good* if its level sets

$$\{x \in \mathcal{X} : I(x) \leq \alpha\}, \quad \text{for all } \alpha \in [0, \infty),$$

are compact subsets of $\mathcal{X}$.

Since we will always consider good rate functions, we will often refer to them simply as "rate functions".

Definition 2.1.2 (Large Deviations Principle)**.** A sequence of random variables $\{X_n\}_{n\geq 1}$ satisfies the *Large Deviations Principle (LDP)* with rate function $I(x)$ if, for any Borel measurable set $A \subseteq \mathcal{X}$,

$$-\inf_{x\in A^\circ} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(X_n \in A^\circ) \leq \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(X_n \in \overline{A}) \leq -\inf_{x\in\overline{A}} I(x),$$

where A° and $\overline{A}$ denote the interior and closure of A, respectively.

In numerous applications, the focus is not restricted to isolated outcomes, but extends to the evolution of entire trajectories of a stochastic process over time. This naturally leads to the study of *pathwise large deviations*, which extend the large deviations framework to stochastic processes. Rather than analyzing a sequence of random variables, the focus is now on the likelihood of rare behaviors in the paths of a process. These paths represent the system's evolution over a time interval, capturing how it behaves at every point in time.

For instance, in the context of physical systems, it may be of interest to understand how a particle moves along an atypical trajectory over time. Similarly, in the domain of finance, the study of the rare event of a stock price following an unexpected trend can provide valuable insight. In such scenarios, pathwise large deviations provide a quantitative framework for evaluating the probabilities of such rare trajectories. A crucial reason why pathwise large deviations are important is that they provide insight into the limiting behavior of stochastic processes. The rate function, which assigns a "cost" to each trajectory, has two main roles. Firstly, it quantifies how unlikely a path is. Secondly, it characterizes the most probable way in which rare events occur. This enables us to describe the asymptotic behavior of the process and identify its limiting properties in a mathematically rigorous way.

In this thesis, we will focus on pathwise large deviations for Markov processes.

2.2 CONNECTION BETWEEN LD THEORY AND HJ EQUATIONS

To explain the connection between large deviations and Hamilton-Jacobi equations, and in particular to understand the method of Feng and Kurtz, we are naturally led to the framework of semigroup theory. The key idea is that the evolution of probabilities in Markov processes can be described by a linear semigroup, while the study of large deviations introduces a nonlinear semigroup that emerges from a logarithmic transformation of the linear one. These semigroups play a central role in deriving large deviation principles and connecting them to Hamilton-Jacobi equations. To set the stage, we introduce the necessary notations and definitions.

Let E denote a Polish space (complete, separable metric space). We consider a Markov process $X = (X_t)_{t\geq 0}$, defined by its transition probabilities $P(t, x, A) = \mathbb{P}(X_t \in A \mid X_0 = x)$, where $t \geq 0$, $x \in E$, and A is a Borel subset of E. Associated with X, we define the following semigroups:

Definition 2.2.1 (Linear Semigroup)**.** Let $B(E)$ be the space of bounded measurable functions on E. The linear semigroup $\{S(t)\}_{t\geq 0}$ corresponding to the Markov process X is given by

$$S(t)f(x) := \int_E f(y)P(t,x,\mathrm{d}y),$$

for $f \in B(E)$.

The semigroup property $S(t+s) = S(t)S(s)$ follows from the Chapman-Kolmogorov equation, which states that the transition probabilities satisfy:

$$P(t+s,x,A) = \int_E P(t,x,\mathrm{d}y)P(s,y,A), \quad \forall x \in E, A \subset E.$$

Definition 2.2.2 (Nonlinear Semigroup)**.** Given a sequence of Markov processes $\{X^n\}_{n\in\mathbb{N}}$, let their transition probabilities be denoted by $P_n(t,x,\mathrm{d}y)$. The nonlinear semigroup $\{V_n(t)\}_{t\geq 0}$ associated with X^n is defined as

$$V_n(t)f(x) := \frac{1}{n}\log \int_E e^{nf(y)}P_n(t,x,\mathrm{d}y),$$

for $f \in C_b(E)$.

Note that we can write

$$V_n(t)f(x) := \frac{1}{n}\log\big(S_n(t)e^{nf}(x)\big).$$

Then, the semigroup property $V_n(t+s)f = V_n(t)(V_n(s)f)$ for V_n holds because:

$$V_n(t+s)f(x) = \frac{1}{n}\log\big(S_n(t+s)e^{nf}(x)\big).$$

Since $S_n(t)$ is a semigroup, $S_n(t+s) = S_n(t)S_n(s)$, so:

$$S_n(t+s)e^{nf}(x) = S_n(t)\big(S_n(s)e^{nf}\big)(x).$$

Substituting this into the expression for $V_n(t+s)$, we get:

$$V_n(t+s)f(x) = \frac{1}{n}\log\big(S_n(t)(e^{nV_n(s)f})(x)\big).$$

Applying the definition of $V_n(t)$ again, this simplifies to:

$$V_n(t+s)f(x) = V_n(t)(V_n(s)f)(x).$$

After introducing the semigroup framework, we now turn to a key technical condition required to establish a large deviation principle: *exponential tightness*. Exponential tightness ensures that the probability measures under consideration do not spread their mass too widely in the space, but instead concentrate their mass in compact regions of the state space, with the probability of escaping these regions decaying exponentially fast as the scaling parameter $n \to \infty$.

Definition 2.2.3 (Exponential Tightness)**.** Let $\{P_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on E. We say that $\{P_n\}_{n\in\mathbb{N}}$ is *exponentially tight* if for every $M > 0$, there exists a compact set $K_M \subset E$ such that:

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(K_M^c) < -M,$$

where $K_M^c = E \setminus K_M$ is the complement of K_M.

We say that a sequence of Markov processes $\{X^n\}$ is exponentially tight if the sequence of their transition probabilities $\{P^n\}$ is exponentially tight.

We now present the first theorem, which provides the necessary conditions for establishing the large deviation principle.

Theorem 2.2.4 (Theorem 5.15 of [FK06])**.** *Let $\{V_n(t)\}_{t\geq 0}$ be the nonlinear semigroups associated with the sequence of Markov processes $\{X^n\}$. Suppose the following hold:*

(a) $\{X^n\}$ is exponentially tight,

(b) There exists $V(t) : C(E) \to C(E)$ such that for any $f_n \in C_b(E)$ and $f \in C(E)$, the convergence

$$\lim_{n\to\infty} \|f_n - f\|_E \to 0,$$

implies the convergence

$$\lim_{n\to\infty} \|V_n(t)f_n - V(t)f\| = 0, \quad \forall t \geq 0, \tag{2.2.1}$$

(c) $\{X^n(0)\}$ satisfies a large deviations principle with rate function $\mathcal{I}_0 : C_E[0,\infty) \to [0,\infty]$. Then, the sequence $\{X_n\}$ satisfies a large deviation principle with a rate function $I : C_E[0,\infty) \to [0,\infty]$, given by:

$$\mathcal{I} := \mathcal{I}_0 + \sup_{k\in\mathbb{N}} \sup_{(t_1,\ldots,t_k)} \sum_{i=1}^{k} \mathcal{I}_{t_i - t_{i-1}}(x(t_i)|x(t_{i-1})), \tag{2.2.2}$$

where

$$\mathcal{I}(x|y) := \sup_{f\in C(E)} \left(f(x) - V(t)f(y)\right).$$

Sketch of the proof. The proof is essentially divided into two steps: first, it follows from the exponential tightness and the convergence of the semigroups that the large deviations principle for the first marginal implies the large deviations principle for the finite dimensional marginals (this is Proposition 3.25 of [FK06]). Then, exponential tightness and large deviations principle for the finite dimensional marginals imply large deviations principle for the entire process (this is Theorem 4.28 of [FK06]). □

Problem 1: The application of Theorem 2.2.4 is often challenging due to the non-linear nature of the semigroups $V_n(t)$, making them difficult to express explicitly. The task of identifying a limiting semigroup $V(t)$ is even more complex.

Solution to Problem 1: A solution to Problem 1 takes inspiration from the classical theory of linear semigroup in which the convergence of a sequence of semigroup is connected to the convergence of their *infinitesimal generators*.

Recall that for a linear semigroup $\{T(t)\}_{t\geq 0}$ on a Banach space X, the infinitesimal generator A of the semigroup $\{T(t)\}$ is defined as the limit:

$$Af = \lim_{t\to 0^+} \frac{T(t)f - f}{t}, \quad \text{for all } f \in D(A),$$

where $D(A)$ is the domain of the generator, consisting of all elements $f \in X$ for which the limit exists. The generator A satisfies the semigroup property, meaning that for all $t \geq 0$,

$$\frac{d}{dt}T(t) = AT(t), \quad T(0) = I.$$

A crucial result from linear semigroup theory is that *the convergence of the generators implies the convergence of the semigroups*. Specifically, if A_n are the generators of the linear semigroups $\{T_n(t)\}$, and if $A_n \to A$ in some appropriate sense, then the semigroups $T_n(t)$ converge to $T(t)$ as $n \to \infty$, where $T(t)$ is the semigroup generated by A.

Now, we ask the crucial question: *Can we apply a similar result to nonlinear semigroups*? That is, can we find a definition of "generators" of the nonlinear semigroups $\{V_n(t)\}$ such that their convergence implies the convergence of the semigroups themselves?

To proceed, let us first recall the generator of a Markov process $X(t)$. Consider a Markov process $\{X(t)\}$ on a state space E, governed by transition probabilities $P(t, x, A)$, where $x \in E$ and $A \subseteq E$ is a measurable set. The generator of the Markov process $X(t)$, denoted L, is defined as the infinitesimal generator of the semigroup $S(t)$ associated to $X(t)$.

Now, we turn our attention to the nonlinear semigroups $\{V_n(t)\}$ associated with a sequence of Markov processes $X^n(t)$.

If we try heuristically to compute the generator as in the linear case, we obtain:

$$\begin{aligned}
\frac{d}{dt}V_n(t)f(x)\bigg|_{t=0} &= \frac{d}{dt}\left(\frac{1}{n}\log\left(S_n(t)e^{nf}(x)\right)\right)\bigg|_{t=0} \\
&= \frac{1}{n}\frac{1}{S_n(0)e^{nf}(x)} \cdot \frac{d}{dt}\left(S_n(t)e^{nf}(x)\right)\bigg|_{t=0} \\
&= \frac{1}{n}\frac{1}{e^{nf}(x)} \cdot \frac{d}{dt}S_n(t)\bigg|_{t=0} \cdot e^{nf}(x) \\
&= \frac{1}{n}\frac{L_n\left(e^{nf}(x)\right)}{e^{nf}(x)}.
\end{aligned}$$

This justifies the following definition.

Definition 2.2.5 (Nonlinear Generator)**.** Let $\{V_n(t)\}_{t\geq 0}$ be the nonlinear semigroup associated to a Markov process $\{X^n\}$. The nonlinear generator H_n is defined as

$$H_n f(x) := \frac{1}{n} e^{-nf(x)} L_n e^{nf(x)},$$

where L_n is the generator of the linear semigroup $S_n(t)$, and H_n acts on functions in the domain $\mathcal{D}(H_n) = \{f \mid e^{nf} \in \mathcal{D}(L_n)\}$.

To show that the convergence of $H_n \to H$ implies the convergence of the nonlinear semigroups $V_n(t) \to V(t)$, we rely on some fundamental results from nonlinear semigroup theory such as the *Crandall-Liggett theorem* (see [CL71]). Before stating the theorem, we first recall the key conditions required for its application.

Definition 2.2.6 (Dissipativity)**.** An operator H on a Banach space X is called dissipative, if for all $x, y \in \mathcal{D}(H)$ and $\lambda > 0$, the inequality

$$\|x - y\| \leq \|x - y - \lambda(Hx - Hy)\|$$

holds.

Definition 2.2.7 (Range Condition)**.** We say that an operator H on a Banach space X satisfies the range condition if, for all $\lambda > 0$, the range of $(\mathbb{1} - \lambda H)$ is dense in X, that is

$$\overline{rg(\mathbb{1} - \lambda H)} = X.$$

Remark 2.2.8. The range condition for dissipative operators implies the existence of a unique (classical) solution to the equation $(1 - \lambda\overline{H})f = h$, for every $h \in D(H)$ and $\lambda > 0$. Indeed, if a dissipative operator H satisfies the range condition, its closure $\overline{H}$ is also dissipative and satisfies

$$rg\left(\mathbb{1} - \lambda\overline{H}\right) = \overline{rg(\mathbb{1} - \lambda H)} = X,$$

that is, for every $h \in X$ there exists $f \in D(\overline{H})$ such that $(\mathbb{1} - \lambda\overline{H})f = h$. Uniqueness follows immediately from dissipativity.

We now present the theorem which answers the question posed earlier about the convergence of nonlinear semigroups. The result, given as Proposition 5.5 in [FK06], has been adapted here to fit our notation. Its proof is based on the Crandall-Liggett theory of nonlinear semigroups. For a detailed discussion and proof, we refer the reader to [FK06].

Theorem 2.2.9 (Proposition 5.5 of [FK06])**.** *Let $\tilde{H}_n : B(E) \to B(E)$ and $H : C(E) \to C(E)$ be dissipative operators and suppose that each satisfies the range condition with the same $\lambda > 0$. Suppose that for all $f \in D(H)$ there exist $f_n \in D(\tilde{H}_n)$ such that*

$$\lim_{n\to\infty} \|f_n - f\|_E = 0 \quad \textit{and} \quad \lim_{n\to\infty} \|\tilde{H}_n f_n - Hf\|_E = 0$$

then, there exist $\tilde{V}_n(t)$ and $V(t)$ such that, for all $f \in \overline{D(H)}$ and $f_n \in \overline{D(\tilde{H}_n)}$ satisfying $\|f_n - f\|_E \to 0$,

$$\lim_{n\to\infty} \|\tilde{V}_n(t) f_n - V(t) f\|_E = 0.$$

The following corollary will solve Problem 1 given above.

Corollary 2.2.10 (Corollary 5.19 of [FK06])**.** *Let $\{X^n\}$ be a sequence of E-valued Markov processes. Let H_n be their nonlinear generators. Suppose the following hold:*
(a) $\{X^n\}$ is exponentially tight,
(b) there exists $H : D(H) \subseteq C(E) \to C(E)$ such that all $f \in D(H)$ there exist $f_n \in D(H_n)$ such that

$$\lim_{n\to\infty} \|f_n - f\|_E = 0 \quad \textit{and} \quad \lim_{n\to\infty} \|H_n f_n - Hf\|_E = 0$$

(c) H satisfies the range condition,
(d) $\{X^n(0)\}$ satisfies a large deviations principle with rate function $\mathcal{I}_0 : C_E[0,\infty) \to [0,\infty]$.
Then, the sequence $\{X_n\}$ satisfies a large deviation principle with a rate function $\mathcal{I} : C_E[0,\infty) \to [0,\infty]$, given by (2.2.2)*.*

Sketch of the proof: Consider the nonlinear generators $H_n = \frac{1}{n}e^{-nf(x)}L_n e^{nf(x)}$ and suppose they converge (in the sense as above) to an operator H that satisfies the range condition. Using some techniques from the linear and nonlinear semigroup theory (such as the Hille–Yosida approximation), it is possible to show the existence of additional operators $\tilde{H}_n$ that are dissipative and verify the range condition and such that $\tilde{H}_n$ still converge to H. Then, from this convergence, H inherits the dissipativity (see Lemma 5.7 of [FK06]). Applying Theorem 2.2.9, we can find $\tilde{V}_n(t)$ converging to a semigroup $V(t)$. It is possible to prove that $\tilde{V}_n(t)$ "approximate" the nonlinear semigroups V_n, and, consequently, the convergence (2.2.1) (see Lemma 5.11 of [FK06]). The large deviations principle then follows from Theorem 2.2.4. □

Problem 2: As mentioned in Remark 2.2.8, the range condition for dissipative operators implies the existence of classical solutions to the equation $(\mathbb{1} - \lambda\overline{H})f = h$. However, this is a strong condition that is often not satisfied in practice. Consequently, assuming the range condition can be too restrictive, making the method challenging or even impossible to apply in many cases.

Solution to Problem 2: From Theorem 2.2.9 and Corollary 2.2.10, it is evident that establishing a large deviation principle requires identifying a limit H of the nonlinear generators H_n that is both dissipative and satisfies the range condition. Instead of enforcing the range condition directly on H, we use an alternative approach: finding conditions on H that guarantee the existence of an auxiliary operator $\hat{H}$, which serves as the appropriate limit operator for applying Corollary 2.2.10. This means that the operator $\hat{H}$, constructed from H, satisfies the following properties:
1. $\hat{H}$ verifies the range condition;
2. $\hat{H}$ is dissipative;
3. $H_n \to \hat{H}$ for $n \to \infty$ (in the appropriate sense).

We now analyze each of these properties and discuss how to construct $\hat{H}$ as above. The idea is to construct $\hat{H}$ "extending" H.

Ensuring the range condition for $\hat{H}$: We saw that the range condition guarantees the existence of a unique solution to $(\mathbb{1} - \lambda\hat{H})f = h$. If we can construct a "weak solution"

$R_{\lambda,h}$ (the resolvent) to $(\mathbb{1} - \lambda H)f = h$, then we can define $\hat{H}$ by "adding" to H this resolvent. More precisely, the operator

$$\hat{H} := H \cup \left\{ \left(R_{\lambda,h}, \frac{R_{\lambda,h} - h}{\lambda} \right), \, h \in C(E) \right\} \tag{2.2.3}$$

automatically satisfies the range condition.

Thus, **the key requirement for constructing $\hat{H}$ that satisfies the range condition is the existence and uniqueness of a "weak" solution to** $(\mathbb{1} - \lambda H)f = h$.

Dissipativity of $\hat{H}$: We saw that the limit operator H inherits the dissipativity as the limit of dissipative operators. However, we must ensure that adding $R_{\lambda,h}$ to H does not violate dissipativity.

In semigroup theory, dissipativity is often implied by the *maximum principle* (This is indeed the case of a nonlinear generator of a Markov process). We give here the proper definition.

Definition 2.2.11 (Maximum principle)**.** We say that an operator $H \subseteq C(E) \times C(E)$ satisfies the *maximum principle* if, for all $f_1, f_2 \in \mathcal{D}(H)$ and $x_0 \in E$ with

$$f_1(x_0) - f_2(x_0) = \sup_{x \in E} \{ f_1(x) - f_2(x) \},$$

we have

$$Hf_1(x_0) \leq Hf_2(x_0)$$

and, analogously, for all $f_1, f_2 \in \mathcal{D}(H)$ and $x_0 \in E$ with

$$f_1(x_0) - f_2(x_0) = \inf_{x \in E} \{ f_1(x) - f_2(x) \},$$

we have

$$Hf_1(x_0) \geq Hf_2(x_0).$$

Now consider the operator $\hat{H}$ as in (2.2.3). To preserve the maximum principle for $\hat{H}$, we take $f \in D(H)$ and assume that x_0 satisfies

$$\sup_x (R_{\lambda,h} - f)(x) = (R_{\lambda,h} - f)(x_0).$$

Then, using the definition of $\hat{H}$, we obtain

$$\begin{aligned} \hat{H} R_{\lambda,h}(x_0) - \hat{H} f(x_0) &= \frac{1}{\lambda} \left(R_{\lambda,h}(x_0) - h(x_0) \right) - Hf(x_0) \\ &= \frac{1}{\lambda} \left(R_{\lambda,h}(x_0) - \lambda Hf(x_0) - h(x_0) \right). \end{aligned}$$

Consequently, for dissipativity to hold, we require

$$R_{\lambda,h}(x_0) - \lambda Hf(x_0) - h(x_0) \leq 0.$$

The attentive reader may have noticed that the above equation corresponds to the definition of viscosity subsolution for $(\mathbb{1} - \lambda H)f = h$ (cf. Definition 1.2.5).

Thus, **to ensure that $\hat{H}$ remains dissipative, we require that $R_{\lambda,h}$ is a viscosity solution to the equation $(\mathbb{1} - \lambda H)f = h$.**

Convergence of H_n to $\hat{H}$: Finally, Condition 3 follows from the convergence of H_n to H as established in Lemma 6.9 of [FK06].

Summary of the key conditions: To ensure that $\hat{H}$ satisfies all three required properties - range condition, dissipativity, and convergence - we need:

1. The nonlinear generators H_n converge to an operator H;
2. There exists a unique viscosity solution $R_{\lambda,h}$ to the equation $(\mathbb{1} - \lambda H)f = h$, for every $\lambda > 0$ and $h \in C(E)$.

The existence of $R_{\lambda,h}$ follows from the convergence $H_n \to H$ (see Lemma 6.9 of [FK06]), while uniqueness is ensured by the comparison principle (cf. Definition 1.2.3).

This discussion leads to the following final theorem.

Theorem 2.2.12 (Simplification of Theorems 6.13 and 6.14 of [FK06])**.** *Let $\{X^n\}$ be a sequence of E-valued Markov processes. Let H_n be their nonlinear generators. Suppose the following hold:*

(a) There exists $H : D(H) \subseteq C(E) \to C(E)$ such that all $f \in D(H)$ there exist $f_n \in D(H_n)$ such that

$$\lim_{n\to\infty} \|f_n - f\|_E = 0 \quad \textit{and} \quad \lim_{n\to\infty} \|H_n f_n - Hf\|_E = 0.$$

(b) For each $\lambda > 0$, the comparison principle holds for

$$(\mathbb{1} - \lambda H)f = h. \tag{2.2.4}$$

(c) $\{X^n\}$ is exponentially tight,

(d) $\{X^n(0)\}$ satisfies a large deviations principle with rate function $\mathcal{I}_0 : C_E[0,\infty) \to [0,\infty]$.

Then,

(a) For each $h \in C(E)$, there exists a unique viscosity solution of (2.2.4) *which we will denote by $f = R_{\lambda,h}$.*

(b) The operator defined as

$$\hat{H} = \bigcup_{\lambda} \left\{ \left(R_{\lambda,h}, \frac{R_{\lambda,h} - h}{\lambda} \right) : h \in C(E) \right\},$$

extends H, is dissipative and satisfies the range condition.

(e) Letting $\{V_n(t)\}$ denote the nonlinear semigroup associated to $\{X^n\}$, there exists $V(t)$ such that, for each $f \in D(H)$ and $f_n \in D(H_n)$ satisfying $\|f_n - \eta_n f\| \to 0$,

$$\lim_{n\to\infty} \|V_n(t) f_n - \eta_n V(t) f\| = 0.$$

(d) $\{X^n\}$ satisfies the large deviations principle with rate function given by (2.2.2).

We can summarize the discussion above with the following diagram.

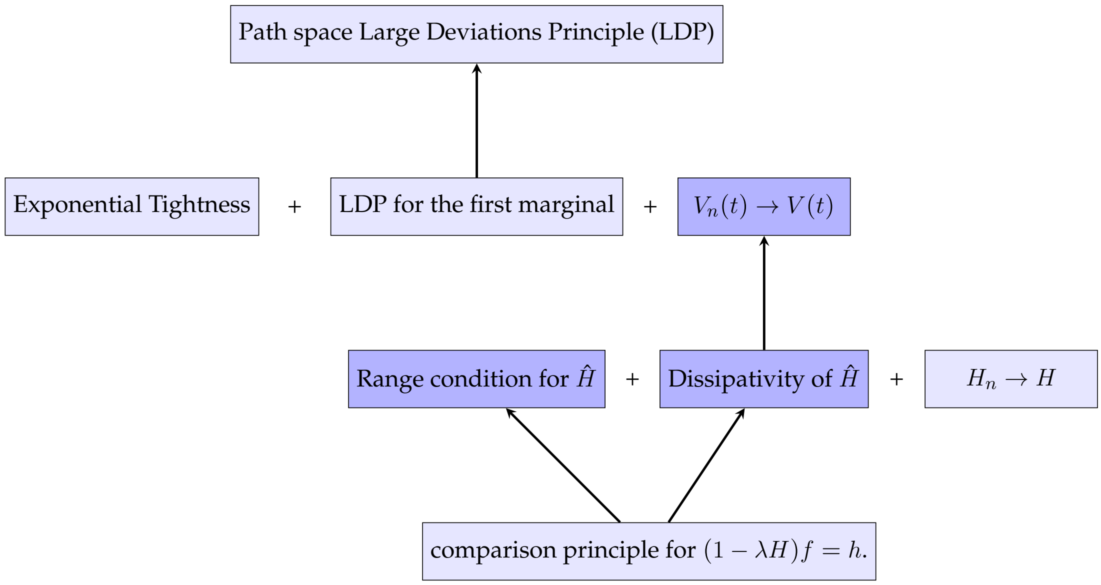

Figure 5: Diagram summarizing the Feng and Kurtz method. An arrow from node A to node B indicates that A implies B. If there is a "+" symbol between two nodes, it means that all elements together imply the subsequent node.

OUTLINE OF THE THESIS

The thesis is organized into three main parts.

The first part, **Introduction and theoretical background**, consists of Chapter 1 and Chapter 2. This part provides an overview of the basic ideas and frameworks necessary to understand the results of this thesis.

Chapter 1 begins by introducing a motivating example from control theory, which illustrates the relevance and applications of Hamilton-Jacobi equations in practical scenarios. This chapter then provides a broad introduction to the mathematical context of Hamilton-Jacobi equations, focusing on the role of viscosity solutions as a key tool for analyzing situations where classical solutions may fail.

Chapter 2 begins with an introduction to the field of large deviations. Then, it provides a discussion about the connection between Large deviations theory and Hamilton-Jacobi equations.

The second part, **Motivations: Examples from a Large Deviations Context**, includes Chapter 3. It presents examples arising from biology and biochemistry, particularly focusing on large deviations for two-scale Markov processes with homogenization. These examples highlight how the connection between large deviations with the Hamilton-Jacobi equations can provide new insights.

The third and last part, **Well-Posedness of Hamilton-Jacobi Equations**, encompasses Chapters 4 to 6 and focuses on the rigorous mathematical analysis of Hamilton-Jacobi equations. Chapter 4 provides a well-posedness result for a general Hamilton-Jacobi-

Bellman equation that includes the one related to the large deviations problem of Chapter 3. Chapter 5 considers the question of existence of viscosity solutions for first-order equations. Chapter 6 turns to second-order Hamilton-Jacobi equations, presenting a new proof of the comparison principle and exploring conditions under which well-posedness can be established for this more complex class of equations. Together, these chapters expand the theoretical boundaries of Hamilton-Jacobi equations by developing techniques that address scenarios beyond the scope of existing literature.

The thesis concludes with a summary of the main findings, reflecting on their contributions to the fields of Hamilton-Jacobi equations and large deviations theory. Appendices include detailed mathematical proofs and supplementary discussions, offering additional insights and support for the core results presented in the thesis.

NOTATIONS

- $\mathbb{R}^N$ – the Euclidean N-dimensional space.
- $B(E)$ – the space of functions $u : E \to \mathbb{R}$ with $\|u\|_\infty < +\infty$.
- $C(E)$ – the space of continuous functions $u : E \to \mathbb{R}$.
- $UC(E)$ – the space of uniformly continuous functions $u : E \to \mathbb{R}$.
- $Lip(E)$ – the space of Lipschitz continuous functions $u : E \to \mathbb{R}$.
- $Lip_{\text{loc}}(E)$ – the space of locally Lipschitz continuous functions $u : E \to \mathbb{R}$.
- $C^k(E)$ – for $k \geq 1$, the space of of functions with continuous partial derivatives up to order k.
- $C_b^k(E)$ - the set of all functions in $C^k(E)$ with bounded derivatives up to order k.
- $C_c^\infty(E)$ - the set of all smooth functions that are constant outside of a compact set.
- $C_u(E)$ and $C_l(E)$ - the set of continuous functions on E that are uniformly bounded from above and below, respectively.
- $C^\infty(E)$ the space of infinitely differentiable functions from $E \subseteq \mathbb{R}^d$ to $\mathbb{R}$.
- $AC(E)$ the space of absolutely continuous functions from $E \subseteq \mathbb{R}^d$ to $\mathbb{R}$.
- $USC(E)$, $LSC(E)$ – the spaces of lower and upper semicontinuous functions $u : E \to \mathbb{R}$.
- $BUSC(E)$, $BLSC(E)$ – the spaces $USC(E) \cap B(E)$ and $LSC(E) \cap B(E)$.
- $\mathbb{P}$ - a probability measure.
- $\pi_\# \mathbb{P}(E) = \mathbb{P}(\pi^{-1}(E))$ - the push-forward measure of $\mathbb{P}$.
- $\mathcal{P}(E)$ – the space of the probability measures on E.
- $\langle x, y \rangle$ – the scalar product of vectors $x = (x_1, \ldots, x_N)$ and $y = (y_1, \ldots, y_N)$.
- $B(x_0, r)$ – the open ball $\{x \in \mathbb{R}^N : |x - x_0| < r\}$.
- $\overline{B}(x_0, r)$ – the closed ball $\{x \in \mathbb{R}^N : |x - x_0| \leq r\}$.
- $\overline{E}$ - the closure of the set E.
- E° - the interior of a set E.
- E^c - the complement of a set E.
- $\operatorname{co} E$ – the convex hull of the set E.
- $\overline{\operatorname{co}} E$ – the closed convex hull of the set E.
- $d(x, E) = \operatorname{dist}(x, E)$ – the distance from x to E.
- $y \wedge o = \min\{y, o\}$ for $y, o \in \mathbb{R}$.
- $y \vee o = \max\{y, o\}$ for $y, o \in \mathbb{R}$.
- $\operatorname{sgn}(r)$ – the sign of $r \in \mathbb{R}$ (1 if $r > 0$, -1 if $r < 0$, 0 if $r = 0$).
- $[x]_{\mathbb{Z}^d} = \{y \in \mathbb{R}^d : x - y \in \mathbb{Z}^d\}$ the equivalence class of x with respect to the relation defined by $\mathbb{Z}^d$.
- $\arg\min u$ – the set of minimum points of $u : E \to \mathbb{R}$.
- $\|u\|_\infty$ – the supremum norm $\sup_{x \in E} |u(x)|$ of a function $u : E \to \mathbb{R}$.
- $w(t)$ – a modulus of continuity.
- $Du(x)$ – the gradient of the function u at x.
- $\Delta u = \nabla^2 u$ – the Laplacian of the function u.
- $D^- u(x), D^+ u(x)$ - the subdifferential and superdifferential of u at x.

- $\mathcal{N}(\mu, \sigma^2)$ - the normal distribution with mean μ and variance σ^2.
- $C_E[0, \infty)$ is the space of functions defined on $[0, \infty)$ and taking value in a metric space E.

Part II

MOTIVATIONS: EXAMPLES ARISING FROM A LARGE DEVIATIONS CONTEXT

3

LARGE DEVIATIONS PRINCIPLE FOR TWO-SCALES MARKOV PROCESSES IN BIOLOGY AND BIOCHEMISTRY

In this chapter, we prove the large deviations principle for two examples of two-scales Markov processes via the method through Hamilton-Jacobi equations explained in Chapter 2 . The examples are motivated by applications in biology and biochemistry.

The results presented in Chapter 3.1 are based on:

S.Della Corte and R.C.Kraaij, "Large deviations for Markov processes with switching and homogenisation via Hamilton-Jacobi equations", *Stochastic Processes and their Applications*, 2024.

The results presented in Chapter 3.2 are based on part of:

S.Della Corte and R.C. Kraaij, " Well-posedness of a Hamilton-Jacobi-Bellman equation in the strong coupling regime", Preprint.

3.1 LARGE DEVIATIONS FOR MARKOV PROCESSES WITH SWITCHING AND HOMOGENIZATION MODELING MOTOR PROTEINS

In biochemical and biophysical processes occurring in a cell, an important role is played by several classes of active enzymatic molecules, generally called *motor proteins* or *molecular motors*. These motors are protein molecules that convert chemical energy into mechanical work and motion (see [JAP97],[KF07],[Kol13],[YFRH02] for more details). In the last decades, such biological phenomena have been largely investigated and this analysis was partly possible due to the contribution of the analysis of particular Markov processes called *"switching Markov processes"* (see for instance [JAP97], [CKK03], [CHK04], [PS07] and [YZ10]).

The process that we will consider is such a process. It is a two–component process (X_t, I_t) where the first component X_t is a drift–diffusion process and the second component I_t is a jump process on a finite set. In the context of molecular motors, the spatial component X_t models the location of the motor, for example on a filament, while I_t models the molecular configuration. The two processes together evolve in accordance with the following stochastic differential equation

$$\begin{aligned} &\mathrm{d}X_t = -\nabla\psi(X_t, I_t)\mathrm{d}t + \mathrm{d}B_t, \qquad (3.1.1)\\ &\mathbb{P}\Big(I(t+\Delta t) = j \,|\, I(t) = i, X(t) = x\Big) = r_{ij}\,(x)\,\Delta t + \mathcal{O}(\Delta t^2) \quad \text{as } \Delta t \to 0, \end{aligned}$$

with $\psi \in C^\infty(\mathbb{R}^d \times \{1, \dots, J\})$, $r_{ij} \in C^\infty(\mathbb{R}^d)$ and ∇ is the gradient with respect to x and B_t is the Brownian motion.

It is clear that the two processes are linked by their rate functions. This means that I_t stays in a first discrete state for a random duration while the diffusion component X_t evolves following a stochastic differential equation with a particular drift. Then, when a switch of the configurational component occurs, the potential ψ changes and therefore X_t diffuses according to a new equation up to another switch of I_t (see Figure 6 for a typical behavior of this type of processes). For more details about the construction of such switching hybrid diffusions see [YZ10].

To allow more flexibility to separate the local dynamics as caused by the internal switching, and macroscopic effects, e.g. modeling the presence of energy molecules in the solution, we will work with ψ_ε and r_ε instead of ψ and r, and we will typically assume that $\{\psi_\varepsilon, r_\varepsilon\}_\varepsilon$ exhibit a separation of scales. The simplest instance of this separation of scale is that

$$\psi_\varepsilon(x,i) = \psi_1(\varepsilon x, i) + \psi_2(x,i),$$
$$r_\varepsilon(x,i,j) = r_1(\varepsilon x, i, j) + r_2(x,i,j),$$

i.e. ψ_1 and r_1 model the global macroscopic scale while ψ_2 and r_2 correspond to the local dynamics. We will also assume that ψ_2 is 1–periodic. Moreover, the most general context that we will consider is such that the sequences of functions ψ_ε and r_ε are actually given by two functions $\psi \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d \times \{1,\ldots,J\})$ and $r_{ij} \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ as

$$\begin{aligned} \psi_\varepsilon(x,i) &= \psi(\varepsilon x, x, i), \\ r_\varepsilon(x,i,j) &= r(\varepsilon x, x, i, j). \end{aligned} \tag{3.1.2}$$

The process arising from the stochastic differential equation with ψ_ε and r_ε as drift and rate function will be called $(X_t^\varepsilon, I_t^\varepsilon)$. However, we are interested in the macroscopic motion of the molecule. Therefore, we work with the rescaled process or "zoomed out" process that we obtain by scaling in space and time by the positive parameter $\varepsilon > 0$. More precisely, we look at $(Y_t^\varepsilon, \bar{I}_t^\varepsilon) := (\varepsilon X^\varepsilon_{\varepsilon^{-1}t}, I^\varepsilon_{\varepsilon^{-1}t})$ that evolves according the stochastic differential equation

$$\mathrm{d}Y_t^\varepsilon = -\nabla\psi\left(Y_t^\varepsilon, \frac{Y_t^\varepsilon}{\varepsilon}, \bar{I}_t\right)\mathrm{d}t + \sqrt{\varepsilon}\,\mathrm{d}B_t$$
$$\mathbb{P}\Big(\bar{I}^\varepsilon(t+\Delta t) = j \,|\, \bar{I}^\varepsilon(t) = i, Y^\varepsilon(t) = x\Big) = \frac{1}{\varepsilon} r_{ij}\left(x, \frac{x}{\varepsilon}\right)\Delta t + \mathcal{O}(\Delta t^2) \quad \text{as } \Delta t \to 0$$

with ψ and r_{ij} given by (3.1.2), and we are interested in the limit $\varepsilon \to 0$.

Intuitively, looking from far away at the process the periodicity becomes smaller and smaller as ε decreases and the internal flip rate diverges. Thus, we expect the periodicity and the internal dynamics to homogenise, effectively obtaining a deterministic limit X_t. The numerical simulation in Figure 7 confirms our intuition. It shows sample paths of numerical approximations of a particular switching process $(Y_t^\varepsilon, \bar{I}_t^\varepsilon)$ for various ε. The figure suggests that for small ε, the spatial component Y_t^ε tends to concentrate around a limiting path that, in the case of the simulated process, is a path with constant velocity.

The aim of this work is to investigate large deviations around such deterministic limit of this kind of process. Showing a large deviations principle, we then will be able to characterize the limit path using a Lagrangian rate function. Indeed, we will show in the main theorem, Theorem 3.1.6, that there exists a non negative rate function (Cf. Definition 2.1.1) $\mathcal{I} : C_{\mathbb{R}^d}[0,\infty) \to [0,\infty]$ with which $\{Y_t^\varepsilon\}_{\varepsilon>0}$ satisfies a path-wise large deviation principle in the sense of Definition 2.1.2. Intuitively, it means that

$$\mathbb{P}\left(Y^\varepsilon \approx x\right) \sim e^{-\mathcal{I}(x)/\varepsilon} \quad \varepsilon \to 0,$$

with $\mathcal{I}$ written in terms of a Lagrangian function. This means that Y^ε has a limit path $\tilde{x} \in \mathbb{C}_{\mathbb{R}^d}$ and this limit is the unique minimizer of the rate function $\mathcal{I}$. Moreover, for any path $x \neq \tilde{x}$ such that $\mathcal{I}(x) > 0$, the probability that X^ε is close to x is exponentially small in ε^{-1}. In Corollary 3.1.8, we characterize the minimum of $\mathcal{I}$ by finding a representation of $\partial_t \tilde{x}$ in terms of the drift ψ similar to what one would expect from an averaging principle.

Our work falls into a long tradition of studying the dynamical large deviations around limiting trajectories starting with [FW98] for small noise diffusions, and [Lip96] for two-scale systems. Then, in the last decades it has been used for the study of different kind of processes (see for instance [BFG13] or [KS20]). Regarding jump-diffusion, there are very few large deviations results but see for example [Pop18] and [KP17]. More recently, in [PS19], the authors prove large deviations for a class of switching Markov processes and apply their result to examples, including the molecular motors model. Regarding this example, our work removes two unnatural assumptions in [PS19]. Our main contributions are the transition from a compact to a non-compact setting and the introduction of the global macroscopic effects in rates with the two components ψ_1 and r_1. These two facts complicate the proof of large deviations principle. Most important, but without going into details, we need to prove the comparison principle for a spatially inhomogeneous Hamilton–Jacobi–Bellman equation where the two above generalizations introduce non-trivial complications.

Indeed, we prove the large deviations property using the method due to Feng and Kurtz ([FK06]), explained in Chapter 2, in which a central role is played by associated Hamilton–Jacobi–Bellmann equations. We will explain in more details the main innovations compared to [PS19] of our work in Section 3.1.9.

The chapter is organized as follows. We first give some preliminary contents in Section 3.1.1. The main theorem, the large deviations principle for the switching process modeling molecular motors, is given in Section 3.1.2 together with its proof. Finally, in Section 3.1.8 we are able to extract the main mathematical structures that we use in the previous section and use them in a large deviations result for a more general class of Switching Markov processes.

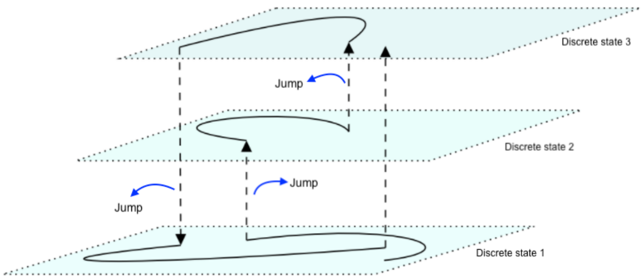

Figure 6: A typical evolution of a process $(X(t), I(t))$

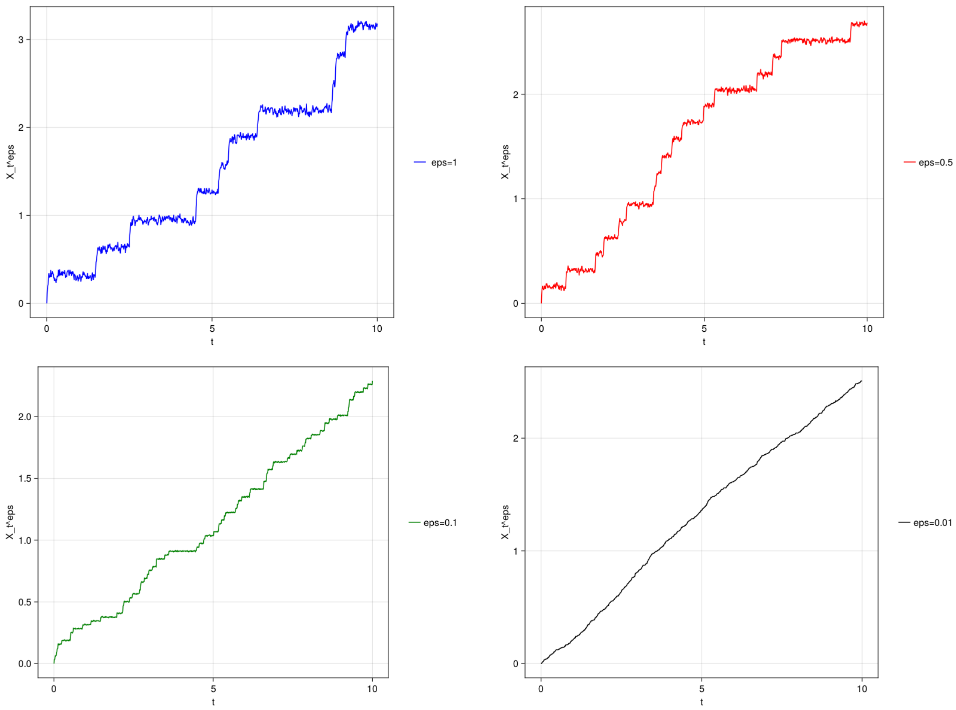

Figure 7: Sample paths of a numerical simulation of the spatial component Y_t^ε of a switching process for different values of ε. We chose a drift ψ_ε^i equal to the periodic part ψ_2 for all $i \in \{1, \dots, 4\}$. We took $\psi_2(\frac{x}{\varepsilon}, i)$ equal to $\sin(x/\varepsilon)$, $\cos(x/\varepsilon)$, $-\sin(x/\varepsilon)$ and $-\cos(x/\varepsilon)$ for $i = 1, 2, 3, 4$ respectively and a rate equal to 1. The jump process switches from a value $i \in \{1, 2, 3\}$ to the value $i + 1$ and from 4 to 1. In this way, the process starts diffusing around a minimum of $\sin(x/\varepsilon)$ $(i = 1)$. The first horizontal part of the "stair" corresponds to this evolution. Then, a switch of $\bar{I}_t^\varepsilon$ takes place, so the value of i becomes $i = 2$, and then the spatial component goes to diffuse around a minimum of $\cos(x/\varepsilon)$, that is the second horizontal part, until another switch.

3.1.1 *Preliminaries*

We begin with the definition of the process that we are going to study. It is a two component Markov process $(X_t^\varepsilon, I_t^\varepsilon)$ to which we refer in all the work with *"molecular motors model"* or *"motor proteins model"*.

Definition 3.1.1 (Molecular motors)**.** Given an integer J, we consider the setting $E = \mathbb{R}^d \times \{1,\dots,J\}$. For all i,j in $\{1,\dots,J\}$, let $r_{ij} \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d;[0,\infty))$ denote nonnegative smooth maps, $\psi^i \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ a smooth and $\nabla\psi^i$ its gradient with respect to x. We suppose that ψ^i grows at most linearly in the first component and is periodic in the second one. Finally, given the following operator

$$\widetilde{A}_\varepsilon f(x,i) := -\nabla\psi^i(\varepsilon x,x)\cdot\nabla_x f(\cdot,i)(x)+\frac{1}{2}\Delta_x f(\cdot,i)(x)+\sum_{j=1}^J r_{ij}(\varepsilon x,x)\left[f(x,j)-f(x,i)\right],$$

we define the E-valued Markov process $(X_t^\varepsilon, I_t^\varepsilon)|_{t\geq 0}$ as the solution to the martingale problem corresponding to $\widetilde{A}_\varepsilon$. More precisely, $(X_t^\varepsilon, I_t^\varepsilon)$ is such that for all $f \in D(\widetilde{A}_\varepsilon)$,

$$f(X^\varepsilon(t),I^\varepsilon(t)) - f(X^\varepsilon(0),I^\varepsilon(0)) - \int_0^t \widetilde{A}_\varepsilon f(X^\varepsilon(s),I^\varepsilon(s))\,ds$$

is a martingale.

Remark 3.1.2. In our case r_{ij} is regular enough that the martingale problem associated to A_ε is well posed (see [EK09] and [SV79]).

Remark 3.1.3. It is straightforward to see that the above defined process solves the stochastic differential equation (.0.1) given in the introduction of this chapter.

We firstly study the above particular model for which we prove the large deviations property. Then, using this model, we lead to a theorem for a general class of processes called *Switching Markov process* (see Section 3.1.8).

As mentioned in the introduction of this chapter, we will work with the rescaled process $(Y_t^\varepsilon, \bar{I}_t^\varepsilon) =$
$= \left(\varepsilon X_{t/\varepsilon}^\varepsilon, I_{t/\varepsilon}^\varepsilon\right)$. Then, by the chain rule, the generator becomes

$$\begin{aligned} &A_\varepsilon f(x,i) \qquad (3.1.3)\\ &= -\nabla\psi^i\left(x,\frac{x}{\varepsilon}\right)\nabla_x f(\cdot,i)(x) + \frac{\varepsilon}{2}\Delta_x f(\cdot,i)(x) + \frac{1}{\varepsilon}\sum_{j=1}^J r_{ij}\left(x,\frac{x}{\varepsilon}\right)[f(x,j)-f(x,i)]. \end{aligned}$$

We will assume in the main theorem that the matrix

$$(R_{ij}(x))_{ij} := (\sup_{y\in\mathbb{R}^d} r_{ij}(x,y))_{ij}$$

is irreducible. Here we give the rigorous definition.

Definition 3.1.4. We say that a matrix $A = (A_{ij}(x))_{ij\in\{1,\dots,J\},x\in\mathbb{R}^d}$ is irreducible if there is no decomposition of $\{1,\dots,J\}$ into two disjoint sets $\mathcal{J}_1$ and $\mathcal{J}_2$ such that $A_{ij} = 0$ on $\mathbb{R}^d$ whenever $i \in \mathcal{J}_1$ and $j \in \mathcal{J}_2$.

The main goal of this work is to prove that the spatial component Y^ε of the above Markov process verifies the large deviation principle. We refer to Chapter 2 for the definitions of Large deviations principle, rate function and exponential tightness.

To prove the exponential tightness we will use the following concept.

Definition 3.1.5 (Compact containment condition)**.** We say that the processes $(Z_\varepsilon(t))$ satisfy the *exponential compact containment condition* if for all $T > 0$ and $a > 0$ there is a compact set $K = K(T, a) \subseteq E$ such that

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left[Z_\varepsilon(t) \notin K \text{ for some } t \in [0, T]\right] \leq -a.$$

3.1.2 *The large deviations principle*

Now we state the main theorem in which we prove sufficient conditions for the large deviation property for the spatial component of the switching process defined in Definition 3.1.1.

Theorem 3.1.6 (Large deviation for the "molecular motors model")**.** *Let $(X_t^\varepsilon, I_t^\varepsilon)$ be the Markov process of Definition 3.1.1. Suppose that the matrix $R_{ij} = (\sup_{y\in\mathbb{R}^d} r_{ij}(y))_{ij}$ is irreducible. Denote $Y_t^\varepsilon = \varepsilon X_{t/\varepsilon}^\varepsilon$ the rescaled process. Suppose further that at time zero, the family of random variables $\{Y^\varepsilon(0)\}_{\varepsilon>0}$ satisfies a large deviation principle in $\mathbb{R}^d$ with good rate function $\mathcal{I}_0 : C_{\mathbb{R}^d}[0,\infty) \to [0,\infty]$. Then, the spatial component $\{Y_t^\varepsilon\}$ satisfies a large deviation principle in $C_{\mathbb{R}^d}[0,\infty)$ with good rate function $\mathcal{I} : C_{\mathbb{R}^d}[0,\infty) \to [0,\infty]$ given in the integral form*

$$\mathcal{I}(x) = \begin{cases} \mathcal{I}_0(x(0)) + \int_0^\infty \mathcal{L}\left(x(t), \dot{x}(t)\right)\, dt & \text{if } x \in AC([0,\infty); \mathbb{R}^d), \\ \infty & \text{else}, \end{cases}$$

with $\mathcal{L}(x,v) = \sup_p\{p\cdot v - \mathcal{H}(x,p)\}$ the Legendre transform of a Hamiltonian $\mathcal{H}(x,p)$ given in variational form by

$$\mathcal{H}(x,p) = \sup_{\mu\in\mathcal{P}(E')} \left[\int_{E'} V_{x,p}(z)\, d\mu(z) - I_{x,p}(\mu)\right], \tag{3.1.4}$$

where $E' = \mathbb{T}^d \times \{1, \ldots, J\}$,

$$V_{x,p}(y,i) = \frac{1}{2}p^2 - p\cdot\nabla_x\psi^i\,(x,y)$$

and the map $I_{x,p} : \mathcal{P}(E') \to [0,\infty]$ is the Donsker–Varadhan function, i.e.

$$I_{x,p}(\mu) = -\inf_{\varphi} \int_{E'} e^{-\varphi} L_{x,p}(e^{\varphi})\, d\mu,$$

where the infimum is taken over vectors of functions $\varphi(\cdot,i) \in C^2(\mathbb{T}^d)$, and $L_{x,p}$ is the operator defined by

$$L_{x,p}u(z,i) = \frac{1}{2}\Delta_z u(z,i) + (p - \nabla_x\psi^i(x,z))\cdot\nabla_z u(z,i) + \sum_{j=1}^J r_{ij}(x,z)\left[u(z,j) - u(z,i)\right]. \tag{3.1.5}$$

Remark 3.1.7. E' captures the periodic behaviour and the internal state. In the homogenisation context described in the introduction, E' is exactly what is being homogenised over while $L_{x,p}$ describes the dynamics on it.

The following corollary characterises the limit process.

Corollary 3.1.8. *Consider the same assumptions of Theorem 3.1.6 for the Markov process* $(Y_t^\varepsilon, I_t^\varepsilon)$*. Then, the spatial component converges almost surely to the path with velocity given by*

$$\partial_t x = \partial_p \mathcal{H}(x,0) = -\int_{E'} \nabla_x \psi^i(x,y)\, d\mu_x^*(y),$$

with μ_x^* *the unique stationary measure of the operator* $L_{x,0}$ *given in* (3.1.5).

Proof. By Theorem A.1 of [PS19], the spatial component Y_t^ε converges almost surely to the set of minimizers of the rate function. More precisely,

$$d(Y_t^\varepsilon, \{\mathcal{I}=0\}) \to 0 \qquad \text{a.s. as } \varepsilon \to 0$$

where $\{\mathcal{I}=0\} = \{x \in C_{\mathbb{R}^d}[0,\infty) \,:\, \mathcal{I}(x) = 0\}$. We now prove that this set is actually a singleton and then characterise the unique element.

With this aim, note that by [Roc70, Theorem 23.5], v is a minimizer of $\mathcal{L}$ if and only if $v \in \partial_p \mathcal{H}(x,0)$. Moreover, by [HUL01, Theorem 4.4.2],

$$\begin{aligned}&\partial_p \mathcal{H}(x,0)\\ &= co\left\{\bigcup_{\mu\in\mathcal{P}(E')} \partial\left[\int_{E'} V_{x,0}\, d\mu - I_{x,0}(\mu)\right] \text{ for all } \mu \text{ s.t. } \mathcal{H}(x,0) = \int_{E'} V_{x,0}\, d\mu - I_{x,0}(\mu)\right\},\end{aligned}$$

where with co we refer to the convex hull of a set and $\partial\left[\int_{E'} V_{x,0}\, d\mu - I_{x,0}(\mu)\right]$ is the differential of the convex functions $\int_{E'} V_{x,p}\, d\mu - I_{x,p}(\mu)$ for $p = 0$.

We know that $\mathcal{H}(x,0) = 0$ and $V_{x,0}(z) = 0$ for all $z \in E'$. Then, if μ_x^* is the optimal measure for $\mathcal{H}(x,0)$, we have that

$$0 = \mathcal{H}(x,0) = \int_{E'} V_{x,0}(z)\, d\mu_x^* - I_{x,0}(\mu_x^*) = I_{x,0}(\mu_x^*).$$

We can conclude that the optimal μ_x^* is the unique stationary measure of $L_{x,0}$ (see Proposition 3.1.38 in the appendix for existence and uniqueness of μ_x^*). We thus find that $\partial_p \mathcal{H}(x,0) = \left\{\frac{\partial}{\partial p}\mathcal{H}(x,p)|_{p=0}\right\}$ and hence $\mathcal{I}(x) = 0 \iff \partial_t x = \frac{\partial}{\partial p}\mathcal{H}(x,0)$ for almost all t and

$$\begin{aligned}\partial_t x = \frac{\partial}{\partial p}\mathcal{H}(x,0) &= \int_{E'} \left.\frac{\partial V_{x,p}(z)}{\partial p}\right|_{p=0} d\mu_x^*(z)\\ &= -\int_{E'} \nabla_x \psi(x,z)\, d\mu_x^*(z).\end{aligned}$$

□

Remark 3.1.9. The above corollary confirms the suggestion of Figure 7 that, when there is no dependence on x in the drift, the spatial component is converging to a path with constant speed. Indeed, for small ε, Y_t^ε tends to concentrate around a path with a constant velocity $v = \partial_p \mathcal{H}(0)$.

Using the discussion of Chapter 2 and Theorem 2.2.12, we can prove Theorem 3.1.6.

Proof of Theorem 3.1.6. We claim the following five facts:

1. The nonlinear generators $H_\varepsilon f = \varepsilon e^{-f/\varepsilon} A_\varepsilon e^{f/\varepsilon}$ of Y_t^ε converge to a multivalued operator $H := \left\{(f, H_{f,\varphi}) : f \in C^2(\mathbb{R}^d),\ H_{f,\varphi} \in C(\mathbb{R}^d \times E') \text{ and } \varphi \in C^2(E')\right\}$,
2. there exists $\tilde{\varphi}$ such that $H_{f,\tilde{\varphi}}(x,z) = H_{\tilde{\varphi}}(x,p,z) = \mathcal{H}(x,p)$ for all $z \in E'$ and $p = \nabla f$,
3. the comparison principle for $(1 - \lambda H)u = h$ holds,
4. Y^ε verifies the exponential tightness property,
5. the rate function (2.2.2) can be represented in the following integral form

$$\mathcal{I}(x) = \mathcal{I}_0(x(0)) + \int_0^\infty \mathcal{L}\left(x(t), \dot{x}(t)\right)\, dt$$

with $\mathcal{L}(x,v) = \sup_{p \in \mathbb{R}^d} \left[p \cdot v - \mathcal{H}(x,p)\right]$ is the Legendre transform of $\mathcal{H}(x,p)$ in (3.1.4).

We prove the above claims respectively in Propositions 3.1.12, 3.1.7, 3.1.25, 3.1.31, 3.1.34 in the following subsections. Then, once the above facts are proved, we can apply Theorem 2.2.12 and the required large deviation property follows. □

3.1.3 *The convergence of generators and an eigenvalue problem*

The first step of the proof of large deviations is based on operator convergence. Since the process and its limit do not live in the same space, we cannot work with the usual definition. In the following, we introduce a new definition of limit for functions and multivalued operator on different spaces.

Definition 3.1.10. Let $f_\varepsilon \in C(\mathbb{R}^d \times \{1, \dots, J\})$ and $f \in C^2(\mathbb{R}^d)$. We say that $LIM f_n = f$ if

$$\|f_\varepsilon - f \circ \eta_\varepsilon\|_{\mathbb{R}^d \times \{1,\dots,J\}} = \sup_{\mathbb{R}^d \times \{1,\dots,J\}} |f_\varepsilon - f \circ \eta_\varepsilon| \to 0 \quad \text{as } \varepsilon \to 0,$$

where $\eta_\varepsilon : \mathbb{R}^d \times \{1, \dots, J\} \to \mathbb{R}^d$ is the projection

$$\eta_\varepsilon(x, i) = x.$$

Definition 3.1.11 (extended limit of multivalued operators)**.** Let $H_\varepsilon \subseteq C(\mathbb{R}^d \times \{1,\dots,J\})$. Define $ex - LIMH_\varepsilon$ as the set

$$
\begin{aligned}
&ex - LIMH_\varepsilon = \\
&= \Big\{ (f,H) \in C^2(\mathbb{R}^d) \times C(\mathbb{R}^d \times \mathbb{T}^d \times \{1,\dots,J\}) \,|\, \exists f_\varepsilon \in D(H_\varepsilon) \,:\, f = LIMf_\varepsilon \\
&\qquad\qquad\qquad\qquad \text{and } \|H \circ \eta'_\varepsilon - H_\varepsilon f_\varepsilon\|_{\mathbb{R}^d \times \{1,\dots,J\}} \to 0 \Big\},
\end{aligned}
$$

where $\eta'_\varepsilon : \mathbb{R}^d \times \{1,\dots,J\} \to \mathbb{R}^d \times \mathbb{T}^d \times \{1,\dots,J\}$ is the function $\eta'_\varepsilon(x,i) = \left(x, \left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}, i\right)$.

The following basic example gives the idea of the intuition behind the definitions above.

Example: Let $H_\varepsilon f(x,i) = \nabla f(x) + \varepsilon \Delta f(x)$. Then, for every $f \in C^2(\mathbb{R}^d)$ and $\varphi \in C^2(\mathbb{T}^d)$, we define

$$
f_\varepsilon(x,i) = f(x) + \varepsilon \varphi\left(\frac{x}{\varepsilon}, i\right) \qquad \text{and} \qquad H(x,y,i) = \Delta \varphi^i(y)
$$

. Then, $(f,H) \in ex - LIMH_\varepsilon$.

Proposition 3.1.12 (Convergence of nonlinear generator)**.** *Let $E = \mathbb{R}^d \times \{1,\dots,J\}$ and let $(Y_t^\varepsilon, \bar{I}_t^\varepsilon)$ be the rescaled Markov process with generator A_ε from (3.1.3) and let H_ε be the nonlinear generators defined in Definition 2.2.5. Then, the multivalued operator $H \subseteq C(\mathbb{R}^d) \times C(\mathbb{R}^d \times \mathbb{T}^d \times \{1,\dots,J\})$ given by*

$$
H := \left\{ (f, H_{f,\varphi}) \,:\, f \in C^2(\mathbb{R}^d), H_{f,\varphi} \in C(\mathbb{R}^d \times E') \text{ and } \varphi \in C^2(E') \right\},
$$

where the images $H_{f,\varphi} : \mathbb{R}^d \times \mathbb{T}^d \times \{1,\dots,J\} \to \mathbb{R}$ are

$$
\begin{aligned}
H_{f,\varphi}(x,y,i) := \frac{1}{2}\Delta_y \varphi^i(y) + \frac{1}{2}\left|\nabla f(x) + \nabla_y \varphi^i(y)\right|^2 - \nabla_x \psi^i(x,y)(\nabla f(x) + \nabla_y \varphi^i(y)) \\
+ \sum_{j=1}^J r_{ij}(x,y) \left[e^{\varphi(y,j) - \varphi(y,i)} - 1 \right],
\end{aligned}
$$

is such that $H \subseteq ex - LIMH_\varepsilon$. Moreover, for all φ parametrising the images we have a map $H_\varphi : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{T}^d \times \{1,\dots,J\} \to \mathbb{R}$ such that for all $f \in \mathcal{D}(H)$ and any $x \in \mathbb{R}^d$, the images $H_{f,\varphi}$ of H are given by

$$
H_{f,\varphi}(x,z') = H_\varphi(x, \nabla f(x), z'), \quad \textit{for all } z' \in \mathbb{T}^d \times \{1,\dots,J\}.
$$

Proof. We want to prove that H_ε converges to H in terms of Definition 3.1.11. With this aim, note that, by the definitions of A_ε and H_ε, we have

$$
\begin{aligned}
H_\varepsilon f(x,i) = \frac{\varepsilon}{2}\Delta_x f^i(x) + \frac{1}{2}\left|\nabla_x f^i(x)\right|^2 - \nabla \psi^i\left(x, \frac{x}{\varepsilon}\right) \nabla_x f^i(x) \\
+ \sum r_{ij}\left(x, \frac{x}{\varepsilon}\right)\left(e^{(f(x,j) - f(x,i))/\varepsilon} - 1\right).
\end{aligned}
$$

Choosing functions $f_\varepsilon(x,i)$ of the form

$$f_\varepsilon(x,i) = f(x) + \varepsilon\,\varphi\left(\left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}, i\right) = f \circ \eta_\varepsilon(x,i) + \varepsilon\,\varphi\left(\left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}, i\right),$$

we obtain,

$$\begin{aligned} H_\varepsilon(f_\varepsilon)(x,i) &= \frac{\varepsilon}{2}\Delta f(x) + \frac{1}{2}\Delta_y \varphi^i\left(\left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}\right) + \frac{1}{2}\left|\nabla f(x) + \nabla_y \varphi^i\left(\left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}\right)\right|^2 \\ -\nabla\psi^i(x,x/\varepsilon)&\left(\nabla f(x) + \nabla_y\varphi^i\left(\left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}\right)\right) + \sum_{j=1}^{J} r_{ij}(x,x/\varepsilon)\left[e^{\varphi([\frac{x}{\varepsilon}]_{\mathbb{Z}^n},j)-\varphi([\frac{x}{\varepsilon}]_{\mathbb{Z}^n},i)} - 1\right], \end{aligned}$$

where ∇_y and Δ_y denote the gradient and Laplacian with respect to the variable $y = x/\varepsilon$. We can conclude that

$$\|f \circ \eta_\varepsilon - f_\varepsilon\|_E = \|f(x) - f_\varepsilon(x,i)\|_E = \varepsilon\|\varphi\|_E \to 0 \text{ as } \varepsilon \to 0,$$

and

$$\begin{aligned} \|H_{f,\varphi} \circ \eta'_\varepsilon - H_\varepsilon f_\varepsilon\|_E &= \sup_{(x,i)\in E}\left|H_{f,\varphi}\left(x, \left[\frac{x}{\varepsilon}\right]_{\mathbb{Z}^n}, i\right) - H_\varepsilon f_\varepsilon(x,i)\right| \\ &= \frac{\varepsilon}{2}\sup_{(x,i)\in E} |\,\Delta f(x)| \xrightarrow{\varepsilon\to 0} 0. \end{aligned}$$

□

Remark 3.1.13. Note that for all $f \in D(H)$ the image H_φ has the representation

$$H_\varphi(x,p,z) = e^{-\varphi(z)}\left[B_{x,p} + V_{x,p} + R_x\right] e^{\varphi}(z)$$

with $p = \nabla f(x)$ and

$$\begin{aligned} (B_{x,p}h)(y,i) &:= \frac{1}{2}\Delta_y h(y,i) + \left(p - \nabla_x\psi^i(x,y)\right)\cdot\nabla_y h(y,i) \\ (V_{x,p}h)(y,i) &:= \left(\frac{1}{2}p^2 - p\cdot\nabla_x\psi^i(x,y)\right)h(y,i), \\ (R_x\,h)(y,i) &:= \sum_{j=1}^{J} r_{ij}(x,y)\left[h(y,j) - h(y,i)\right]. \end{aligned}$$

Proposition 3.1.14 (Existence of an eigenvalue). *Let $E' = \mathbb{T}^d \times \{1,\ldots,J\}$ and let H_φ : $\mathbb{R}^d \times \mathbb{R}^d \times E' \to R$ the images of H given in Proposition 3.1.12. Then, for all $p \in \mathbb{R}^d$ there exists an eigenfunction $g_{x,p} \in C^2(E' \times \{1,\ldots,J\})$ with $g^i_{x,p} > 0$ and an eigenvalue $\lambda_{x,p}$ such that*

$$\left[B_{x,p} + V_{x,p} + R_x\right] g_{x,p} = \lambda_{x,p} g_{x,p}.$$

Proof. We want to solve the following eigenvalue problem

$$[L_{x,p} + R_x]\, g_{x,p} = \lambda_{x,p} g_{x,p} \tag{3.1.6}$$

where $L_{x,p}$ is a diagonal matrix with $(L_{x,p})_{ii} = (B_{x,p})_i + (V_{x,p})_i$ and $(R_x)_{ij} = r_{ij}$ for $i \neq j$ and $(R_x)_{ii} = \sum_{j=1}^{J} r_{ij}$.
Guido Sweers showed (see [Swe92]) that there exists $\gamma_{x,p}$ and $g_{x,p} > 0$ such that

$$[-L_{x,p} - R_x]\, g_{x,p} = \gamma_{x,p} g_{x,p}$$

when $L_{x,p}$ is a diagonal matrix with $(L_{x,p})_{ii}$ of the type $-\Delta + p \cdot \nabla + c$. Hence, in our case, the equality (3.1.6) is verified by taking $\lambda_{x,p} = -\gamma_{x,p}$. □

In the next proposition we prove that the images H_φ depend only on x and p.

Proposition 3.1.15. *Consider the same setting of Proposition 3.1.14 and let $\mathcal{H}(x,p)$ be the constant depending on p and x given in (3.1.4). Then, for all $x, p \in \mathbb{R}^d$ there exist a function $\varphi_{x,p} \in C^2(E')$ such that*

$$H_{\varphi_{x,p}}(x,p,z) = \mathcal{H}(x,p) \qquad \text{for all } z \in E'. \tag{3.1.7}$$

Proof. By Proposition 3.1.14, there exists a function $g_{x,p}$ and a constant $\lambda_{x,p}$ that satisfy the eigenvalue problem for the operator $L_{x,p} + R_{x,p}$ defined in (3.1.6). By the variational representation established by Donsker and Varadhan in [DV75], the eigenvalue is equal to the constant $\mathcal{H}(x,p)$ defined in (3.1.4). Then, equality (3.1.7) follows from Remark 3.1.13 and Proposition 3.1.14 by choosing $\varphi_{x,p} = \log g_{x,p}$. □

3.1.4 *Regularity of the Hamiltonian*

Before proving the comparison principle, we first show that the map $p \mapsto \mathcal{H}(x,p)$, constructed out of the eigenvalue problem in Propositions 3.1.14 and 3.1.15, is convex, coercive and continuous uniformly with respect to x.

Proposition 3.1.16 (Convexity and Coercivity of $\mathcal{H}$)**.** *The map $\mathcal{H} : (x,p) \mapsto \mathcal{H}(x,p)$ in (3.1.4) is convex in p and coercive in p uniformly with respect to x. Precisely,*

$$\lim_{|p| \to \infty} \inf_{x \in K} \mathcal{H}(x,p) = \infty$$

for every K compact set. Moreover, $\mathcal{H}(x,0) = 0$ for all $x \in \mathbb{R}^d$.

Proof. By Proposition 3.1.15 the eigenvalue $\mathcal{H}(x,p)$ admits the representation

$$\begin{aligned}
\mathcal{H}(x,p) &= -\sup_{g>0} \inf_{z' \in E'} \left\{ \frac{1}{g(z')} \left[(-B_{x,p} - V_{x,p} - R_x) g \right](z') \right\} \\
&= \inf_{g>0} \sup_{z' \in E'} \left\{ \frac{1}{g(z')} \left[(B_{x,p} + V_{x,p} + R_x) g \right](z') \right\} \\
&= \inf_{\varphi} \sup_{z' \in E'} \left\{ e^{-\varphi(z')} \left[(B_{x,p} + V_{x,p} + R_x) e^{\varphi} \right](z') \right\} =: \inf_{\varphi} \sup_{z' \in E'} F(x,p,\varphi)(z'),
\end{aligned}$$

where the map F is given by

$$F(x,p,\varphi)(y,i)=\frac{1}{2}\Delta\varphi^i(y)+\frac{1}{2}|\nabla\varphi^i(y)+p|^2-\nabla_x\psi^i(x,y)(\nabla\varphi^i(y)+p)\\+\sum_{j=1}^{J}r_{ij}(x,y)\left[e^{\varphi^j(y)-\varphi^i(y)}-1\right].$$

Note that F is jointly convex in p and φ. By Proposition 3.1.15, for every x,p there exists $\varphi_{x,p}$ such that equality holds, i.e. for any $z'\in E'$, we have $\mathcal{H}(x,p)=F(x,p,\varphi_{x,p})(z')$. Therefore, we obtain for $\xi\in[0,1]$ and any $p_1,p_2\in\mathbb{R}^d$ with corresponding eigenfunctions e^{φ_1} and e^{φ_2} that

$$\begin{aligned}\mathcal{H}(x,\xi p_1+(1-\xi)p_2)&=\inf_{\varphi}\sup_{E'}F\left(x,\xi p_1+(1-\xi)p_2,\varphi\right)\\&\le\sup_{E'}F\left(x,\xi p_1+(1-\xi)p_2,\xi\varphi_1+(1-\xi)\varphi_2\right)\\&\le\sup_{E'}\left[\xi F(x,p_1,\varphi_1)+(1-\xi)F(x,p_2,\varphi_2)\right]\\&\le\xi\sup_{E'}F(x,p_1,\varphi_1)+(1-\xi)\sup_{E'}F(x,p_2,\varphi_2)\\&=\xi\mathcal{H}(x,p_1)+(1-\xi)\mathcal{H}(x,p_2).\end{aligned}$$

Regarding coercivity of $\mathcal{H}(x,p)$, we isolate the p^2 term in $V_{x,p}$, to obtain

$$\mathcal{H}(x,p)=\frac{p^2}{2}+\inf_{\varphi}\sup_{E'}\left\{e^{-\varphi}\left[B_{x,p}-p\cdot\nabla_x\psi+R_x\right]e^{\varphi}\right\}.$$

Any $\varphi\in C^2(E')$ admits a minimum (y_m,i_m) on the compact set E', and with the thereby obtained uniform lower bound

$$\begin{aligned}&\Gamma(x,p,\varphi)=\sup_{E'}\left\{e^{-\varphi(z_m)}\left[B_{x,p}-p\cdot\nabla_x\psi+R_x\right]e^{\varphi(z_m)}\right\}\\&\ge e^{-\varphi(z_m)}\left[B_{x,p}-p\cdot\nabla_x\psi+R_x\right]e^{\varphi(z_m)}\\&=\underbrace{\frac{1}{2}\Delta_y\varphi(y_m,i_m)}_{\ge0}+\frac{1}{2}|\underbrace{\nabla_y\varphi(y_m,i_m)}_{=0}|^2+(p-\nabla_x\psi^{i_m}(x,y_m))\cdot\underbrace{\nabla_y\varphi(y_m,i_m)}_{=0}\\&+\sum_{j\neq i}r_{ij}(x,y_m)\underbrace{\left[e^{\varphi(y_m,j)-\varphi(y_m,i_m)}-1\right]}_{\ge0}-p\cdot\nabla_x\psi^{i_m}(x,y_m)\ge-p\cdot\nabla_x\psi^{i_m}(x,y_m).\end{aligned}$$

Using the lower bound $\Gamma(x,p,\varphi)\ge-p\cdot\nabla_x\psi^{i_m}(x,y_m)\ge\inf_{E'}(-p\cdot\nabla_x\psi)$, it follows that, if K is a compact set

$$\begin{aligned}\inf_{x\in K}\mathcal{H}(x,p)&\ge\frac{p^2}{2}-\sup_{x\in K}\sup_{E'}(p\cdot\nabla_x\psi^i(x,y))\\&\ge\frac{1}{4}p^2-\sup_{x\in K}\sup_{E'}|\nabla_x\psi^i(x,y)|^2\xrightarrow{|p|\to\infty}\infty.\end{aligned}$$

Regarding $\mathcal{H}(x,0) = 0$, note that $\Gamma(x,0,\varphi) \leq 0$ for all x and φ. Then we have the first inequality $\mathcal{H}(x,0) \geq \inf_\varphi \Gamma(x,0,\varphi) \geq 0$. For the opposite inequality we choose the function $\varphi = (1,\ldots,1)$ in the representation of $\mathcal{H}$. □

Proposition 3.1.17 (Continuity of $\mathcal{H}$). *The map $\mathcal{H} : (x,p) \in \mathbb{R}^d \times \mathbb{R}^d \to \mathcal{H}(x,p) \in \mathbb{R}$ is continuous.*

We will prove the continuity of $\mathcal{H}$ by showing that it is lower and upper semicontinuous. For that, we need the following auxiliary results. In particular, for the lower semicontinuity we will make use of the Γ–convergence in the sense expressed in the following lemma in which we prove that property in a general context. Later, we will use it for $\mathcal{J}(x,p,\theta) = I_{x,p}(\theta)$.

Lemma 3.1.18 (Γ-convergence). *Given two sets $U, V \subseteq \mathbb{R}^d$ and a constant $M \geq 0$ we define $\Theta_{U,V,M}$ as*

$$\Theta_{U,V,M} = \bigcup_{x\in U, p\in V} \{\theta \in \Theta | \mathcal{J}(x,p,\theta) \leq M\}.$$

Let $\mathcal{J} : \mathbb{R}^d \times \mathbb{R}^d \times \Theta \to [0,\infty]$ satisfy the following assumptions:

(i) The map $(x,p,\theta) \mapsto \mathcal{J}(x,p,\theta)$ is lower semi-continuous on $\mathbb{R}^d \times \mathbb{R}^d \times \Theta$.

(ii) For every x and p fixed and $M \geq 0$, there exist U_x and U_p open and bounded neighbourhoods and a constant M' such that

$$\mathcal{J}(y,q,\theta) \leq M' \qquad \text{for all } y \in U_x,\ q \in U_p \text{ and } \theta \in \Theta_{\{x\},\{p\},M}$$

.

(iii) For all compact sets $K_1 \subseteq \mathbb{R}^d$ and $K_2 \subseteq \mathbb{R}^d$ and each $M \geq 0$ the collection of functions $\{\mathcal{J}(\cdot,\cdot,\theta)\}_{\theta\in\Theta_{K_1,K_2,M}}$ is equi-continuous.

Then if $x_n \to x$ and $p_n \to p$, the functionals $\mathcal{J}_n$ defined by

$$\mathcal{J}_n(\theta) := \mathcal{J}(x_n,p_n,\theta)$$

converge in the Γ-sense to $\mathcal{J}_\infty(\theta) := \mathcal{J}(x,p,\theta)$. That is:

1. *If $x_n \to x$, $p_n \to p$ and $\theta_n \to \theta$, then $\liminf_{n\to\infty} \mathcal{J}(x_n,p_n,\theta_n) \geq \mathcal{J}(x,p,\theta)$,*
2. *For $x_n \to x$ and $p_n \to p$ and all $\theta \in \Theta$ there are controls $\theta_n \in \Theta$ such that $\theta_n \to \theta$ and $\limsup_{n\to\infty} \mathcal{J}(x_n,p_n,\theta_n) \leq \mathcal{J}(x,p,\theta)$.*

Proof. Let $x_n \to x$ and $p_n \to p$ in $\mathbb{R}^d$. If $\theta_n \to \theta$, then by lower semicontinuity (i),

$$\liminf_{n\to\infty} \mathcal{J}(x_n,p_n,\theta_n) \geq \mathcal{J}(x,p,\theta).$$

For the lim-sup bound, let $\theta \in \Theta$. If $\mathcal{J}(x,p,\theta) = \infty$, there is nothing to prove. Thus suppose that $\mathcal{J}(x,p,\theta)$ is finite, i.e., $\theta \in \Theta_{\{x\},\{p\},M}$ for some M. Then, by (ii), there exist a bounded neighborhood U_x of x, a bounded neighborhood U_p of p and a constant M' such that for any $y \in U_x$ and $q \in U_p$,

$$\mathcal{J}(y,q,\theta) \leq M'.$$

Since $x_n \to x$ and $p_n \to p$, the sequences x_n and p_n are, for n large, contained in U_x and U_p, respectively. Taking the constant sequence $\theta_n := \theta$, we thus get that $\mathcal{J}(x_n, p_n, \theta_n) \leq M'$ for all n large enough. By (iii), the family of functions $\{\mathcal{J}(\cdot, \cdot, \theta)\}_{\theta \in \Theta_{\bar{U}_x, \bar{U}_p, M'}}$ is equicontinuous, and hence

$$\lim_{n\to\infty} |\mathcal{J}(x_n, p_n, \theta_n) - \mathcal{J}(x, p, \theta)| \leq 0,$$

and the lim-sup bound follows. □

We can now prove that the function $I_{x,p}$ in (3.1.4) is Γ-convergent.

Proposition 3.1.19 (Γ-convergence of $I_{x,p}$). *Let $I_{x,p} : \Theta \to [0, \infty]$ the function defined in* (3.1.4)*. If $x_n \to x$ and $p_n \to p$, the functionals $I_n(\theta) := I_{x_n,p_n}(\theta)$ converge in the Γ-sense to $I_\infty(\theta) := I_{x,p}(\theta)$.*

Proof. Using Lemma 3.1.18, we only need to prove that $I_{x,p}$ verifies the assumptions.

Assumption (i). For any fixed function $u \in \mathcal{D}(L_{x,p})$ such that $u > 0$, the function $(L_{x,p}u/u)$ is continuous. Thus, for any such fixed $u > 0$ it follows that

$$(x, p, \theta) \mapsto \int_{E'} \frac{L_{x,p}u}{u}\, d\theta$$

is continuous on $\mathbb{R}^d \times \mathbb{R}^d \times \Theta$. As a consequence $I(x, p, \theta)$ is lower semicontinuous as the supremum over continuous functions.

Assumption (ii). Fix x, p and $M \geq 0$. Let $\theta \in \Theta_{x,p,M}$. Then, $I_{x,p}(\theta) = I(x, p, \theta) \leq M$. It follows from [Pin07, Theorem 3] that the density $\frac{d\theta}{dz}$ exists. Moreover, by the same theorem, for all y and q there exist constants $c_1(y, q), c_2(y, q)$ positive, depending continuously on y and q, but not on θ, such that

$$I_{y,q}(\theta) \leq c_1(y, q) \int_{E'} |\nabla g_\theta|^2\, dz + c_2(y, q),$$

where $g_\theta = (d\theta/dz)^{1/2}$ is the square root of the Radon–Nykodym derivative. As the dependence is continuous in y and q, we can find two open neighbourhoods, $U \subseteq \mathbb{R}^d$ of x and $V \subseteq \mathbb{R}^d$ of p, such that there exist constants c_1, c_2 positive, that do not depend on θ, such that for any $y \in U$ and $q \in V$:

$$I_{y,q}(\theta) \leq c_1 \int_{E'} |\nabla g_\theta|^2\, dz + c_2 := M',$$

obtaining then (ii).

Assumption (iii). By the continuity of r_{ij} and ψ, assumption (iii) follows from Theorem 4 of [Pin07].

□

The following technical lemma will give us the upper semi-continuity of $\mathcal{H}$.

Lemma 3.1.20 (Lemma 17.30 in [AB06]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces. Let $\phi : \mathcal{X} \to \mathcal{K}(\mathcal{Y})$, where $\mathcal{K}(\mathcal{Y})$ is the space of non-empty compact subsets of $\mathcal{Y}$. Suppose that ϕ is upper hemi-continuous, that is if $x_n \to x$ and $y_n \to y$ and $y_n \in \phi(x_n)$, then $y \in \phi(x)$. Let f : Graph(ϕ) $\to \mathbb{R}$ be upper semi-continuous. Then the map $m(x) = \sup_{y\in\phi(x)} f(x,y)$ is upper semi-continuous.*

We can finally prove the continuity of $\mathcal{H}(x, p)$.

Proof of Proposition 3.1.17. We have already showed that $I_{x,p}(\mu)$ is lower semicontinuous and, since $V_p(x, i)$ is continuous and bounded, $\int_{E'} V_{x,p}\, d\mu$ is continuous. Then, $f(x, p, \mu) := \int_{E'} V_{x,p}\, d\mu - I_{x,p}(\mu)$ is upper semi-continuous.
Let $x, p \in \mathbb{R}^d$. We know, by Proposition 3.1.38 in the appendix, that there exists a unique stationary measure $\theta^0_{x,p}$ such that for all $g \in D(L_{x,p})$,

$$\int_{E'} L_{x,p} g(z, i) d\theta^0_{x,p} = 0. \tag{3.1.8}$$

Let $L^\lambda_{x,p} = \lambda(\lambda - L_{x,p})^{-1} L_{x,p}$ the Hille-Yosida approximation of $L_{x,p}$. Then we have

$$\begin{aligned}
&-\int_{E'} \frac{L_{x,p} u}{u} d\theta^0_{x,p} = -\int_{E'} \frac{L^\lambda_{x,p} u}{u} d\theta^0_{x,p} + \int_{E'} \frac{\left(L^\lambda_{x,p} - L_{x,p}\right) u}{u} d\theta^0_{x,p} \\
&\leq -\int_{E'} \frac{L^\lambda_{x,p} u}{u} d\theta^0_{x,p} + \frac{1}{\inf_{E'} u} \|(L^\lambda_{x,p} - L_{x,p}) u\|_{E'} \\
&\leq -\int_{E'} L^\lambda_{x,p} \log u \, d\theta^0_{x,p} + o(1).
\end{aligned}$$

Sending $\lambda \to 0$ and using (3.1.8) we have that $I_{x,p}(\theta^0_{x,p}) = 0$. Then, $\mathcal{H}(x, p) \geq \int_{E'} V_{x,p} d\theta^0_{x,p}$. Thus, it suffices to restrict the supremum over $\theta \in \phi(x, p)$ where

$$\phi(x, p) := \left\{ \theta \in \mathcal{P}(E') \,\middle|\, I_{x,p}(\theta) \leq 2\|\Pi(x, p, \cdot)\|_{\mathcal{P}(E')} \right\},$$

where $\|\cdot\|_{\mathcal{P}(E')}$ denotes the supremum norm on $\mathcal{P}(E')$ and we called for simplicity $\Pi(x, p, \theta) = \int_{E'} V_{x,p} d\theta$.
Note that $\|\Pi(x, p, \theta)\|_{\mathcal{P}(E')} < \infty$ by definition of $V_{x,p}$. It follows that

$$\mathcal{H}(x, p) = \sup_{\theta\in\phi(x,p)} \left[\int_{E'} V_{x,p}\, d\mu - I_{x,p}(\mu) \right].$$

$\phi(x, p)$ is non-empty as $\theta^0_{x,p} \in \phi(x, p)$ and it is compact because any closed subset of $\mathcal{P}(E')$ is compact. We are left to show that ϕ is upper hemi-continuous. Let $(x_n, p_n, \theta_n) \to (x, p, \theta)$ with $\theta_n \in \phi(x_n, p_n)$. We establish that $\theta \in \phi(x, p)$. By the lower semi-continuity of I and the definition of ϕ we find

$$I_{x,p}(\theta) \leq \liminf_n I_{x_n,p_n}(\theta_n) \leq \liminf_n 2\|\Pi(x_n, p_n, \cdot)\|_{\mathcal{P}(E')} = 2\|\Pi(x, p, \cdot)\|_{\mathcal{P}(E')}$$

which implies indeed that $\theta \in \phi(x,p)$. Thus, upper semi-continuity follows by an application of Lemma 3.1.20.

We proceed with proving lower semi-continuity of $\mathcal{H}$. Suppose that $(x_n,p_n) \to (x,p)$, we prove that $\liminf_n \mathcal{H}(x_n,p_n) \geq \mathcal{H}(x,p)$. Let θ be the measure such that $\mathcal{H}(x,p) = \Pi(x,p,\theta) - I_{x,p}(\theta)$.

By Proposition 3.1.19, there are θ_n such that $\theta_n \to \theta$ and $\limsup_n I_{x_n,p_n}(\theta_n) \leq I_{x,p}(\theta)$. Moreover, $\Pi(x_n,p_n,\theta_n)$ converges to $\Pi(x,p,\theta)$ by continuity. Therefore,

$$\begin{aligned}\liminf_{n\to\infty} \mathcal{H}(x_n,p_n) &\geq \liminf_{n\to\infty} \left[\Pi(x_n,p_n,\theta_n) - I_{x_n,p_n}(\theta_n)\right] \\ &\geq \liminf_{n\to\infty} \Pi(x_n,p_n,\theta_n) - \limsup_{n\to\infty} I_{x_n,p_n}(\theta_n) \\ &\geq \Pi(x,p,\theta) - I_{x,p}(\theta) = \mathcal{H}(x,p),\end{aligned}$$

establishing that $\mathcal{H}$ is lower semi-continuous. □

3.1.5 *Comparison principle*

In this section we prove the comparison principle for the Hamilton–Jacobi equation in terms of H by using a similar argument presented in Chapter 1.2.2, that is, relating it to a set of Hamilton-Jacobi equations constructed from $\mathcal{H}$ (Figure 8). We introduce the operators $H_\dagger, H_\ddagger$ and H_1, H_2. In both cases, the new Hamiltonians will serve as natural upper and lower bounds for $\mathbf{H}f(x) = \mathcal{H}(x,\nabla f(x))$ and H respectively, where $\mathcal{H}$ and H are the operators introduced in Propositions 3.1.15 and 3.1.12. These new Hamiltonians are defined in terms of a containment function Υ, which allows us to restrict our analysis to compact sets. Here we give the rigorous definition.

Definition 3.1.21 (Containment function)**.** A function $\Upsilon : \mathbb{R}^d \to [0,\infty)$ is a containment function for $V_{x,p}$ in (3.1.4), if $\Upsilon \in C^1(\mathbb{R}^d)$ and it is such that

- Υ has compact sub-level sets, i.e. for every $c \geq 0$ the set $\{x | \Upsilon(x) \leq c\}$ is compact ;
- $\sup_{x\in\mathbb{R}^d, z\in E'} V_{x,\nabla\Upsilon(x)}(z) < \infty$.

Lemma 3.1.22. *The function* $\Upsilon(x) = \frac{1}{2}\log\left(1+|x|^2\right)$ *is a containment function for* $V_{x,p}$*.*

Proof. Firstly note that Υ has compact sub-level sets. Regarding the second property, by the definition of $V_{x,p}$, we have for every $x \in \mathbb{R}^d$ and $z = (y,i) \in \mathbb{T}^d \times \{1,\ldots,J\}$,

$$V_{x,\nabla\Upsilon(x)}(y,i) = \frac{x^2}{2(1+|x|^2)^2} - \nabla_x\psi^i(x,y)\frac{x}{1+|x|^2}.$$

Recalling that ψ grows at most linearly in x, we can conclude that $\sup_{x,z} V_{x,\nabla\Upsilon}(z) < \infty$. □

Using the above lemma we are now able to define the auxiliary operators in terms of Υ. In the following we will denote by $C_l^\infty(E)$ the set of smooth functions on E that have a lower bound and by $C_u^\infty(E)$ the set of smooth functions on E that have an upper bound.

Definition 3.1.23. Fix $\eta \in (0,1)$ and given $\Upsilon(x) = \frac{1}{2}\log\left(1+|x|^2\right)$, $C_\Upsilon := \sup_{x,z} V_{x,\nabla\Upsilon(x)}(z)$ and $\mathbf{H}f(x) = \mathcal{H}(x, \nabla f(x))$, we define

- For $f \in C_l^\infty(E)$,

$$f_\dagger^\eta := (1-\eta)f + \eta\Upsilon,$$
$$H_{\dagger,f}^\eta(x) := (1-\eta)\mathbf{H}f(x) + \eta C_\Upsilon,$$

and set

$$H_\dagger := \left\{ (f_\dagger^\eta, H_{\dagger,f}^\eta) \,\middle|\, f \in C_l^\infty(E), \eta \in (0,1) \right\}.$$

- For $f \in C_u^\infty(E)$,

$$f_\ddagger^\eta := (1+\eta)f - \eta\Upsilon,$$
$$H_{\ddagger,f}^\eta(x) := (1+\eta)\mathbf{H}f(x) - \eta C_\Upsilon,$$

and set

$$H_\ddagger := \left\{ (f_\ddagger^\eta, H_{\ddagger,f}^\eta) \,\middle|\, f \in C_u^\infty(E), \eta \in (0,1) \right\}.$$

Definition 3.1.24. Fix $\eta \in (0,1)$ and given $\Upsilon(x) = \frac{1}{2}\log\left(1+|x|^2\right)$, $C_\Upsilon := \sup_{x,z} V_{x,\nabla\Upsilon(x)}(z)$ and $\mathbf{H}f(x) = \mathcal{H}(x, \nabla f(x))$, we define

- For $f \in C_l^\infty(E)$, $\varphi \in C^2(E')$, $\eta \in (0,1)$ set

$$f_1^\eta := (1-\eta)f + \eta\Upsilon,$$
$$H_{1,f,\varphi}^\eta(x,z) := (1-\eta)H_{f,\varphi}(x,z) + \eta C_\Upsilon,$$

and set

$$H_1 := \left\{ (f_1^\eta, H_{1,f,\varphi}^\eta) \,\middle|\, f \in C_l^\infty(E), \varphi \in C^2(E'), \eta \in (0,1) \right\}.$$

- For $f \in C_u^\infty(E)$, $\varphi \in C^2(E')$, $\eta \in (0,1)$ set

$$f_2^\eta := (1+\eta)f - \eta\Upsilon,$$
$$H_{2,f,\varphi}^\eta(x,z) := (1+\eta)H_{f,\varphi}(x,z) - \eta C_\Upsilon,$$

and set

$$H_2 := \left\{ (f_2^\eta, H_{2,f,\varphi}^\eta) \,\middle|\, f \in C_u^\infty(E), \varphi \in C^2(E'), \eta \in (0,1) \right\}.$$

We now prove the comparison principle for $f - \lambda H f = h$ based on the results summarized in Figure 8.

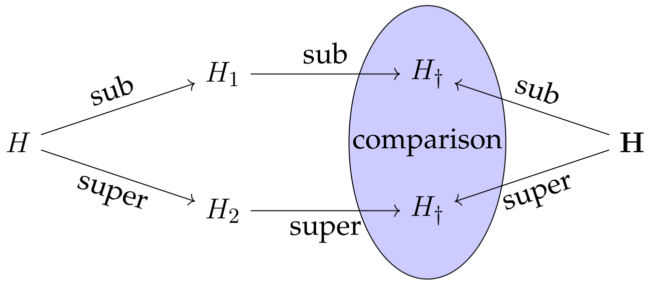

Figure 8: An arrow connecting an operator A with operator B with subscript 'sub' means that viscosity subsolutions of $(1-\lambda A)f = h$ are also viscosity subsolutions of $(1-\lambda B)f = h$. Similarly for arrows with a subscript 'super'. The box around the operators $H_\dagger$ and $H_\ddagger$ indicates that the comparison principle holds for subsolutions of $(1-\lambda H_\dagger)f = h$ and supersolutions of $(1-\lambda H_\ddagger)f = h$.

Theorem 3.1.25 (Comparison principle)**.** *Let $h \in C_b(E)$ and $\lambda > 0$. Let u and v be, respectively, any subsolution and any supersolution to $(1-\lambda H)f = h$. Then we have that*

$$\sup_x u(x) - v(x) \leq 0.$$

Proof. Fix $h \in C_b(E)$ and $\lambda > 0$. Let u be a viscosity subsolution and v be a viscosity supersolution to $(1-\lambda H)f = h$. By Figure 8, the function u is a viscosity subsolution to $(1-\lambda H_\dagger)f = h$ and v is a viscosity supersolution to $(1-\lambda H_\ddagger)f = h$. Hence by the comparison principle for $H_\dagger, H_\dagger$ established in Theorem 3.1.26 below, $\sup_x u(x) - v(x) \leq 0$, which finishes the proof. □

The rest of this subsection is devoted to establishing Figure 8. More precisely, we establish Figure 8 in results 3.1.26, 3.1.27, 3.1.28 and 3.1.33.

The next theorem contains the comparison principle for $H_\dagger$ and $H_\ddagger$. The proof follows the ideas presented in Chapter 1 (see also [BC97] and [CIL92]). In order to be able to use both the subsolution and supersolution properties in the estimate of $\sup_x u(x) - v(x)$, we use the following strategy based on the introduction of double variables.

1. First of all, note that the supremum over x of $u(x) - v(x)$ can be replaced, sending $\varepsilon \to 0$, with the supremum over x and y of the double variables function $u(x) - v(y) - (2\varepsilon)^{-1}(x-y)^2$
2. Once the supremum (x,y) is found, we are able to use the sub-super solution properties in the following way:
   - fixing y and optimising over x, it can be used in the application of the subsolution property of u
   - fixing x and optimising over y, it can be used in the application of the supersolution property of v.

Theorem 3.1.26. *Let $h \in C_b(E)$ and $\lambda > 0$. Let u be any subsolution to $(1 - \lambda H_\dagger) f = h$ and let v be any supersolution to $(1 - \lambda H_\ddagger) f = h$. Then we have that*

$$\sup_x u(x) - v(x) \leq 0.$$

Proof. Following the above steps we define the double variables function

$$\Phi_{\varepsilon,\beta}(x,y) = \frac{u(x)}{1-\beta} - \frac{v(y)}{1+\beta} - \frac{|x-y|^2}{2\varepsilon} - \frac{\beta}{1-\beta}\Upsilon(x) - \frac{\beta}{1+\beta}\Upsilon(y).$$

Note that the containment function Υ is introduced in order to be able to work in a compact set, and the positive constant β will allow us to use the convexity of $\mathcal{H}$. Since $\Phi_{\varepsilon,\beta}$ is upper semicontinuous and $\lim_{|x|+|y|\to\infty} \Phi(x,y) = -\infty$, for every $\varepsilon \in (0,1)$ there exists $(x_\varepsilon, y_\varepsilon)$ such that

$$\Phi_{\varepsilon,\beta}(x_\varepsilon, y_\varepsilon) = \sup_{\mathbb{R}^d \times \mathbb{R}^d} \Phi_{\varepsilon,\beta}(x,y). \tag{3.1.9}$$

Suppose by contradiction that $\delta = u(\tilde{x}) - v(\tilde{x}) > 0$ for some $\tilde{x}$. We choose β such that $\frac{2\beta}{(1-\beta)(1+\beta)}\Upsilon(\tilde{x}) < \delta/2$ and $\frac{2\beta}{1-\beta^2}(||h|| + C_\Upsilon) < \delta/2$. Then,

$$\Phi_{\varepsilon,\beta}(x_\varepsilon, y_\varepsilon) \geq \Phi_{\varepsilon,\beta}(\tilde{x}, \tilde{x}) > \delta - \frac{2\beta}{(1-\beta)(1+\beta)}\Upsilon(\tilde{x}) > \frac{\delta}{2} > 0, \tag{3.1.10}$$

and

$$\frac{\beta}{1-\beta}\Upsilon(x_\varepsilon) + \frac{\beta}{1+\beta}\Upsilon(y_\varepsilon) \leq \sup\left(\frac{u}{1-\beta}\right) + \sup\left(\frac{-v}{1+\beta}\right) < \infty.$$

Therefore there exists $R_\beta > 0$ such that x_ε and y_ε belong to $B(0, R_\beta)$.

Next we observe that by Lemma 3.1 of [CIL87a],

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{\varepsilon} \to 0 \qquad \text{as } \varepsilon \to 0^+,$$

and, as a consequence, $|x_\varepsilon - y_\varepsilon| \to 0$ as $\varepsilon \to 0^+$. Define the functions $\varphi_1^{\varepsilon,\beta} \in D(H_\dagger)$ and $\varphi_2^{\varepsilon,\beta} \in D(H_\ddagger)$ by

$$\varphi_1^{\varepsilon,\beta}(x) = (1-\beta)\left[\frac{v(y_\varepsilon)}{1+\beta} + \frac{|x-y_\varepsilon|^2}{2\varepsilon} + \frac{\beta}{1-\beta}\Upsilon(x) + \frac{\beta}{1+\beta}\Upsilon(y_\varepsilon) + (1-\beta)(x-x_\varepsilon)^2\right]$$

$$\varphi_2^{\varepsilon,\beta}(y) = (1+\beta)\left[\frac{u(x_\varepsilon)}{1-\beta} - \frac{|x_\varepsilon-y|^2}{2\varepsilon} - \frac{\beta}{1-\beta}\Upsilon(x_\varepsilon) - \frac{\beta}{1+\beta}\Upsilon(y) - (1+\beta)(y-y_\varepsilon)^2\right].$$

Using (3.1.9), observe that $u - \varphi_1^{\varepsilon,\beta}$ attains its supremum at $x = x_\varepsilon$, and thus

$$\sup_E (u - \varphi_1^{\varepsilon,\beta}) = (u - \varphi_1^{\varepsilon,\beta})(x_\varepsilon).$$

Moreover, by addition of the $(1-\beta)(x-x_\varepsilon)^2$ term, this supremum is the unique optimizer of $u-\varphi_1^{\varepsilon,\beta}$. Then, by the subsolution and supersolution properties, taking into account Lemma 1.2.6,

$$u(x_\varepsilon)-\lambda\left[(1-\beta)\mathcal{H}\left(x_\varepsilon,\frac{x_\varepsilon-y_\varepsilon}{\varepsilon}\right)+\beta\,C_{\mathrm{Y}}\right]\leq h(x_\varepsilon).$$

With a similar argument for u_2 and φ_2^ε, we obtain by the supersolution inequality that

$$v(y_\varepsilon)-\lambda\left[(1+\beta)\mathcal{H}\left(y_\varepsilon,\frac{x_\varepsilon-y_\varepsilon}{\varepsilon}\right)-\beta C_{\mathrm{Y}}\right]\geq h(y_\varepsilon). \tag{3.1.11}$$

By the coercivity property obtained in Proposition 3.1.16 on page 55 and by the inequality (3.1.11), $p_\varepsilon:=\frac{x_\varepsilon-y_\varepsilon}{\varepsilon}$ is bounded in ε, allowing us to extract a converging subsequence p_{ε_k}.

We conclude that for each β

$$\begin{aligned}
&\liminf_{\varepsilon\to0}\Phi(x_\varepsilon,y_\varepsilon)\\
&\leq\liminf_{\varepsilon\to0}\frac{u(x_\varepsilon)}{1-\beta}-\frac{v(y_\varepsilon)}{1+\beta}\\
&\leq\liminf_{k\to\infty}\lambda\mathcal{H}\left(x_{\varepsilon_k},p_{\varepsilon_k}\right)+\frac{\beta}{1-\beta}C_{\mathrm{Y}}-\lambda\mathcal{H}\left(y_{\varepsilon_k},p_{\varepsilon_k}\right)+\frac{\beta}{1+\beta}C_{\mathrm{Y}}\\
&\qquad+\frac{h(x_{\varepsilon_k})}{1-\beta}-\frac{h(y_{\varepsilon_k})}{1+\beta}\\
&\leq\liminf_{k\to\infty}\lambda\left[\mathcal{H}\left(x_{\varepsilon_k},p_{\varepsilon_k}\right)-\mathcal{H}\left(y_{\varepsilon_k},p_{\varepsilon_k}\right)\right]+\frac{h(x_{\varepsilon_k})-h(y_{\varepsilon_k})}{1-\beta^2}+\frac{2\beta}{1-\beta^2}\left(||h||+C_{\mathrm{Y}}\right)\\
&\leq\frac{2\beta}{1-\beta^2}\left(||h||+C_{\mathrm{Y}}\right).
\end{aligned}$$

As β is chosen such that $\frac{2\beta}{1-\beta^2}\left(||h||+C_{\mathrm{Y}}\right)<\delta/2$, we obtain a contradiction with (3.1.10), establishing the comparison principle. □

Below, we complete the figure by proving the left-hand side of Figure 8.

Lemma 3.1.27. *For all $h\in C(\mathbb{R}^d)$ and $\lambda>0$, viscosity subsolutions of $(1-\lambda H)f=h$ are viscosity subsolutions of $(1-\lambda H_1)f=h$, and viscosity supersolutions of $(1-\lambda H)f=h$ are viscosity supersolutions of $(1-\lambda H_2)f=h$.*

Proof. Fix $\lambda>0$ and $h\in C_b(E)$. Let u be a subsolution to $(1-\lambda H)f=h$. We prove it is also a subsolution to $(1-\lambda H_1)f=h$. Fix $\eta\in(0,1)$, $\varphi\in C^2(E')$ and $f\in C_l^\infty(E)$, so that $(f_1^\eta,H_{1,f,\varphi}^\eta)\in H_1$ with f_1^η and $H_{1,f,\varphi}^\eta$ as in Definition 3.1.24. We will prove that there are (x_n,z_n) such that

$$\lim_n u(x_n)-f_1^\eta(x_n)=\sup_x u(x)-f_1^\eta(x), \tag{3.1.12}$$

$$\limsup_n u(x_n)-\lambda H_{1,f,\varphi}^\eta(x_n,z_n)-h(x_n)\leq0. \tag{3.1.13}$$

Given $M := \eta^{-1} \sup_y u(y) - (1-\eta) f(y) < \infty$, as u is bounded and $f \in C_l^\infty(E)$, we have that the sequence x_n along which the limit in (3.1.12) is attained, is contained in the compact set $K := \{x | \Upsilon(x) \leq M\}$. We define $\gamma : \mathbb{R} \to \mathbb{R}$ as a smooth increasing function such that

$$\gamma(r) = \begin{cases} r & \text{if } r \leq M, \\ M+1 & \text{if } r \geq M+2. \end{cases}$$

Denote by f_η the function on E defined by

$$f_\eta(x) = \gamma((1-\eta) f(x) + \eta \Upsilon(x)) = \gamma(f_1^\eta(x)).$$

By construction, f_η is smooth and constant outside a compact set and thus lies in $\mathcal{D}(H)$. We conclude that $(f_\eta, H_{f_\eta,(1-\eta)\varphi}) \in H$. As u is a viscosity subsolution for $(1-\lambda H)u = h$, there exist $x_n \in E$ and $z_n \in E'$ with

$$\begin{aligned} &\lim_n u(x_n) - f_\eta(x_n) = \sup_x u(x) - f_\eta(x), \\ &\limsup_n u(x_n) - \lambda H_{f_\eta,(1-\eta)\varphi}(x_n, z_n) - h(x_n) \leq 0. \end{aligned} \tag{3.1.14}$$

Since f_1^η equals f_η in $K = \{x | \Upsilon(x) \leq M\}$, we also have that

$$\lim_n u(x_n) - f_1^\eta(x_n) = \sup_x u(x) - f_1^\eta(x),$$

establishing (3.1.12). Convexity of $H_{f,\varphi}(x,z) = H_\varphi(x, \nabla f(x), z)$ in p and φ yields for arbitrary (x,z) the elementary estimate

$$\begin{aligned} H_{f_\eta,(1-\eta)\varphi}(x,z) =& H_{(1-\eta)\varphi}(x, (1-\eta)\nabla f(x) + \eta \nabla \Upsilon(x), z) \\ &\leq (1-\eta) H_\varphi(x, \nabla f(x), z) + \eta H_0(x, \nabla \Upsilon(x), z) \\ &= (1-\eta) H_\varphi(x, \nabla f(x), z) + \eta V_{x, \nabla \Upsilon(x)}(z) \\ &\leq H_{1,f,\varphi}^\eta(x,z). \end{aligned}$$

Combining the above inequality with (3.1.14), we have

$$\begin{aligned} &\limsup_n u(x_n) - \lambda H_{1,f,\varphi}^\eta(x,z) - h(x_n) \\ &\leq \limsup_n u(x_n) - \lambda H_{f_\eta,(1-\eta)\varphi}(x_n, z_n) - h(x_n) \\ &\leq 0, \end{aligned}$$

establishing (3.1.13). The supersolution statement follows in the same way. □

Lemma 3.1.28. *Fix $\lambda > 0$ and $h \in C_b(E)$.*

(a) Every subsolution to $(1-\lambda H_1)f = h$ is also a subsolution to $(1-\lambda H_\dagger)f = h$.

(b) Every supersolution to $(1-\lambda H_1)f = h$ is also a supersolution to $(1-\lambda H_\ddagger)f = h$.

The definition of viscosity solutions, Definition 1.2.5, is written down in terms of the existence of a sequence of points that maximizes $u - f$ or minimizes $v - f$. Lemma 1.2.6, instead, gives, for a class of test functions, a definition in terms of a point that maximizes $u - f$ or minimize $v - f$. However, to prove the lemma above, we would like to have the subsolution and supersolution inequalities for any point that maximizes or minimizes the difference. This is achieved by the following auxiliary lemma.

Lemma 3.1.29. *Fix $\lambda > 0$ and $h \in C_b(E)$.*

(a) Let u be a subsolution to $(1 - \lambda H_1)f = h$, then for all $(f, g) \in H_1$ and $x_0 \in E$ such that

$$u(x_0) - f(x_0) = \sup_x u(x) - f(x)$$

there exists a $z \in E'$ such that

$$u(x_0) - \lambda g(x_0, z) \leq h(x_0).$$

(b) Let v be a supersolution to $(1 - \lambda H_2)f = h$, then for all $(f, g) \in H_2$ and $x_0 \in E$ such that

$$v(x_0) - f(x_0) = \inf_x v(x) - f(x)$$

there exists a $z \in E'$ such that

$$v(x_0) - \lambda g(x_0, z) \geq h(x_0).$$

For a proof of the above Lemma see Lemma 5.7 of [KS20].

Proof of Lemma 3.1.28. We only prove the subsolution statement. Fix $\lambda > 0$ and $h \in C_b(E)$. Let u be a subsolution of $(1 - \lambda H_1)f = h$. We prove that it is also a subsolution of $(1 - \lambda H_\dagger)f = h$. Let $f_1^\eta = (1 - \eta)f + \eta \Upsilon \in \mathcal{D}(H_1)$ and let x_0 be such that

$$u(x_0) - f_1^\eta(x_0) = \sup_x u(x) - f_1^\eta(x).$$

For each $\delta > 0$, since $\mathcal{H}(x, p)$ is a principal eigenvalue for $L_{x,p} + R_x$ (as remarked in Proposition 3.1.15, there exists a function g such that

$$\mathcal{H}(x_0, \nabla f(x_0)) = g^{-1}\left(L_{x_0, \nabla f(x_0)} + R_{x_0}\right) g. \tag{3.1.15}$$

As

$$\left(f_1^\eta, (1 - \eta) g^{-1}\left(L_{x_0, \nabla f(x_0)} + R_{x_0}\right) g + \eta C_\Upsilon\right) \in H_1,$$

we find by the subsolution property of u and that there exists z such that

$$\begin{aligned} h(x_0) &\geq u(x_0) - \lambda\left((1 - \eta) g^{-1}\left(L_{x_0, \nabla f(x_0)} + R_{x_0}\right) g + \eta C_\Upsilon\right) \\ &= u(x_0) - \lambda\left((1 - \eta)\mathcal{H}(x_0, \nabla f(x_0)) + \eta C_\Upsilon\right) \end{aligned}$$

where the second inequality follows by (3.1.15) and it establishes that u is a subsolution for $(1 - \lambda H_\dagger)f = h$. □

We conclude this subsection proving the right part of Figure 8.

Proposition 3.1.30. *Let the map $\mathcal{H} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the eigenvalue (3.1.4) and let $\mathbf{H} : \mathcal{D}(\mathbf{H}) \subseteq C^1(\mathbb{R}^d) \to C(\mathbb{R}^d)$ be the operator $\mathbf{H}f(x) := \mathcal{H}(x, \nabla f(x))$. Then, for all $\lambda > 0$ and $h \in C(\mathbb{R}^d)$, every viscosity subsolution of $(1 - \lambda\mathbf{H})f = h$ is also a viscosity subsolutions of $(1 - \lambda H_\dagger)f = h$ and every viscosity supersolution of $(1 - \lambda\mathbf{H})f = h$ is also a viscosity supersolution of $(1 - \lambda H_\ddagger)f = h$.*

Proof. Fix $\lambda > 0$ and $h \in C_b(E)$. Let u be a subsolution to $(1 - \lambda\mathbf{H})f = h$. We prove it is also a subsolution to $(1 - \lambda H_\dagger)f = h$. Fix $\eta > 0$ and $f \in C_\ell^\infty(E)$ and let $(f_\dagger^\varepsilon, H_{\dagger,f}^\eta) \in H_\dagger$ as in Definition 3.1.23. We will prove that

$$\left(u - f_\dagger^\eta\right)(x) = \sup_{x \in E} \left(u(x) - f_\dagger^\eta(x)\right),$$

implies

$$u(x) - \lambda H_{\dagger,f}^\eta(x) - h(x) \le 0. \tag{3.1.16}$$

As u is a viscosity subsolution for $(1 - \lambda\mathbf{H})f = h$ and $f_\dagger^\eta \in D(\mathbf{H})$, if

$$\left(u - f_\dagger^\eta\right)(x) = \sup_{x} \left(u(x) - f_\dagger^\eta(x)\right),$$

then,

$$u(x) - \lambda\mathbf{H}f_\dagger^\eta(x) - h(x) \le 0. \tag{3.1.17}$$

Convexity of $p \mapsto \mathcal{H}(x, p)$ yields the estimate

$$\begin{aligned}\mathbf{H}f_\eta(x) &= \mathcal{H}(x, \nabla f_\eta(x)) \\ &\le (1-\eta)\mathcal{H}(x, \nabla f(x)) + \eta\mathcal{H}(x, \nabla \Upsilon(x)) \\ &\le (1-\eta)\mathcal{H}(x, \nabla f(x)) + \eta C_\Upsilon = H_{\dagger,f}^\eta(x).\end{aligned}$$

Combining this inequality with (3.1.17), we have

$$u(x) - \lambda H_{\dagger,f}^\eta(x) - h(x) \le u(x) - \lambda\mathbf{H}f_\dagger^\eta(x) - h(x) \le 0,$$

establishing (3.1.16). The supersolution statement follows in a similar way. □

3.1.6 *Exponential tightness*

To establish exponential tightness, we first note that by [FK06, Corollary 4.19] it suffices to establish the exponential compact containment condition (cf. Definition 3.1.18). This is the content of the next proposition.

Proposition 3.1.31. *For all $K \subset E$ compact, $T > 0$ and $a > 0$ there is a compact set $\hat{K}_{K,T,a} \subset E$ such that*

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\left[\bigcup_{t\in[0,T]} \left\{Y^\varepsilon(t) \notin \hat{K}_{K,T,a}\right\} \neq \emptyset\right] \leq \max\{-a, \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}(X_\varepsilon(0) \notin K)\}. \tag{3.1.18}$$

Remark 3.1.32. Note that, since $Y^\varepsilon(0)$ satisfies the large deviations principle by assumption, inequality (3.1.18) gives the searched compact containment condition.

Proof of Proposition 3.1.31. First of all let's consider $\varphi \equiv 0$. Note that, by Lemma 3.1.22, we have $\sup_{x,z} H_0(x, \nabla \mathrm{Y}, z) = \sup_{x,z} V_{x,\nabla \mathrm{Y}(x)}(z) \leq C_{\mathrm{Y}}$. Choose $\beta > 0$ such that $TC_{\mathrm{Y}} - \beta \leq -a$. Since Y is continuous, there is some c such that the set $G := \{x \,|\, \mathrm{Y}(x) < c + \beta\}$ is non empty. Note that G is open and let $\overline{G}$ be the closure of G. Then, $\overline{G}$ is compact. Let $f(x) := \iota \circ \mathrm{Y}$ where ι is some smooth increasing function such that

$$\iota(r) = \begin{cases} r & \text{if } r \leq \beta + c, \\ 2\beta + c & \text{if } r \geq \beta + c + 2. \end{cases}$$

It follows that $\iota \circ \mathrm{Y}$ equals Y on $\overline{G}$ and is constant outside of a compact set. Set $f_\varepsilon = f \circ \eta_\varepsilon$, $g_\varepsilon = H_\varepsilon f_\varepsilon$ and $g = H_{f,\varphi}$. Note that $g(x,z) = H_\varphi(x, \nabla \mathrm{Y}(x), z)$ if $x \in \overline{G}$. Therefore, we have $\sup_{x\in\overline{G}, z\in E'} g(x,z) \leq C_{\mathrm{Y}}$. Let τ be the stopping time $\tau := \inf\left\{t \geq 0 \,\middle|\, Y^\varepsilon(t) \notin \overline{G}\right\}$ and let

$$M_\varepsilon(t) := \exp\left\{\frac{1}{\varepsilon}\left(f(Y^\varepsilon(t)) - f(Y^\varepsilon(0)) - \int_0^t g_\varepsilon(Y^\varepsilon(s), \bar{I}_\varepsilon(t))\mathrm{d}s\right)\right\}.$$

By construction M_ε is a martingale. Let $K \subset E$ be compact. We have

$$\begin{aligned}
&\mathbb{P}\left[\bigcup_{t\in[0,T]} \left\{Y^\varepsilon(t) \notin \overline{G}\right\} \neq \emptyset\right] \\
&\leq \mathbb{P}\left(Y^\varepsilon(0) \in K, \bigcup_{t\in[0,T]} \left\{Y^\varepsilon(t) \notin \overline{G}\right\}\right) + \mathbb{P}\left(Y^\varepsilon(0) \notin K\right) \\
&= \mathbb{E}\left[\mathbb{1}_{\{Y^\varepsilon(0)\in K\}} \mathbb{1}_{\left\{\bigcup_{t\in[0,T]}\left\{Y^\varepsilon(t)\notin\overline{G}\right\}\right\}} M_\varepsilon(\tau) M_\varepsilon(\tau)^{-1}\right] + \mathbb{P}\left(Y^\varepsilon(0) \notin K\right) \\
&\leq \exp\left\{-\frac{1}{\varepsilon}\left(\inf_{y_1\notin\overline{G}} f(y_1) - f(Y^\varepsilon(0))\right.\right. \\
&\qquad\qquad \left.\left. -T \sup_{y_2\in\overline{G}, i\in\{1,\ldots,J\}} g_\varepsilon(y_2, i)\right)\right\} \\
&\qquad \times \mathbb{E}\left[\mathbb{1}_{\{Y^\varepsilon(0)\in K\}} \mathbb{1}_{\left\{\bigcup_{t\in[0,T]}\left\{Y^\varepsilon(t)\notin\overline{G}\right\}\right\}} M_\varepsilon(\tau)\right] + \mathbb{P}\left(Y^\varepsilon(0) \notin K\right).
\end{aligned}$$

Since $\sup_{x\in\overline{G},z\in E'} g(x,z) \leq C_{\Upsilon,\varphi}$ and g is the limit of g_ε for $\varepsilon \to 0$ in the sense of Definition 3.1.11, we obtain that the term in the exponential is bounded by $\frac{1}{\varepsilon}(C_\Upsilon T - \beta) \leq -\frac{1}{\varepsilon}a$ for sufficiently small ε. The expectation is bounded by 1 due to the martingale property of $M_\varepsilon(\tau)$. We can conclude that

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left[\bigcup_{t\in[0,T]} \{Y^\varepsilon(t) \notin K_{T,a}\} \neq \emptyset\right] \leq \max\{-a, \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}\left(Y^\varepsilon(0) \notin K\right)\}$$

where $\hat{K}_{K,T,a} = \overline{G}$. □

3.1.7 *Action-integral representation of the rate function*

In this section we establish a representation of the rate function as an integral of a Lagrangian function $\mathcal{L}$. We refer to this representation as the "action-integral representation" of the rate function $\mathcal{I}$. We argue on basis of Section 8 of [FK06] for which we need to check the following two conditions.

Lemma 3.1.33. *Let $\mathcal{H} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the map given in (3.1.4) and $\mathbf{H} : \mathcal{D}(\mathbf{H}) \subseteq C^1(\mathbb{R}^d) \to C(\mathbb{R}^d)$ the operator $\mathbf{H}f(x) := \mathcal{H}(x, \nabla f(x))$. Then:*

(i) The Legendre-Fenchel transform $\mathcal{L}(x,v) := \sup_{p\in\mathbb{R}^d}(p\cdot v - \mathcal{H}(x,p))$ and the operator $\mathbf{H}$ satisfy Conditions 8.9, 8.10 and 8.11 of [FK06].

(ii) For all $\lambda > 0$ and $h \in C(\mathbb{R}^d)$, the comparison principle holds for $(1-\lambda\mathbf{H})u = h$.

Proof. To prove the first aim, we will show that following items (a), (b) and (c) imply Condition 8.9, 8.10 and 8.11 of [FK06]. Then, the proof of (a), (b), (c) is shown in [PS19, Proposition 6.1].

(a) The function $\mathcal{L} : \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ is lower semicontinuous and for every $C \geq 0$, the level set $\{(x,v) \in \mathbb{R}^d \times \mathbb{R}^d \,:\, \mathcal{L}(x,v) \leq C\}$ is relatively compact in $\mathbb{R}^d \times \mathbb{R}^d$.

(b) For all $f \in \mathcal{D}(H)$ there exists a right continuous, nondecreasing function $\psi_f : [0,\infty) \to [0,\infty)$ such that for all $(x,v) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$|\nabla f(x) \cdot v| \leq \psi_f(\mathcal{L}(x,v)) \qquad \text{and} \qquad \lim_{r\to\infty} \frac{\psi_f(r)}{r} = 0.$$

(c) For each $x_0 \in \mathbb{R}^d$ and every $f \in \mathcal{D}(\mathbf{H})$, there exists an absolutely continuous path $x : [0,\infty) \to \mathbb{R}^d$ such that $x_0 = x(0)$ and

$$\int_0^t \mathcal{H}(x(s), \nabla f(x(s)))\, ds = \int_0^t \left[\nabla f(x(s)) \cdot \dot{x}(s) - \mathcal{L}(x(s), \dot{x}(s))\right] ds.$$

Then regarding Condition 8.9, the operator $Af(x,v) := \nabla f(x)\cdot v$ on the domain $\mathcal{D}(A) = \mathcal{D}(H)$ satisfies (1). For (2), we can choose $\Gamma = \mathbb{T}^d \times \mathbb{R}^d$, and for $x_0 \in \mathbb{T}^d$, take the pair (x,λ) with $x(t) = x_0$ and $\lambda(dv \times dt) = \delta_0(dv) \times dt$. Part (3) is a consequence of (a) from above. Part (4) can be verified as follows. Let Υ the containment function used in

Definition 3.1.23 and note that the sub-level sets of Υ are compact. Let $\gamma \in \mathcal{AC}$ with $\gamma(0) \in K$ and such that the control

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

implies $\gamma(t) \in \hat{K}$ for all $t \leq T$, with $\hat{K}$ compact. Then,

$$\begin{aligned} \Upsilon(\gamma(t)) &= \Upsilon(\gamma(0)) + \int_0^t \langle \nabla\Upsilon(\gamma(s)), \dot{\gamma}(s)\rangle \; ds \\ &\leq \Upsilon(\gamma(0)) + \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s)) + \mathcal{H}(\gamma(s), \nabla\Upsilon(\gamma(s)))\, ds \\ &\leq \sup_{y\in K} \Upsilon(y) + M + \int_0^T \sup_z \mathcal{H}(z, \nabla\Upsilon(z))\, ds \\ &:= C < \infty. \end{aligned}$$

Hence, we can take $\hat{K} = \{z \in \mathbb{R}^d | \Upsilon(z) \leq C\}$.

Part (5) is implied by (b) from above. Condition 8.10 is implied by Condition 8.11 and the fact that $\mathbf{H}1 = 0$, by Theorem 3.1.16 (see Remark 8.12 (e) in [FK06]). Finally, Condition 8.11 is implied by (c) above, with the control $\lambda(dv \times dt) = \delta_{\dot{x}(t)}(dv) \times dt$.

The comparison principle for $\mathbf{H}$ follows from Proposition 3.1.30 and Theorem 3.1.26. □

In the following, we prove the integral representation of the rate function. Firstly, let's recall that Theorem 2.2.12 gives the existence of a semigroup $V(t)$ and a family of functions $R(\lambda)$ and let $\mathbf{V}(t) : C(\mathbb{R}^d) \to C(\mathbb{R}^d)$ be the Nisio semigroup with cost function $\mathcal{L}$, that is

$$\mathbf{V}(t)f(x) = \sup_{\substack{\gamma\in \mathrm{AC}_{\mathbb{R}^d}[0,\infty) \\ \gamma(0)=x}} \left[f(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(t), \dot{\gamma}(s))\, ds \right].$$

Let $\mathbf{R}(\lambda)h$ be the operator given by

$$\mathbf{R}(\lambda)h(x) = \sup_{\substack{\gamma\in\mathcal{AC} \\ \gamma(0)=x}} \int_0^\infty \lambda^{-1} e^{-\lambda^{-1}t} \left[h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s)) \right] \mathrm{d}t.$$

The proof of the result below is based on the following four main steps.

- Figure 8 on page 62 shows that $R(\lambda)h$ is the unique function that is a sub- and supersolution to the equations $(1 - \lambda H_\dagger)f = h$ and $(1 - \lambda H_\ddagger)f = h$ respectively.
- $\mathbf{R}(\lambda)h$ has been proven to be the unique viscosity solution to $(1 - \lambda\mathbf{H})f = h$. Then, again by Figure 8, we must have $R(\lambda)h = \mathbf{R}(\lambda)h$.
- Starting from the equality of resolvents we work to an equality for the semigroups $V(t)$ and $\mathbf{V}(t)$.

- Recalling that the rate function in Theorem 3.1.6 is given by,

$$I(x) = I_0(x(0)) + \sup_{k\in\mathbb{N}} \sup_{(t_1,\ldots,t_k)} \sum_{i=1}^{k} I_{t_i - t_{i-1}}(x(t_i)|x(t_{i-1}))$$

with $I_t(z|y) = \sup_{f\in C(E)}[f(z) - V(t)f(y)]$, it is not difficult to realise that, if $V(t) = \mathbf{V}(t)$, it follows that $I_t(y|z) = \inf_{\substack{\gamma:\gamma(0)=z,\\ \gamma(t)=y}} \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\, ds$.

Theorem 3.1.34 (Integral representation of the rate function). *The rate function of Theorem 3.1.6 has the following representation*

$$\mathcal{I}(x) = \begin{cases} \mathcal{I}_0(x(0)) + \int_0^\infty \mathcal{L}\left(x(t),\dot x(t)\right)\, dt & \text{if } x \in AC([0,\infty);\mathbb{R}^d), \\ \infty & \text{else}, \end{cases}$$

where $\mathcal{L}(x,v) = \sup_{p\in\mathbb{R}^d}\left[p\cdot v - \mathcal{H}(x,p)\right]$ *is the Legendre transform of* $\mathcal{H}$.

Proof. Following the above mentioned steps, we recall that, as stated by Theorem 2.2.12, there exists a family of operators $R(\lambda) : C_b(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$, such that for $\lambda > 0$ and $h \in C_b(\mathbb{R}^d)$, the function $R(\lambda)h$ is the unique function that is a viscosity solution to $(1-\lambda H)f = h$ and such that

$$\lim_{m\to\infty} \left|\left| R\left(\frac{t}{m}\right)^m f - V(t)f \right|\right| = 0 \qquad \text{for all } f \text{ in a dense set } D \subseteq C_b(\mathbb{R}^d). \quad (3.1.19)$$

See also [Kra20, Theorem 7.10] or [FK06, Theorem 7.17] for the construction of the operators $R(\lambda)$. By [KS20, Proposition 6.1] (or [FK06, Chapter 8]), $\mathbf{R}(\lambda)$ is the unique viscosity solution to $(1-\lambda\mathbf{H})f = h$. Then, Figure 8 on page 62 shows that it must equal $R(\lambda)h$. Moreover, we find by [FK06, Lemma 8.18] (whose assumptions are implied by Lemma 3.1.33 above) that for all $f \in C_b(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$

$$\lim_{m\to\infty} \mathbf{R}\left(\frac{t}{m}\right)^m f(x) = \mathbf{V}(t)f(x). \quad (3.1.20)$$

We conclude from (3.1.19) and (3.1.20) that $V(t)f = \mathbf{V}(t)f$ for all t and $f \in D$. Now recall that D is sequentially strictly dense so that equality for all $f \in C_b(\mathbb{R}^d)$ follows if $V(t)$ and $\mathbf{V}(t)$ are sequentially continuous. The first statement follows by Theorems [Kra22, Theorem 7.10] and [Kra20, Theorem 6.1]. The second statement follows by [FK06, Lemma 8.22]. We conclude that $V(t)f = \mathbf{V}(t)f$ for all $f \in C_b(\mathbb{R}^d)$ and $t \geq 0$. Using Theorem 8.14 of [FK06] and the convexity of $v \mapsto \mathcal{L}(x,v)$ we get the integral representation. □

3.1.8 *A more general theorem*

Analysing the proofs in the previous sections, we can state the following facts:

- In the proof of large deviations principle, the main steps are:

1. Convergence of the nonlinear operators H_ε to a multivalued operator H,
2. comparison principle for $(1-\lambda H)f = h$.

- The existence of an eigenvalue $\mathcal{H}(x,p)$ and its convexity, coercivity and continuity are crucial for our approach to comparison principle and
  - the arguments for existence, convexity and coercivity (proofs of Propositions 3.1.3 and 3.1.16) are based on the fact that $\mathcal{H}(x,p)$ is the eigenvalue of an operator of the type $B_{x,p} + V_{x,p} + R_x$ with the three operators that verify particular properties such as coercivity and the maximum principle,
  - to show the continuity of $\mathcal{H}$ the representation (3.1.4) is needed. In particular, some properties of V and I, like Γ- convergence, are necessary.

The above observations allow for a straightforward generalization in Theorem 3.1.37 and justify the assumptions of the next subsection. In this section we indeed prove the large deviation principle for a general switching Markov process. In particular, we will study the Markov process $(Y_t^\varepsilon, \bar{I}_t^\varepsilon)$, that is the solution to the Martingale problem corresponding to the following operator

$$A_\varepsilon f(x,i) := A_\varepsilon^{(i)} f(\cdot,i)(x) + \sum_{j=1}^{J} r_{ij}(x, x/\varepsilon)\,[f(x,j) - f(x,i)] \tag{3.1.21}$$

with $A_\varepsilon^{(i)} : \mathcal{D}(A_\varepsilon^{(i)}) \subseteq C(\mathbb{R}^d) \to C(\mathbb{R}^d)$ be the generator of a strong $\mathbb{R}^d$-valued Markov process, with domain $\mathcal{D}(A_\varepsilon^{(i)})$.

Here we give the assumptions needed.

Assumption 3.1.35. The nonlinear generators $H_\varepsilon f = \varepsilon e^{-f/\varepsilon} A_\varepsilon e^{f/\varepsilon}$ admits an extended limit $H \subseteq ex - LIM H_\varepsilon$ with H of the type

$$H := \left\{ (f, H_{f,\varphi}) \,:\, f \in C^2(\mathbb{R}^d),\ H_{f,\varphi} \in C(\mathbb{R}^d \times E') \text{ and } \varphi \in C^2(E') \right\}.$$

For all φ there exist a map $H_\varphi : \mathbb{R}^d \times \mathbb{R}^d \times E' \to \mathbb{R}$ such that for all $f \in D(H)$, $x \in \mathbb{R}^d$ and $z \in E'$, $H_{\varphi,f}(x,z) = H_\varphi(x, \nabla f, z)$. Moreover, the image H_φ has the representation

$$H_\varphi(x,p,z) = e^{-\varphi(z)}\,[B_{x,p} + V_{x,p} + R_x]\, e^{\varphi}(z)$$

with $p = \nabla f(x)$ and $B_{x,p}, V_{x,p}, R_x$ such that

(i) For all $p \in \mathbb{R}^d$ there exists an eigenfunction $g_{x,p} \in C^2(\mathbb{R}^d \times J)$ with $g^i_{x,p} > 0$ and an eigenvalue $\mathcal{H}(x,p)$ such that $[B_{x,p} + V_{x,p} + R_x]\, g_{x,p} = \mathcal{H}(x,p) g_{x,p}$.

(ii) $T_{x,p} = B_{x,p} + R_x$ verifies the maximum principle :
if $(i_m, y_m) = \arg\min \varphi$ then $e^{-\varphi(i_m,y_m)} T_{x,p} e^{\varphi(i_m,y_m)} \geq 0$.

(iii) $p \mapsto V_{x,p}$ is coercive uniformly with respect to x.

(iv) $p \mapsto B_{x,p}$ and $p \mapsto V_{x,p}$ are convex uniformly on compact sets.

The above assumption implies the convergence of the nonlinear operators and the existence of the principal eigenvalue $\mathcal{H}$. Moreover, it will imply convexity and coercivity of $\mathcal{H}(x,p)$.

Assumption 3.1.36. The eigenvalue $\mathcal{H}$ is of the type

$$\mathcal{H}(x,p) = \sup_{\mu \in \mathcal{P}(E')} \left[\Lambda(x,p,\mu) - I_{x,p}(\mu)\right]$$

with

$$\Lambda(x,p,\mu) = \int_{E'} V_{x,p}\, d\mu, \qquad \text{and} \qquad I_{x,p}(\mu) = -\inf_{u>0} \int_{E'} \frac{(B_{x,p}+R_x)u}{u}\, d\mu,$$

and the following properties hold

(i) $I_{x,p}$ satisfies the assumption of Lemma 3.1.18,

(ii) $\Lambda(x,p,\mu)$ is continuous and $\|\Lambda(x,p,\mu)\|_\Theta < \infty$,

(iii) there exists a containment function Υ for Λ in the sense of Definition 3.1.21,

(iv) for all x, there exists a unique measure μ_x^* such that $I_{x,0}(\mu_x^*) = 0$.

Assumption 3.1.36 implies the continuity of $\mathcal{H}$.

We are ready to state the general theorem.

Theorem 3.1.37 (Large deviation for a Switching Markov process)**.** *Let $(Y_t^\varepsilon, \bar{I}_t^\varepsilon)$ be the solution of the Martingale problem corresponding to the operator given in* (3.1.21)*. If Assumptions 3.1.35 and 3.1.36 hold and suppose further that at time zero, the family of random variables $\{Y^\varepsilon(0)\}_{\varepsilon>0}$ satisfies a large deviation principle in $\mathbb{R}^d$ with good rate function $\mathcal{I}_0 : \mathbb{R}^d \to [0,\infty]$. Then, the spatial component $\{Y_t^\varepsilon\}$ satisfies a large deviation principle in $C_{\mathbb{R}^d}[0,\infty)$.*

The proof of the above theorem follows the same lines of what is done in Subsection 3.1.5.

3.1.9 *Conclusions and comparison with previous works*

To conclude our work, in the following we summarize all the main novelties of our results.

1. We prove large deviations principle for the Markov process defined in Definition 3.1.1. The main steps of the proof are:
   a) Convergence of the nonlinear generators H_ε.
   b) Proof of continuity of the Hamiltonian $\mathcal{H}$.
   c) Comparison principle for $(1-\lambda H)u = h$.
   d) Proof of exponential tightness for X^ε.
   e) Proof of the integral representation of the rate function.
2. We prove the Law of large numbers for the path characterizing the limit process by calculating its speed. To do so, we also prove existence and uniqueness of the stationary measure of the operator (3.1.5).
3. We give a general result for a Switching Markov process.

The first result can be seen as an extension of part of the work [PS19]. We elaborate on the primary distinctions between our work and the previously cited one and how these distinctions contribute to the increased complexity of the proof of large deviations.

- We work on $\mathbb{R}^d$ and not on the torus $\mathbb{T}^d$. This transition from a compact to a non-compact set leads to the following significant consequences:
  (i) Firstly, in order to prove comparison principle, i.e step 1c, we need to construct four Hamiltonians in terms of a containment function that allows us to restrict part of our analysis on compact sets. Hence, we need to prove diagram 8. In [PS19], they only need two additional operators defined as multi-valued limit of the Hamiltonian H.
  (ii) Secondly, in a compact setting step 1d is trivial. Indeed, the exponential tightness is implied by exponential compact containment condition that is always verified in a compact set.
  (iii) In the proof of the integral representation of the rate function, step 1e, some details are not needed in a compact setting as part of condition 8.9 in [FK06].
- We introduce a spatial component x in the rates of the process that forces us to work with a Hamiltonian depending on both variables x and p. For this reason, we need to work with a spatially inhomogeneous Hamilton–Jacobi equation. In particular:
  (i) Proving comparison principle one usually wants to bound the difference between sub-solution and super-solution $\sup_E u_1 - u_2$ by using a doubling variables procedure and typically ends up with an estimate of the following type

$$\begin{aligned}\sup_E (u_1 - u_2) \leq& \lambda \liminf_{\varepsilon \to 0} \left[\mathcal{H}(x_\varepsilon, \alpha(x_\varepsilon - y_\varepsilon)) - \mathcal{H}(y_\varepsilon, \alpha(x_\varepsilon - y_\varepsilon))\right] \\ &+ \sup_E (h_1 - h_2).\end{aligned} \tag{3.1.22}$$

  If the Hamiltonian does not depend on x, the final estimate is

$$\begin{aligned}\sup_E (u_1 - u_2) \leq& \lambda \liminf_{\varepsilon \to 0} \left[\mathcal{H}(\alpha(x_\varepsilon - y_\varepsilon)) - \mathcal{H}(\alpha(x_\varepsilon - y_\varepsilon))\right] + \sup_E (h_1 - h_2) \\ =& \sup_E (h_1 - h_2),\end{aligned}$$

  that gives immediately comparison principle. This means that step 1c is partially immediate.
  (ii) To bound (3.1.22) in the non-spatially homogeneous case, we instead need to prove some regularity of the Hamiltonian. In particular, we need to prove continuity of $\mathcal{H}$ using some notions as Γ-convergence. If the Hamiltonian depends only on p, continuity, that is step 1b, follows immediately from convexity in p.

3.1.10 *Appendix: Uniqueness of the stationary measure*

We give here the proof of the existence and uniqueness of the stationary measure.

Proposition 3.1.38. *Under the ussumptions of Theorem 3.1.6, there exists a unique stationary measure of the operator*

$$L_{x,p}u(z,i) = \frac{1}{2}\Delta_z u(z,i) + (p - \nabla_x \psi^i(x,z)) \cdot \nabla_z u(z,i) + \sum_{j=1}^{J} r_{ij}(x,z)\left[u(z,j) - u(z,i)\right].$$

Proof. First of all, note that $L_{x,p}$ is of the type $T_{x,p} + R_x$ where $T_{x,p}$ is a diagonal matrix with diagonal elements $(T_{x,p})_{ii}$ of the type $\Delta + p \cdot \nabla + c$ and $(R_x)_{ii} = \sum_{j=1}^{J} r_{ij}$.

Let us consider, for some $\delta \in \mathbb{R}$, the operator $T_{x,p} + R_x + \delta$. For the latter operator, Conditions 1, 2 and 3 of [Swe92] hold. Then, by [Swe92, Theorem 1.1], there exists a unique function $\Psi \gg 0$ such that $(T_{x,p} + R_x + \delta)\Psi = \lambda\Psi$ for some $\lambda > 0$. It follows that $\lambda = \delta$ and Ψ is the identity function. Hence, $\ker(T_{x,p} + R_x)$ is one-dimensional and it is spanned by the identity function, i.e., it consists of constants. Let P_t be the semigroup associated to the generator $L_{x,p}$. By [EN00, Corollary V.4.6], P_t is mean ergodic, that means that the Cesàro mean

$$C(r) = \frac{1}{r}\int_0^r P_s \, ds,$$

has a limit $P : C_b(E') \to C_b(E')$ for $r \to \infty$. Moreover, by [EN00, Lemma V.4.2], $Pf \in \ker(L_{x,p})$ for every $f \in C_b(E')$.

Let $T : c\mathbb{1} \in \ker(L_{x,p}) \mapsto c \in \mathbb{R}$. Then, $T \circ P : C_b(E') \to \mathbb{R}$ is a linear continuous function on $C_b(E')$. Then, by Riesz–Representation theorem, there exists a unique measure μ such that $(T \circ P)f = \langle f, \mu \rangle$. We show now that μ is the unique invariant measure for P_t. For all $f \in C_b(E')$ we have

$$\langle P_t f, \mu \rangle = (T \circ P)(P_t f) = (T \circ P)f = \langle f, \mu \rangle,$$

where in the second equality we used that $P = P_t P = P P_t = P^2$ (see [EN00, Lemma V.4.4]). Moreover, if μ^* is an invariant probability measure for P_t, let Q be the projection $Q : f \in C_b(E') \mapsto \langle f, \mu^* \rangle \in \ker(L_{x,p})$. We show that $P = QP = Q$, obtaining then the uniqueness. On one hand, for $f \in C_b(E')$

$$QPf = \langle Pf, \mu^* \rangle = \langle \langle f, \mu \rangle, \mu^* \rangle = \langle f, \mu \rangle = Pf.$$

On the other hand,

$$QPf = \lim_{r \to \infty} \langle C(r)f, \mu^* \rangle = \lim_{r \to \infty} \langle f, C^*(r)\mu^* \rangle = \langle f, \mu^* \rangle = Qf.$$

□

3.2 LARGE DEVIATIONS PRINCIPLE FOR TWO-SCALES MARKOV PROCESSES MODELING CHEMICAL REACTIONS NETWORKS

Over the past few decades, there has been significant research conducted on multi-scale chemical reactions networks (see for example [KK13] and [Bal+06]). They are usually described by the use of continuous - time Markov chains with generators of the following form

$$Af(z) = \sum_{\gamma\in\Gamma} r(z,\gamma)\left[f(z+\gamma) - f(z)\right],$$

with z in the space $\mathbb{N}^J$, with J the set of chemical species, and $r \in C^1(\mathbb{N}^J \times \Gamma)$ is a non-negative smooth function. In the above, the state z is a vector whose components describe the number of molecules of a chemical species, r is the transition rate of the reaction, and Γ is the set representing all reactions. In particular, every $\gamma \in \Gamma$ describes a reaction in the sense that its i–th component represents the number of molecules of the i–th species that are used (if the component is negative) or obtained (if the component is positive) in the reaction.

Motivated by a huge class of examples arising from biochemistry in which two dominant time-scales occur (see, for instance, Example 3.2.2), we will consider a two - scale process $Z = (X, Y)$ in the space $E^0 = \mathbb{N}^l \times \mathbb{N}^m$. The amount of molecules of the first type is an order of magnitude greater than the amount of the second type. For this reason and to be able to study the limit behavior of the process, we consider scaled species $X_N = X/N$ and $Y_N = Y$ in the space $E^0_N = \left(\frac{1}{N}\mathbb{N}\right)^l \times \mathbb{N}^m$. The time-scale separation between the slow process X_N and the fast process Y_N is then N and the generator of the rescaled process $Z_N = (X_N, Y_N)$ is given by

$$A_N f(x,y) = N \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma} r(x,y,\gamma)[f(x+N^{-1}\gamma_x, y+\gamma_y)], \tag{3.2.1}$$

for $f \in D(A_N) \subseteq C(E^0_N)$.

In order to gain a more comprehensive understanding of the system and simplify the limit procedure that will follow, we divide the generator into three distinct parts. These parts will specifically describe reactions occurring on the macroscopic, microscopic, and a combination of the two scales, respectively. The generator of (X_N, Y_N) is then,

$$\begin{aligned} A_N f(x,y) =& N \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_1} r(x,y,\gamma)\left[f(x+N^{-1}\gamma_x, y) - f(x,y)\right] + \\ & N \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_2} r(x,y,\gamma)\left[f(x+N^{-1}\gamma_x, y+\gamma_y) - f(x,y)\right] + \\ & N \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_3} r(x,y,\gamma)\left[f(x, y+\gamma_y) - f(x,y)\right], \end{aligned} \tag{3.2.2}$$

where we write

$$\begin{aligned} \Gamma_1 &= \left\{\gamma = (\gamma_x,\gamma_y) \in \mathbb{Z}^l \times \mathbb{Z}^m \,:\, \gamma_{y_i} = 0\,\forall i \in \{1,\dots,m\}\right\}; \\ \Gamma_2 &= \{\gamma = (\gamma_x,\gamma_y) \in \mathbb{Z}^l \times \mathbb{Z}^m : \exists i \in \{1,\dots,l\}, j \in \{1,\dots,m\} | \\ &\qquad\qquad\qquad\qquad \gamma_{x_i} \neq 0, \gamma_{y_j} \neq 0\}; \\ \Gamma_3 &= \left\{\gamma = (\gamma_x,\gamma_y) \in \mathbb{Z}^l \times \mathbb{Z}^m \,:\, \gamma_{x_i} = 0\,\forall i \in \{1,\dots,l\}\right\}. \end{aligned}$$

The model is subjected to the following assumption.

Assumption 3.2.1. The molecules that are part of the fast process are subjected to a conservation law. More precisely, there exists a constant $M > 0$ such that

$$\sum_{i=1}^{m} Y_i = M, \qquad \text{and} \qquad \sum_{i=1}^{m} \gamma_{y,i} = 0 \qquad \forall \gamma \in \Gamma_2 \cup \Gamma_3.$$

Assumption 3.2.1 allows us to restrict the set of values of Y_N to the set

$$F_M = \{n \in \mathbb{N}^m : \sum_{i=1}^{m} n_i = M\},$$

and, hence, we consider for our analysis of Z_N the set

$$E_N = \left(\frac{1}{N}\mathbb{N}\right)^l \times F_M.$$

To show an example, we describe in the following the model for enzyme kinetics with an inflow of the substrate, also called Michaelis–Menten model, studied in [Pop18].

Example 3.2.2. *Consider four types of molecules. Namely, S, E, ES and P representing respectively the substrate, enzyme, enzyme–substrate complex and the product. The following four reactions occur.*

(1) $\emptyset \xrightarrow{k_0} S$
(2) $E + S \xrightarrow{k_1} ES$
(3) $ES \xrightarrow{k_2} E + S$
(4) $ES \xrightarrow{k_3} P + E.$

Let X^1, X^2, Y^1, Y^2 represent the amount of S, P, E, ES respectively.

In real-world physical scenarios, the quantities of enzyme molecules and enzyme - substrate complexes are typically small when compared to the number of substrate and product molecules. Consequently, it is reasonable to assume that X^1 and X^2 are an order of magnitude greater then Y^1 and Y^2. In this way, we lead to the scaled amounts represented by a slow process $X_N = (X^1/N, X^2/N)$ and a fast process $Y_N = (Y^1, Y^2)$.

The generator of the two–scales process $Z_N = (X_N, Y_N)$ is as in (3.2.2), with the following rates, each describing one of the reactions above.

(1) $r(x, y, (1, 0, 0, 0)) = k_0 \in \mathbb{R}_+;$
(2) $r(x, y, (-1, 0, -1, 1)) = k_1 x_1 y_1 \qquad k_1, \in \mathbb{R}_+;$
(3) $r(x, y, (1, 0, 1, -1)) = k_2 y_2 \qquad k_2 \in \mathbb{R}_+;$
(4) $r(x, y, (0, 1, 1, -1)) = k_3 y_2 \qquad k_3 \in \mathbb{R}_+.$

The above model falls into the class of examples described by (3.2.2).

We are interested in the limit behavior of the two component E_N-valued Markov process (X_N, Y_N). In the limit regime, the fast component Y_N converges to equilibrium and the slow component X_N converges to a deterministic limit. To characterize the speed of convergence, we are interested in the large deviation behavior for the slow process.

We saw in Chapters 2 and 3.1 that the rate function function $\mathcal{I}$ of the large deviations principle is characterized by the unique solution of the Hamilton-Jacobi equation

$$f - \lambda \mathbf{H} f = h \qquad (\text{or } \partial_t f - \mathbf{H} f = 0),$$

with $\mathbf{H}f = \mathcal{H}(f, \nabla f)$ found in three steps:

1. Given the generator A_N of the process (X_N, I_N), define the non linear generator $H_N f = \frac{1}{N} e^{-Nf} A_N e^{Nf}$ for f such that $e^{Nf} \in D(A_N)$;
2. Given $f \in C(E)$ and $h \in C(E \times \mathbb{R}^l)$ define $f_N(x,y) = f(x) + N^{-1} h(x,y)$ and find H_h such that $\lim_{N\to\infty} H_N f_N(x,y) = H_h(x, \nabla f(x), y)$;
3. For every x, find h_x such that H_{h_x} does not depend on y. Using $p = \nabla f(x)$, this is equivalent to solving an eigenvalue problem : for all $x \in E$ and $p \in \mathbb{R}^n$, there exist $\mathcal{H}(x,p)$ and $\bar{h}$ such that $H(x,p,y)\bar{h}(x,y) = \mathcal{H}(x,p)\bar{h}(x,y)$.

In the following, we use the three steps above to obtain the Hamiltonian linked to this model. Later, in Chapter 4, Theorem 4.6.1, we prove the comparison principle for the Hamilton-Jacobi equation in terms of this Hamiltonian, implying then Large deviations principle for the process. For these aims, we need two additional assumptions.

Assumption 3.2.3. The matrix $(R_x)_{y_1,y_2} = \sum_{k\in\{2,3\}} \sum_{\gamma\in\Gamma_k : y_2 = y_1+\gamma_y} r(x, y_1, \gamma)$ is irreducible for every x.

Assumption 3.2.4. For all z and γ there exist continuous functions $\phi_1^{z,\gamma}, \phi_2^{z,\gamma} : [0,\infty)^l \to \mathbb{R}$ such that $r(x,z,\gamma) = \phi_1^{z,\gamma}(x)\phi_2^{z,\gamma}(x)$ and such that

1. $\inf \phi_2^{z,\gamma} > 0$;
2. if $\langle \gamma_x, x - y\rangle > 0$ then $\phi_1^{z,\gamma}(x) < \phi_1^{z,\gamma}(y)$;
3. if $\gamma_{x_i} < 0$ and $x_i = 0$ then $\phi_1(x) = 0$.

Assumption 3.2.3 guarantees the existence of an eigenvalue (i.e. step 3 above).

Assumption 3.2.4 restricts the possible rates, and hence the Hamiltonians that we are able to treat with our results. First of all, it removes cases of type $\mathcal{H}(x,p) = x(e^p - 1)$, with $x \in [0,\infty)$, for which it is well known that comparison principle fails (see e.g. Example E in [SW05]). Moreover, it is straightforward to show Assumption 3.2.4 in many cases in one dimension, e.g. $\mathcal{H}(x,p) = x(e^{-p}-1)$, $\mathcal{H}(x,p) = xe^{\beta x}(e^{-p}-1)$ in $[0,\infty)$. It is also straightforward to verify the above assumptions for the two-dimensional Example 3.2.2. Regarding higher dimensional cases, the assumption excludes examples in which there is an interaction between two molecules of the slow process. These cases, indeed, produce rates of the type $r_{z,\gamma}(x) = x_i x_j$ with $i,j \in \{1,\dots,l\}$ or such that $\langle \gamma_x, x-y\rangle > 0$ and not equal to $\phi_1(y) - \phi_1(x)$, for which Assumption 3.2.4 does not necessary hold. However, to our knowledge, examples of this type are largely unexplored.

Cases in higher dimension for which the above assumption holds are e.g. Hamiltonians of the type $\mathcal{H}(x,p) = x_1(1+x_2)(e^{p_1}-1)$.

Proposition 3.2.5. *Consider the Markov process (X_N, Y_N) having the operator (3.2.1) as generator. Suppose Assumptions 3.2.3 and 3.2.4. Then, the following hold:*

1. *Let* $f \in C(E)$ *and* $h \in C(E \times F_M)$. *Define* $f_N(x,y) = f(x) + N^{-1}h(x,y)$ *and* $H_N f = \frac{1}{N}e^{-Nf}A_N e^{Nf}$ *provide* $e^{Nf} \in D(A_N)$. *Then,*

$$\lim_N H_N f_N(x,y) = V(y;x,\nabla_x f(x)) + e^{-h(x,y)} L_{x,\nabla_x f(x)} e^{h(x,y)},$$

with

$$V(y;x,p) = \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_1} r(x,y,\gamma)(e^{\langle p,\gamma_x\rangle} - 1) \\ + \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_2} r(x,y,\gamma)(e^{\langle p,\gamma_x\rangle} - 1), \tag{3.2.3}$$

and

$$L_{x,p} f(x,y) = \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_2} r(x,y,\gamma)e^{\langle p,\gamma_x\rangle}[f(x,y+\gamma_y) - f(x,y)] \\ + \sum_{\gamma=(\gamma_x,\gamma_y)\in\Gamma_3} r(x,y,\gamma)[f(x,y+\gamma_y) - f(x,y)]. \tag{3.2.4}$$

2. *There exists a unique constant* $\mathcal{H}(x,p)$ *and a unique function* $g(x,y)$ *that solve the eigenvalue problem* $(V(y;x,p) + L_{x,p})g(x,y) = \mathcal{H}(x,p)g(x,y)$.
3. *The map* $\mathcal{H}(x,p) : E \times \mathbb{R}^d \to \mathbb{R}$ *has the following representation*

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta}\left[\int_{F_M} V(y;x,p)\,d\theta(y) + \inf_{\varphi\in C^2(F_M)}\int_{F_M} e^{-\varphi}L_{x,p}e^{\varphi}\,d\theta\right],$$

with $\Theta = \mathcal{P}(F_M)$.

Proof sketch. Recalling the generator A_N in (3.2.2), note that the exponential generator H_N acting on the test functions f_N are

$$\begin{aligned} H_N f_N(x,y) = &\sum_{\gamma\in\Gamma_1} r(x,y,\gamma)\left[e^{N\left(f\left(x+\frac{1}{N}\gamma_x\right)-f(x)\right)+h\left(x+\frac{1}{N}\gamma_x,y\right)-h(x,y)} - 1\right] \\ &+ \sum_{\gamma\in\Gamma_2} r(x,y,\gamma)\left[e^{N\left(f\left(x+\frac{1}{N}\gamma_x\right)-f(x)\right)+h\left(x+\frac{1}{N}\gamma_x,y+\gamma_y\right)-h(x,y)} - 1\right] \\ &+ \sum_{\gamma\in\Gamma_3} r(x,y,\gamma)\left[e^{h(x,y+\gamma_y)-h(x,y)} - 1\right]. \end{aligned}$$

As a consequence, its limit is

$$\begin{aligned} \lim_N H_N f_N(x,y) = &\sum_{\gamma\in\Gamma_1} r(x,y,\gamma)[e^{\langle\nabla f(x),\gamma_x\rangle} - 1] \\ &+ \sum_{\gamma\in\Gamma_2} r(x,y,\gamma)[e^{\langle\nabla f(x),\gamma_x\rangle}e^{h(x,y+\gamma_y)-h(x,y)} - 1] \\ &\sum_{\gamma\in\Gamma_3} r(x,y,\gamma)[e^{h(x,y+\gamma_y)-h(x,y)} - 1] \end{aligned}$$

and the first claim is proven.

The second point follows from Assumption 3.2.3 and Perron – Frobenius Theorem.

The third point follows from the fact that, being the eigenvalue of $V_{x,p} + L_{x,p}$, $\mathcal{H}(x,p)$ can be written in the following way

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left[\int_{F_M} V(y;x,p)\, d\theta(y) + \inf_{\varphi\in C^2(F_M)} \int_{F_M} e^{-\varphi} L_{x,p} e^{\varphi}\, d\theta \right], \tag{3.2.5}$$

with $\Theta = \mathcal{P}(F_M)$ (see [DV75] for more details about the representation of the eigenvalue). □

We want to point out that in [Pop18], the author studies the limit behavior of a similar example following the method developed by Feng and Kurtz described in the steps above. However, in contrast with what has been done in the work cited above, with our approach it is not necessary to find and calculate an explicit expression of the eigenvalue $\mathcal{H}$ found in step 3. Therefore, our method presents an effective solution that can be applied to similar scenarios in which the eigenvalue problem cannot be easily solved.

3.3 CONCLUSIONS AND FUTURE PERSPECTIVES

In this chapter, we presented two examples of Markov processes from biological and biochemical contexts, where proving large deviations leads to a comparison principle for the Hamilton-Jacobi equation. In both cases, the Hamiltonian takes the form:

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left\{ \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta) \right\}.$$

Motivated by these examples, which arise naturally from large deviation principles, we proceed in the next chapter to prove the comparison principle for HJB equations with Hamiltonians of the type described above, filling the gap between large deviation theory and the Hamilton-Jacobi framework.

We now conclude by outlining some promising directions for future work that arise from the results of this chapter.

- **Open problem 1:** Extending to an infinite state set for the switching process of Chapter 3, i.e., $J = \infty$ in (3.1.1), is not a trivial generalization. It indeed introduces many technical complications in the method used above. We list here the two main issue that arise:
  - The Hamiltonian of the associated HJ equation is derived by solving an eigenvalue problem, but this approach is often challenging or infeasible for an infinite state space (see [DV75] for related discussions).
  - The introduction of an unbounded set of controls adds significant technical difficulties to the comparison principle proof, which requires careful, case-specific adjustments.

We then need different techniques to deal with the above problems such as the introduction of an upper and lower bound of the Hamiltonian in terms of a Lyapunov function that allows to restrict on compact sets.

- **Open problem 2:** Consider the same molecular motors model as above, but now with the diffusion replaced by an underdamped diffusion. This modification provides a more realistic representation of particle motion. In this case, the model becomes a multiscale process $(X_t^\epsilon, Y_t^\epsilon, I_t^\epsilon)$, governed by the following system of stochastic differential equations:

$$
\begin{aligned}
&dY_t^\epsilon = X_t^\epsilon \, dt, \\
&dX_t^\epsilon = F^\varepsilon(X_t^\varepsilon, Y_t^\varepsilon, I_t^\varepsilon) \, dt + dB_t \\
&\mathbb{P}\Big(I^\epsilon(t+\Delta t) = j \,\Big|\, I^\epsilon(t) = i, (X^\epsilon(t), Y^\varepsilon(t)) = (x,y)\Big) \\
&= r_{ij}^\epsilon(x,y)\Delta t + \mathcal{O}(\Delta t^2) \quad \text{as } \Delta t \to 0.
\end{aligned}
$$

  As before, we introduce a rescaling parameter $\epsilon > 0$ and study the behavior of the system in the limit as $\epsilon \to 0$. However, this case introduces a more complex class of Hamilton-Jacobi equations. Specifically, the resulting Hamiltonian lacks the coercivity and regularity properties seen in the earlier setting. Moreover, the associated control problem is no longer on a compact space but on the space $\mathcal{P}(\mathbb{R}^d \times \mathbb{T}^d \times \{1, \ldots, J\})$. Consequently, analyzing this example requires alternative techniques, such as incorporating a Lyapunov control framework, to handle the additional complexity.
- **Open problem 3:** Recently, the study of Markov processes modeling population dynamics (e.g. population growth, epidemic models, prey-predator systems) has garnered significant interest. The application of the method explained above to these models leads to the study of a Hamilton-Jacobi equation with domain that have corners (see e.g. [KM20], [DIS90]). One has then to deal with possible irregularities of the Hamiltonian at the boundary. Recent works explore the case of jump processes. To the best of our knowledge cases of diffusion processes with discontinuities at the boundary is open.

Part III

WELL- POSEDNESS OF HAMILTON-JACOBI EQUATIONS

4

COMPARISON PRINCIPLE FOR A FIRST-ORDER HAMILTON-JACOBI-BELLMAN EQUATION IN THE STRONG COUPLING REGIME

In Chapter 2, we saw that well-posedness of Hamilton-Jacobi equations plays an important role in the Large deviations theory. In particular, we presented two examples arising from biology and biochemistry in which proving large deviations leads to well-posedness in the viscosity sense of a Hamilton-Jacobi equation with, in both cases, an Hamiltonian of the following type

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left\{\Lambda(x,p,\theta) - \mathcal{I}(x,p)\right\}. \tag{4.0.1}$$

In this chapter, we explore the comparison principle for Hamilton-Jacobi-Bellman equations with a Hamiltonian as in (4.0.1)

The results presented in this chapter are based on:

S. Della Corte and R.C. Kraaij, "Well-posedness of a Hamilton-Jacobi-Bellman equation in the strong coupling regime", Preprint, 2023.

4.1 INTRODUCTION

In this chapter, we study well-posedness of the following Hamilton-Jacobi-Bellman equation on a subset $E \subseteq \mathbb{R}^d$,

$$u(x) - \lambda\mathcal{H}(x, \nabla u(x)) = h(x), \tag{4.1.1}$$

where λ is a positive constant and h is a continuous bounded function, and for the time–dependent version

$$\begin{cases} \partial_t u(x,t) - \mathcal{H}(x, \nabla_x u(t,x)) = 0, & \text{if } t > 0, \\ u(0,x) = u_0(x) & \text{if } t = 0. \end{cases} \tag{4.1.2}$$

In the entire chapter we consider a Hamiltonian of the type

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left[\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)\right]. \tag{4.1.3}$$

The main goal of this chapter is to prove the *comparison principle* for viscosity solutions of the above equations (4.1.1) and (4.1.2), implying also the uniqueness of solutions.

Comparison principle for viscosity solutions has been largely studied in the past years with an increasingly complex Hamiltonian. Above all, we mention [BC97] for a proof of comparison principle for equations arising from *optimal control* problems and [KP17] for Hamiltonians coming from the theory of large deviations for Markov processes. We already mentioned in Chapter 1 that, in these settings, the standard assumptions used to obtain the comparison principle are usually either the *modulus continuity* of $\mathcal{H}$ i.e.

$$|\mathcal{H}(x,p) - \mathcal{H}(y,p)| \leq \omega(|x-y|(1+|p|)),$$

or uniformly coercivity of $\mathcal{H}$, that is

$$\sup_{x \in K} \mathcal{H}(x,p) \to \infty \qquad \text{if } |p| \to \infty.$$

In the case in which $\mathcal{H}$ is in a variational representation as in (4.1.3), the above assumptions can be derived from conditions on Λ and $\mathcal{I}$ such that uniformly coercivity or pseudo-coercivity of Λ and regularity and boundedness of the cost function $\mathcal{I}$ and the existence of a modulus of continuity for it. In [BC97], it is also proved comparison principle for Hamilton-Jacobi equations with Hamiltonian that satisfies the above conditions locally.

Later, [CIL87a] improved the conditions above by introducing new assumptions on the growth of the Hamiltonian in the second variable. This conditions are, however, still typically verified by examples in which at least the cost function is not depending on momenta.

However, there is a wide class of examples violating the above assumptions. In particular, this is the case of Hamiltonians arising in the study of systems with multiple time–scales such as the ones presented in Chapter 3 (see for example also [Pop18]).

We also want to mention that in works as [BS91], [CS97] and [CS05] it is showed comparison principle for Hamiltonians with possible unbounded, but still with a modulus of continuity and not depending on momenta, cost function.

More recently, in [KS21] the authors prove well–posedness for viscosity solutions of a general Hamilton–Jacobi–Bellman equation that can be applied in many of the above contexts. It is proved comparison principle under more generic and weaker assumptions then the common ones explained above. To be more precised, the authors in [KS21] prove for the first time comparison principle for an Hamilton–Jacobi–Bellman equation with Hamiltonian of the type (4.1.3) with Λ that can be non coercive, non pseudo–coercive and non Lipschitz and $\mathcal{I}$ that can be unbounded and discontinuous, but not depending on momenta p. Our work can then be seen as an extension of the above mentioned work as we introduce a cost function $\mathcal{I}$ depending on momenta p.

We saw in Chapter 3 that Hamiltonians with a cost function that is unbounded and discontinuous and depending on momenta are typically the ones that arise in a large deviations context. In this setting, indeed, it is common to get a cost function that is of the "Donsker-Varadhan type" (see [DV75]), that is

$$\mathcal{I}(x,p,\theta) = -\inf_{\phi>0} \int \frac{L_{x,p}\phi}{\phi}\, d\theta$$

with $L_{x,p}$ an operator. These types of Hamiltonians are not covered by the previous mentioned works because of its irregularity and the presence of the momenta in the cost function.

In section 4.6 we explain in more details how our work includes all the examples previously covered in works such as [BC97] and [KS21] and we give two extra examples arising from a large deviations context to show that our work can study new examples not included previously.

The introduction of the momenta in the function $\mathcal{I}$ makes on one hand the setting even more general including examples arising from problems in homogenisation theory that could not be treated before, and on the other hand the Hamiltonian more difficult to treat as it takes into account contributions from both parts Λ and $\mathcal{I}$. For this reason, it is necessary a change of the starting assumptions based on the difference $\Lambda - \mathcal{I}$ and not on the two separate functions.

In the following we present a concise overview of our strategy, without delving deep into specific details.

We saw in Chapter 1 that proving comparison principle one usually wants to bound the difference between subsolution and supersolution $\sup_E u_1 - u_2$ by using a doubling variables procedure and typically ends up with an estimate of the following type

$$\sup_E (u_1 - u_2) \leq \sup_E (h_1 - h_2) + \lambda \liminf_{\varepsilon \to 0} \liminf_{\alpha \to \infty} \Big[\mathcal{H}\left(x_{\alpha,\varepsilon}, \mathrm{d}_x \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})\right) - \mathcal{H}\left(y_{\alpha,\varepsilon}, -\mathrm{d}_y \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})\right)\Big]. \tag{4.1.4}$$

Therefore, the aim is usually to bound the difference of Hamiltonians in two sequences of points, $x_{\alpha,\varepsilon}$ and $y_{\alpha,\varepsilon}$, obtained as optimizers in the doubling variables procedure, and corresponding momenta $p^1_{\alpha,\varepsilon} = \mathrm{d}_x \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = -\mathrm{d}_y \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = p^2_{\alpha,\varepsilon}$.

Unlike the approach taken in [KS21] and in [BC97], where Λ and $\mathcal{I}$ were worked on independently, in the strong coupling regime, where $\mathcal{I}$ depends on p, we need to consider their difference as a single function. Indeed, we give assumptions on $\Lambda - \mathcal{I}$. Our main assumptions are as follows:

- Firstly, we rely on the *continuity estimate* of $\Lambda - \mathcal{I}$ that is morally the comparison principle for $\Lambda - \mathcal{I}$ for fixed θ. Indeed, it enables us to control the difference of $\mathcal{H}$ in (4.1.4) by managing the difference of

  $$\left(\Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon})\right) - \left(\Lambda(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon})\right)$$

  for $\theta_{\alpha,\varepsilon}$ optimizing $\mathcal{H}$ and well chosen as explained in the next point. See Assumption 4.3.3 (VII) for the rigorous notions.
- In order to get comparison for the Hamilton–Jacobi-Bellman equation in terms of $\mathcal{H}$ by the continuity estimate for $\Lambda - \mathcal{I}$, we also need to control the $\theta_{\alpha,\varepsilon}$. For this reason, we assume the compactness of the level sets of $\mathcal{I} - \Lambda$. Using this assumption we are indeed able to prove that the above sequence $\theta_{\alpha,\varepsilon}$ is relatively compact, i.e. (C3). This assumption is made rigorous in Assumption 4.3.3 (V).

- We assume Γ–convergence for $\mathcal{I} - \Lambda$ to prove regularity of $\mathcal{H}$. This assumption is typically true for the most treated examples, e.g. when Λ and $\mathcal{I}$ are continuous or when $\mathcal{I}$ arises as a Donsker–Varadhan functional (see [DV75]). See Assumption 4.3.3 (IV).

The rest of the chapter is structured as follows: In Section 4.2 we firstly give some preliminaries and definitions and an overview of the general setting. In Section 4.3, we state Theorem 4.3.1 and the assumptions needed to prove it. Then, we prove continuity of the Hamiltonian in Section 4.4. In Section 4.5 we give the proof of comparison principle and we state the existence of solutions. Finally, in Section 4.6 we treat two new examples to show that our assumptions are well-posed and we compare our work with the literature.

4.2 PRELIMINARIES AND MAIN DEFINITIONS

Throughout the chapter, E will be the set on which we base our Hamilton-Jacobi equations. We assume that E is an open subset of $\mathbb{R}^d$. We give now some definitions that we will use to give our assumptions in Section 4.3.

For clarity, we reiterate the definition of our Hamiltonian $\mathcal{H} : E \times \mathbb{R}^d \to \mathbb{R}$ given by

$$\mathcal{H}(x,p) = \sup_{\theta \in \Theta} \left[\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)\right]. \tag{4.2.1}$$

We start with the definition of *continuity estimate*. We will apply the definition below for $\mathcal{G} = \Lambda - \mathcal{I}$.

The reader familiar with comparison principle recognize the key conditions for establishing comparison for $f - \lambda \mathcal{G} f = h$ and $\partial_t f - \mathcal{G} f = 0$. The continuity estimate is indeed exactly the estimate that one would perform when proving the comparison principle for the Hamilton-Jacobi equation in terms of the Hamiltonian (4.2.1) (disregarding the supremum over θ). Indeed, in standard proofs of comparison principle one usually wants to control the difference of Hamiltonians calculated in particular collections of points.

Definition 4.2.1 (Continuity estimate)**.** Let $\mathcal{G} : E \times \mathbb{R}^d \times \Theta \to \mathbb{R}$, $(x,p,\theta) \mapsto \mathcal{G}(x,p,\theta)$ be a function.

Suppose that for each ε and $\alpha \in (0,\infty)^2$, $\alpha = (\alpha_1, \alpha_2)$, we have variables $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ in E^2 and variables $\theta_{\alpha,\varepsilon}$ in Θ. We say that this collection is *fundamental* for $\mathcal{G}$ with if:

(C1) For each ε, there is a compact set $K_\varepsilon \subseteq E$ such that for all α we have $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in K_\varepsilon$ and for all $\varepsilon > 0$, $\alpha_2 > 0$ there is a compact set $\widehat{K}_{\varepsilon,\alpha_2} \subseteq \Theta$ such that $\theta_{(\alpha_1,\alpha_2),\varepsilon} \in \widehat{K}_{\varepsilon,\alpha_2}$.

(C2) For each $\varepsilon > 0$ and $\alpha_2 > 0$ we have $\lim_{\alpha_1 \to \infty} \alpha_1 d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = 0$. For any limit point $(x_{\alpha_2,\varepsilon}, y_{\alpha_2,\varepsilon})$ of $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ as $\alpha_1 \to \infty$, we have $\mathrm{d}(x_{\alpha_2,\varepsilon}, y_{\alpha_2,\varepsilon}) = 0$.

(C3) We have for all $\varepsilon > 0$

$$\sup_\alpha \mathcal{G}\left(y_{\alpha,\varepsilon}, -\sum_{i=1}^2 \frac{\alpha_i}{2} \mathrm{d}_y d^2(x_{\alpha,\varepsilon}, \cdot)(y_{\alpha,\varepsilon}), \theta_{\alpha,\varepsilon}\right) < \infty,$$

$$\inf_\alpha \mathcal{G}\left(x_{\alpha,\varepsilon}, \sum_{i=1}^2 \frac{\alpha_i}{2} \mathrm{d}_x d^2(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}), \theta_{\alpha,\varepsilon}\right) > -\infty.$$

We say that $\mathcal{G}$ satisfies the *continuity estimate* if for every fundamental collection of variables we have that

$$\liminf_{\varepsilon\to 0} \liminf_{\alpha_2\to\infty} \liminf_{\alpha_1\to\infty} \mathcal{G}\left(x_{\alpha,\varepsilon}, \sum_{i=1}^2 \alpha_i \mathrm{d}_x \frac{1}{2} d^2(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}), \theta_{\alpha,\varepsilon}\right) - \mathcal{G}\left(y_{\alpha,\varepsilon}, -\sum_{i=1}^2 \alpha_i \mathrm{d}_y \frac{1}{2} d^2(x_{\alpha,\varepsilon}, \cdot)(y_{\alpha,\varepsilon}), \theta_{\alpha,\varepsilon}\right) \leq 0.$$

Typically, the control on $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ that is assumed in (C1) and (C2) is obtained from choosing $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ as optimizers in the doubling of variables procedure (see Section 1.2.3 or Lemma 4.5.5), and the control that is assumed in (C3) is obtained by using the viscosity sub- and supersolution properties in the proof of the comparison principle. To obtain these sequences in the doubling variable procedure we will include in the test functions a *containment function*.

Definition 4.2.2 (Containment function)**.** We say that a function $\Upsilon : E \to [0, \infty]$ is a *containment function* for $\mathcal{H}$ if $\Upsilon \in C^1(E)$ and there exists a constant c_Υ such that

- For every $c \geq 0$, the set $\{x \,|\, \Upsilon(x) \leq c\}$ is compact;
- $\sup_\theta \sup_x \left(\Lambda(x, \nabla\Upsilon(x), \theta) - \mathcal{I}(x, \nabla\Upsilon(x), \theta)\right) \leq c_\Upsilon$.

To prove the main results we will also make use of the continuity of $\mathcal{H}$ that will be proved in Proposition 4.4.1 by making use of the notion of Γ-convergence for the function $\mathcal{I} - \Lambda$.

Definition 4.2.3 (Γ-convergence)**.** Let $J : E \times \mathbb{R}^d \times \Theta \to \mathbb{R} \cup \{\infty\}$. We say that J is Γ–convergent in terms of (x, p), if

1. If $x_n \to x$ in E, $p_n \to p$ in $\mathbb{R}^d$ and $\theta_n \to \theta$ then $\liminf_n J(x_n, p_n, \theta_n) \geq J(x, p, \theta)$,
2. For $x_n \to x$ and $p_n \to p$ and for all $\theta \in \Theta$ there are θ_n such that $\theta_n \to \theta$ and $\limsup_n J(x_n, p_n, \theta_n) \leq J(x, p, \theta)$.

For two constants M_1, M_2 and a compact set $K \subseteq E$ we write

$$\Theta_{M_1,M_2,K} := \bigcup_{x,y\in K} \bigcup_{\alpha>1} \left\{\theta \,|\, \mathcal{I}\left(x, \partial_x \frac{\alpha}{2} d^2(x,y), \theta\right) - \Lambda\left(x, \partial_x \frac{\alpha}{2} d^2(x,y), \theta\right) \leq M_1, \right.$$
$$\left. \Lambda\left(y, \partial_y \frac{-\alpha}{2} d^2(x,y), \theta\right) - \mathcal{I}\left(y, \partial_y \frac{-\alpha}{2} d^2(x,y), \theta\right) \leq M_2 \right\}.$$

4.3 THE COMPARISON PRINCIPLE

We will state our assumption afterwards.

Theorem 4.3.1 (Comparison principle). *Consider $\mathcal{H} : E \times \mathbb{R}^d \to \mathbb{R}$ given by*

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left[\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)\right], \tag{4.3.1}$$

with Θ a metric space and Λ and $\mathcal{I}$ satisfying Assumption 4.3.3. Define the operator $\mathbf{H}f(x) := \mathcal{H}(x, \nabla f(x))$ with domain $\mathcal{D}(\mathbf{H}) = C_{cc}^\infty(E)$. Then:

(a) For any $h_1 \in C_l(E)$, $h_2 \in C_u(E)$ and $\lambda > 0$, the comparison principle holds for

$$f - \lambda \mathbf{H} f = h_1$$
$$f - \lambda \mathbf{H} f = h_1.$$

(b) For any $f_0 \in C_b(E)$, the comparison principle holds for

$$\begin{cases} \partial_t f(t,x) - \mathbf{H}f(t,\cdot)(x) = 0, & \text{if } t > 0 \\ f(0,x) = f_0(x) & \text{if } t = 0. \end{cases}$$

Remark 4.3.2 (Domain). The comparison principle holds with any domain that satisfies $C_{cc}^\infty(E) \subseteq \mathcal{D}(\mathbf{H}) \subseteq C_b^1(E)$.

We consider the following assumption.

Assumption 4.3.3. The functions $\mathcal{H}$ and $\Lambda - \mathcal{I}$ verify the following properties.

(I) The map $p \mapsto \mathcal{H}(x,p)$ is convex and $\mathcal{H}(x,0) = 0$ for every $x \in E$.

(II) The function $\theta \mapsto \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)$ is bounded from above for every x, p.

(III) There exists a containment function $\Upsilon : E \to [0,\infty]$, cf. Definition 4.2.2.

(IV) The function $\mathcal{I} - \Lambda$ is Γ–convergent in terms of (x,p), cf. Definition 4.2.3.

(V) $\forall M_1, M_2 \in \mathbb{R}$ and $\forall K$ compact, the set $\Theta_{M_1,M_2,K}$ is relatively compact.

(VI) $\forall x \in E$, $\forall p \in \mathbb{R}^d$ and for all small neighborhood of x and p, U_x V_p, there exists a continuous function $g : U_x \times V_p \to \mathbb{R}$ such that the set

$$\phi_g(y,q) = \{\theta \in \Theta \,|\, \mathcal{I}(y,q,\theta) - \Lambda(y,q,\theta) \leq g(y,q)\}$$

is non-empty and compact $\forall\, y \in U_x, q \in V_p$.

(VII) The function $\Lambda - \mathcal{I}$ verifies the continuity estimate, cf. Definition 4.2.1.

4.3.1 *Comments about the assumptions*

We give here some conditions that are often easier to show in many examples and that imply our assumptions.

Lemma 4.3.4. *Let $\mathcal{H}$ be as in (4.3.1). Suppose the following properties*

(I) *The function* $(x,p,\theta) \mapsto \Lambda(x,p,\theta)$ *is continuous and the map* $\theta \mapsto \Lambda(x,p,\theta)$ *is bounded for every* x, p.
(II) $\mathcal{I}$ *is* Γ*-convergent in terms of* (x,p).
(III) $\mathcal{I}$ *has compact sublevel sets.*
(IV) $\mathcal{I} \geq 0$ *and for every* x *and* p *there exists a* $\theta_{x,p}$ *such that* $\mathcal{I}(x,p,\theta_{x,p}) = 0$.

Then, Assumptions 4.3.3 (II),(IV) and (VI) hold.

Proof. Proof of (II): Since Λ in bounded in θ and continuous and $\mathcal{I}$ has compact sublevel sets, $\Lambda - \mathcal{I}$ is bounded in θ.
Proof of (IV): Since $\mathcal{I}$ is Γ-convergent and Λ is continuous, $\mathcal{I} - \Lambda$ is Γ-convergent.
Proof of (VI): Fix x and p. Let $\theta_{x,p}$ be such that $\mathcal{I}(x,p,\theta_{x,p}) = 0$. Then, the set

$$\phi(x,p) = \{\theta \,:\, \mathcal{I}(x,p,\theta) - \Lambda(x,p,\theta) \leq -\Lambda(x,p,\theta_{x,p})\}$$

is not empty since $\theta_{x,p} \in \phi(x,p)$. Moreover, the above set is a subset of the set

$$\{\theta \,:\, \mathcal{I}(x,p,\theta) \leq 2\|\Lambda\|_{\Theta}\},$$

that is compact since it is a sublevel set of $\mathcal{I}$. Then, it follows that $\phi(x,p)$ is relatively compact. □

Lemma 4.3.5. *Let* $\mathcal{H}$ *be as in* (4.3.1)*. Suppose the following properties*
(I) *There exists a function* $\phi : E \times \mathbb{R}^d \to \mathbb{R}$ *such that*
 (i) ϕ *is coercive in* p*, uniformly in* x *in a compact, that is*

$$\lim_{p\to\infty} \sup_{x\in K} \phi(x,p) = \infty.$$

 (ii) $\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta) \geq \phi(x,p)$ *for all* $\theta \in \Theta$.
(II) *There exists a function* $J : E \times \mathbb{R}^d \times \Theta \to \mathbb{R}$ *such that*
 (i) $J(x,p,\cdot)$ *has compact sublevel sets;*
 (ii) $\mathcal{I}(x,p,\theta) \geq J(x,p,\theta)$ *for all* $x \in E$*,* $p \in \mathbb{R}^d$ *and* $\theta \in \Theta$*;*
 (iii) $\lim_{\theta\to\infty} \frac{\Lambda(x,p,\theta)}{J(x,p,\theta)} = 0$ *for all* $x \in E$*,* $p \in \mathbb{R}^d$*.*

Then, Assumptions 4.3.3 (V) and (VII) hold.

Proof. Proof of (V): First, recall the definition of $\Theta_{M_1,M_2,K}$ as

$$\begin{aligned}\Theta_{M_1,M_2,K} &:= \\ \bigcup_{x,y\in K}\bigcup_{\alpha>1}\Big\{\theta \,|\, &\mathcal{I}\left(x,\partial_x \frac{\alpha}{2} d^2(x,y),\theta\right) - \Lambda\left(x,\partial_x \frac{\alpha}{2} d^2(x,y),\theta\right) \leq M_1, \\ &\Lambda\left(y,\partial_y \frac{-\alpha}{2} d^2(x,y),\theta\right) - \mathcal{I}\left(y,\partial_y \frac{-\alpha}{2} d^2(x,y),\theta\right) \leq M_2\Big\}.\end{aligned}$$

Let $\theta \in \Theta_{M_1,M_2,K}$ and call $p = \partial_x \frac{\alpha}{2} d^2(x,y) = \partial_y \frac{-\alpha}{2} d^2(x,y)$. Since $\Lambda(y,p,\theta) - \mathcal{I}(y,p,\theta) \geq \phi(y,p)$ for some coercive function ϕ, and $\Lambda(y,p,\theta) - \mathcal{I}(y,p,\theta) \leq M_2$, we can conclude that

$$\phi(x,p) \leq M_2.$$

By coercivity of ϕ, we can conclude that p lies in a compact set. Then, we can write that

$$\Theta_{M_1,M_2,K} \subseteq \bigcup_{x\in K_1} \bigcup_{p\in K_2} \{\theta \,|\, J(x,p,\theta) - \Lambda(x,p,\theta) \leq M_1\} := \Theta^*.$$

Then, by the fact that J grows faster then Λ in θ and that J has compact sublevel sets, we can conclude that Θ^* is compact. Then, $\Theta_{M_1,M_2,K}$ is relatively compact.
Proof of (VII): The continuity estimate follows from the coercivity of $\phi \leq \Lambda - \mathcal{I}$ as in Proposition 5.15 of [KS21]. □

4.4 CONTINUITY OF THE HAMILTONIAN

We start showing that the Hamiltonian is continuous. This is the content of the following proposition.

Proposition 4.4.1 (Continuity of the Hamiltonian). *Consider the map $\mathcal{H} : E \times \mathbb{R}^d \to \mathbb{R}$ as in* (4.3.1)*. Suppose Assumption 4.3.3. Then,*

(a) $(x,p) \mapsto \mathcal{H}(x,p)$ *is continuous;*
(b) $(x,v) \mapsto \mathcal{L}(x,v) := \sup_p \langle p, v\rangle - \mathcal{H}(x,p)$ *is lower semi-continuous;*
(c) For every x, p, there exists $\theta \in \Theta$ such that $\mathcal{H}(x,p) = \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)$.

We will use the following technical result to establish upper semicontinuity of $\mathcal{H}$.

Lemma 4.4.2 (Lemma 17.30 in [AB06]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces. Let $\phi : \mathcal{X} \to \mathcal{K}(\mathcal{Y})$, where $\mathcal{K}(\mathcal{Y})$ is the space of non-empty compact subsets of $\mathcal{Y}$. Suppose that ϕ is upper hemi-continuous, that is if $x_n \to x$ and $y_n \to y$ and $y_n \in \phi(x_n)$, then $y \in \phi(x)$.*

Let $f : Graph(\phi) \to \mathbb{R}$ be upper semi-continuous. Then the map $m(x) = \sup_{y\in\phi(x)} f(x,y)$ is upper semi-continuous.

Proof of Proposition 4.4.1. Item (c) follows by the boundness and the upper semi-continuity of $\Lambda - \mathcal{I}$ that are Assumption 4.3.3 (II) and (IV).

We establish the upper semi-continuity arguing on the basis of Lemma 4.4.2. Firstly, note that $f(x,p,\theta) = \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)$ is upper semi-continuous by Assumption 4.3.3 (IV). Moreover, by Assumption 4.3.3 (VI), for all small neighborhood of x and p there exists a continuous function g in these neighborhoods such that there exists $\theta^0(x,p)$ for which

$$\mathcal{I}(x,p,\theta^0(x,p)) - \Lambda(x,p,\theta^0(x,p)) \leq g(x,p).$$

Hence, we can write the supremum over $\theta \in \Theta$ as the supremum over $\theta \in \phi(x,p)$ where

$$\phi_g(x,p) = \overline{\{\theta \in \Theta \,|\, \mathcal{I}(x,p,\theta) - \Lambda(x,p,\theta) \leq g(x,p)\}}.$$

$\phi_g(x,p)$ is non empty, since $\theta^0(x,p) \in \phi_g(x,p)$, and it is compact for Assumption 4.3.3 (VI). We are left to show that ϕ is upper hemi-continuous. Thus, let $(x_n,p_n) \to (x,p)$ and $\theta_n \to \theta$ with $\theta_n \in \phi_g(x_n,p_n)$. We establish that $\theta \in \phi_g(x,p)$. By the definition of

$\phi_g(x_n, p_n)$, $\mathcal{I}(x_n, p_n, \theta_n) - \Lambda(x_n, p_n, \theta_n) \leq g(x,p)$. Then, by the lower semi-continuity of $\mathcal{I} - \Lambda$, we can write

$$\begin{aligned}\mathcal{I}(x,p,\theta) - \Lambda(x,p,\theta) &\leq \liminf_n \mathcal{I}(x_n, p_n, \theta_n) - \Lambda(x_n, p_n, \theta_n) \\ &\leq \liminf_n g(x_n, p_n) = g(x,p),\end{aligned}$$

which implies indeed that $\theta \in \phi_g(x,p)$. Thus, upper semi-continuity follows by an application of Lemma 4.4.2.

We prove now the lower semi–continuity of $\mathcal{H}$. Precisely, we want to show that if $(x_n, p_n) \to (x,p)$ then $\liminf_n \mathcal{H}(x_n, p_n) \geq \mathcal{H}(x,p)$. Let $\theta \in \Theta$ be such that $\mathcal{H}(x,p) = \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)$. By Assumption 4.3.3 (IV), the function $\mathcal{I} - \Lambda$ is Γ–convergent in the sense of Definition 4.2.3. That means that there exist θ_n converging to θ such that $\limsup_n \mathcal{I}(x_n, p_n, \theta_n) - \Lambda(x_n, p_n, \theta_n) \leq \mathcal{I}(x,p,\theta) - \Lambda(x,p,\theta)$. Therefore,

$$\begin{aligned}\liminf_n \mathcal{H}(x_n, p_n) &\geq \liminf_n \left[\Lambda(x_n, p_n, \theta_n) - \mathcal{I}(x_n, p_n, \theta_n)\right] \\ &= -\limsup_n \left[\mathcal{I}(x_n, p_n, \theta_n) - \Lambda(x_n, p_n, \theta_n)\right] \\ &\geq -\left[\mathcal{I}(x,p,\theta) - \Lambda(x,p,\theta)\right] \\ &= \mathcal{H}(x,p),\end{aligned}$$

establishing the lower semi-continuity of $\mathcal{H}$ and, hence, the continuity. Moreover, since the Lagrangian $\mathcal{L}$ is the Legendre transform of $\mathcal{H}$, it is lower semi-continuous. □

4.5 PROOF OF THE MAIN THEOREM

In this section we establish Theorem 4.3.1. To prove the comparison principle for $f - \lambda \mathbf{H} f = h$ and $\partial_t f - \mathbf{H} f = 0$, we relate them to a set of Hamilton–Jacobi–Bellman equation with Hamiltonians constructed from $\mathbf{H}$. To do this, we introduce two operators $H_\dagger$ and $H_\ddagger$ that will be respectively an upper and lower bound for $\mathbf{H}$. The two new Hamiltonians, defined in Subsection 4.5.1, are constructed in terms of a containment function Υ that allows us to restrict our analysis to a compact set. Schematically, we will establish the diagram in Figure 8 on pag. 62.

The arrows will be established in Subsection 4.5.1. Finally, we will establish the comparison principle for $H_\dagger$ and $H_\ddagger$ in Subsection 4.5.2. The combination of these two results imply the comparison principle for $\mathbf{H}$ as shown in the following.

Proof of Theorem 4.3.1. We only prove the first item. The proof for the time-dependent case follows the same lines. Fix $h_1, h_2 \in C_b(E)$ and $\lambda > 0$. Let u, v be a viscosity sub- and supersolution to $f - \lambda \mathbf{H} f = h_1$ and $f - \lambda \mathbf{H} f = h_2$ respectively. By Lemma 4.5.3 proven in Section 4.5.1, u and v are a sub- and supersolution to $f - \lambda H_\dagger f = h_1$ and $f - \lambda H_\ddagger f = h_2$ respectively. Thus $\sup_E u - v \leq \sup_E h_1 - h_2$ by Proposition 4.5.4 of Section 4.5.2. □

4.5.1 *Auxiliary operators*

In this section, we repeat the definition of $\mathbf{H}$, and introduce the operators $H_\dagger$ and $H_\ddagger$.

Definition 4.5.1. The operator $\mathbf{H} \subseteq C_b^1(E) \times C_b(E)$ has domain $\mathcal{D}(\mathbf{H}) = C_{cc}^\infty(E)$ and satisfies $\mathbf{H}f(x) = \mathcal{H}(x, \nabla f(x))$, where $\mathcal{H}$ is the map

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left[\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)\right].$$

We proceed by introducing $H_\dagger$ and $H_\ddagger$ serving as natural upper and lower bounds to $\mathbf{H}$. Recall Assumption (III) and the constant

$$C_\Upsilon := \sup_\theta \sup_x \Lambda(x, \nabla\Upsilon(x), \theta) - \mathcal{I}(x, \nabla\Upsilon(x), \theta).$$

Denote by $C_\ell^\infty(E)$ the set of smooth functions on E that have a lower bound and by $C_u^\infty(E)$ the set of smooth functions on E that have an upper bound. The definitions below are motivated by the convexity of the map $p \mapsto \mathcal{H}(x,p)$.

Definition 4.5.2 (The operators $H_\dagger$ and $H_\ddagger$)**.** For $f \in C_\ell^\infty(E)$ and $\varepsilon \in (0,1)$ set

$$\begin{aligned} f_\dagger^\varepsilon &:= (1-\varepsilon) f + \varepsilon\Upsilon \\ H_{\dagger,f}^\varepsilon(x) &:= (1-\varepsilon)\mathcal{H}(x, \nabla f(x)) + \varepsilon C_\Upsilon. \end{aligned}$$

and set

$$H_\dagger := \left\{ (f_\dagger^\varepsilon, H_{\dagger,f}^\varepsilon) \,\middle|\, f \in C_\ell^\infty(E), \varepsilon \in (0,1) \right\}.$$

For $f \in C_u^\infty(E)$ and $\varepsilon \in (0,1)$ set

$$\begin{aligned} f_\ddagger^\varepsilon &:= (1+\varepsilon) f - \varepsilon\Upsilon \\ H_{\ddagger,f}^\varepsilon(x) &:= (1+\varepsilon)\mathcal{H}(x, \nabla f(x)) - \varepsilon C_\Upsilon. \end{aligned}$$

and set

$$H_\ddagger := \left\{ (f_\ddagger^\varepsilon, H_{\ddagger,f}^\varepsilon) \,\middle|\, f \in C_u^\infty(E), \varepsilon \in (0,1) \right\}.$$

The operator $\mathbf{H}$ is related to $H_\dagger, H_\ddagger$ by the following Lemma whose proof is standard and can be found for example in [KS21]. We include it for completeness.

Lemma 4.5.3. *Fix $\lambda > 0$ and $h \in C_b(E)$.*

(a) Every subsolution to $f - \lambda\mathbf{H}f = h$ is also a subsolution to $f - \lambda H_\dagger f = h$.
(b) Every supersolution to $f - \lambda\mathbf{H}f = h$ is also a supersolution to $f - \lambda H_\ddagger f = h$.
(c) Every subsolution to $\partial_t f - \mathbf{H}f = 0$ is also a subsolution to $\partial_t f - H_\dagger f = 0$.
(d) Every supersolution to $\partial_t f - \mathbf{H}f = 0$ is also a supersolution to $\partial_t f - H_\ddagger f = 0$.

Proof. We only prove (a) as the other claims can be carried out analogously. Fix $\lambda > 0$ and $h \in C_b(E)$. Let u be a subsolution to $f - \lambda \mathbf{H} f = h$. We prove it is also a subsolution to $f - \lambda H_\dagger f = h$.

Fix $\varepsilon > 0$ and $f \in C_\ell^\infty(E)$ and let $(f_\dagger^\varepsilon, H_{\dagger,f}^\varepsilon) \in H_\dagger$ as in Definition 4.5.2. We will prove that there are $x_n \in E$ such that

$$\lim_{n\to\infty} \left(u - f_\dagger^\varepsilon\right)(x_n) = \sup_{x \in E} \left(u(x) - f_\dagger^\varepsilon(x)\right), \tag{4.5.1}$$

$$\limsup_{n\to\infty} \left[u(x_n) - \lambda H_{\dagger,f}^\varepsilon(x_n) - h(x_n)\right] \leq 0. \tag{4.5.2}$$

As the function $[u - (1-\varepsilon) f]$ is bounded from above and $\varepsilon \Upsilon$ has compact sublevel-sets, the sequence x_n along which the first limit is attained can be assumed to lie in the compact set

$$K := \left\{ x \,|\, \Upsilon(x) \leq \varepsilon^{-1} \sup_x \left(u(x) - (1-\varepsilon) f(x)\right) \right\}.$$

Set $M = \varepsilon^{-1} \sup_x \left(u(x) - (1-\varepsilon) f(x)\right)$. Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a smooth increasing function such that

$$\gamma(r) = \begin{cases} r & \text{if } r \leq M, \\ M+1 & \text{if } r \geq M+2. \end{cases}$$

Denote by f_ε the function on E defined by

$$f_\varepsilon(x) := \gamma\left((1-\varepsilon) f(x) + \varepsilon \Upsilon(x)\right).$$

By construction f_ε is smooth and constant outside of a compact set and thus lies in $\mathcal{D}(H) = C_{cc}^\infty(E)$. As u is a viscosity subsolution for $f - \lambda H f = h$ there exists a sequence $x_n \in K \subseteq E$ (by our choice of K) with

$$\lim_n \left(u - f_\varepsilon\right)(x_n) = \sup_x \left(u(x) - f_\varepsilon(x)\right),$$

$$\limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \leq 0.$$

As f_ε equals $f_\dagger^\varepsilon$ on K, we have from (5.5.3) that also

$$\lim_n \left(u - f_\dagger^\varepsilon\right)(x_n) = \sup_{x \in E} \left(u(x) - f_\dagger^\varepsilon(x)\right),$$

establishing (4.5.1). Convexity of $p \mapsto \mathcal{H}(x,p)$ yields for arbitrary points $x \in K$ the estimate

$$\begin{aligned} \mathbf{H} f_\varepsilon(x) &= \mathcal{H}(x, \nabla f_\varepsilon(x)) \\ &\leq (1-\varepsilon)\mathcal{H}(x, \nabla f(x)) + \varepsilon \mathcal{H}(x, \nabla \Upsilon(x)) \\ &\leq (1-\varepsilon)\mathcal{H}(x, \nabla f(x)) + \varepsilon C_\Upsilon = H_{\dagger,f}^\varepsilon(x). \end{aligned}$$

Combining this inequality with (5.5.4) yields

$$\begin{aligned}\limsup_n \left[u(x_n) - \lambda H^{\varepsilon}_{\dagger,f}(x_n) - h(x_n)\right] &\leq \limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \\ &\leq 0,\end{aligned}$$

establishing (4.5.2). This concludes the proof. □

4.5.2 *The comparison principle*

In the following we prove the comparison principle for the operators $H_\dagger$ and $H_\ddagger$.

Proposition 4.5.4. *Fix $\lambda > 0$ and $h_1, h_2 \in C_b(E)$. The following holds:*

(a) Let u be a viscosity subsolution to $f - \lambda H_\dagger f = h_1$ and let v be a viscosity supersolution to $f - \lambda H_\ddagger f = h_2$. Then we have $\sup_x u(x) - v(x) \leq \sup_x h_1(x) - h_2(x)$.

(b) Let u be a viscosity subsolution to $\partial_t f - H_\dagger f = 0$ and let v be a viscosity supersolution to $\partial_t f - H_\ddagger f = 0$. Then we have $\sup_{t\in[0,T],x} u(x,t) - v(x,t) \leq \sup_x u(x,0) - v(x,0)$ for all $T > 0$.

The strategy of the proof is the same for both equations. In both cases, the aim is to prove that it is possible to bound the difference of the Hamiltonians, in well-chosen sequences of points, by using the continuity estimate for $\Lambda - \mathcal{I}$. This is the content of Proposition 4.5.7. With this aim, we use two variants of a classical estimate, that was proven e.g. in [CIL92, Proposition 3.7], given respectively in Lemma 4.5.5 for the stationary equation and Lemma 4.5.6 for the evolutionary case.

Lemma 4.5.5. *Let u be bounded and upper semi-continuous, let v be bounded and lower semi-continuous and let Υ be a containment function.*

Fix $\varepsilon > 0$. For every $\alpha > 0$ there exists $(x_\alpha, y_\alpha) = (x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \in E \times E$ such that

$$\begin{aligned}&\frac{u(x_\alpha)}{1-\varepsilon} - \frac{v(y_\alpha)}{1+\varepsilon} - \frac{\alpha}{2}d^2(x_\alpha, y_\alpha) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_\alpha) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_\alpha) \\ &\qquad = \sup_{x,y\in E}\left\{\frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \frac{\alpha}{2}d^2(x,y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y)\right\}.\end{aligned}$$

Additionally, for every $\varepsilon > 0$ we have that

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

(b) All limit points of $\{(x_\alpha, y_\alpha)\}_{\alpha>0}$ as $\alpha \to \infty$ are of the form (z,z) and for these limit points we have

$$\frac{u(z)}{1-\varepsilon} - \frac{v(z)}{1+\varepsilon} - \frac{2\varepsilon}{1-\varepsilon^2}\Upsilon(z) = \sup_{x\in E}\left\{\frac{u(x)}{1-\varepsilon} - \frac{v(x)}{1+\varepsilon} - \frac{2\varepsilon}{1-\varepsilon^2}\Upsilon(x)\right\}.$$

(c) We have

$$\lim_{\alpha\to\infty} \alpha d^2(x_\alpha, y_\alpha) = 0.$$

Lemma 4.5.6. *Let u be bounded and upper semi-continuous, let v be bounded and lower semi-continuous and let Υ be a containment function. Fix $\varepsilon > 0$, $\beta > 0$ and $T > 0$. For every $\alpha > 0$ and $\gamma > 0$ there exists*

$$(x_{\alpha,\gamma,\varepsilon,\beta}, t_{\alpha,\gamma,\varepsilon,\beta}, y_{\alpha,\gamma,\varepsilon,\beta}, s_{\alpha,\gamma,\varepsilon,\beta}) \in E \times [0,T], \times E \times [0,T],$$

denoted by $(x_{\alpha,\gamma}, t_{\alpha,\gamma}, y_{\alpha,\gamma}, s_{\alpha,\gamma})$, such that

$$\begin{aligned}
&\frac{u(t_{\alpha,\gamma}, x_{\alpha,\gamma})}{1-\varepsilon} - \frac{v(s_{\alpha,\gamma}, y_{\alpha,\gamma})}{1+\varepsilon} - \frac{\alpha}{2} d^2(x_{\alpha,\gamma}, y_{\alpha,\gamma}) - \frac{\gamma}{2}(s_{\alpha,\gamma} - t_{\alpha,\gamma})^2 \\
&- \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\gamma}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\gamma}) - \frac{\beta}{2}(t_{\alpha,\gamma} + s_{\alpha,\gamma}) + \beta T \\
&= \sup_{s,t\in[0,T],x,y} \left\{ \frac{u(t,x)}{1-\varepsilon} - \frac{v(s,y)}{1+\varepsilon} - \frac{\alpha}{2} d^2(x,y) - \frac{\gamma}{2}(s-t)^2 \right. \\
&\qquad\qquad \left. - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) - \frac{\beta}{2}(t+s) + \beta T \right\}.
\end{aligned}$$

Additionally, for every $\varepsilon > 0$ and $\beta > 0$ we have that

(a) For any $\gamma > 0$,
- *(i) the set $\{x_{\alpha,\gamma}, y_{\alpha,\gamma} \,|\, \alpha > 0\}$ is relatively compact in E,*
- *(ii) $\lim_{\alpha\to\infty} \alpha d^2(x_{\alpha,\gamma}, y_{\alpha,\gamma}) = 0$,*
- *(iii) all limit points of $\{(x_{\alpha,\gamma}, y_{\alpha,\gamma}, t_{\alpha,\gamma}, s_{\alpha,\gamma})\}_{\alpha>0}$ as $\alpha \to \infty$ are of the form $(z_\gamma, z_\gamma, t_\gamma, s_\gamma)$.*

(b) Let $(z_\gamma, z_\gamma, t_\gamma, s_\gamma)$ be a limit point as in (a)(iii). Then
- *(i) the set $\{z_\gamma \,|\, \gamma > 0\}$ is relatively compact in E,*
- *(ii) all limit point of $\{(z_\gamma, z_\gamma, t_\gamma, s_\gamma)\}_{\gamma>0}$ as $\gamma \to \infty$ are of the form (z, z, w, w) and for these limit points we have*

$$\begin{aligned}
&\frac{u(w,z)}{1-\varepsilon} - \frac{v(w,z)}{1+\varepsilon} - \frac{2\varepsilon}{1-\varepsilon^2}\Upsilon(z) + \beta(T-w) \\
&\qquad = \sup_{x\in E, t\in[0,T]} \left\{ \frac{u(t,x)}{1-\varepsilon} - \frac{v(t,x)}{1+\varepsilon} - \frac{2\varepsilon}{1-\varepsilon^2}\Upsilon(x) + \beta(T-t) \right\}.
\end{aligned}$$

In the following proposition we prove the continuity estimate for $\mathcal{H}$ by using the continuity estimate of $\Lambda - \mathcal{I}$.

Proposition 4.5.7. *Consider $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ found in Lemma 4.5.5 or Lemma 4.5.6 (for which we fix γ and β) and denote $p^1_{\alpha,\varepsilon} := \alpha \mathrm{d}_x \frac{1}{2} d^2(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})$ and $p^2_{\alpha,\varepsilon} := -\alpha \mathrm{d}_y \frac{1}{2} d^2(x_{\alpha,\varepsilon}, \cdot)(y_{\alpha,\varepsilon})$. Suppose that*

$$\inf_\alpha \mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) > -\infty \tag{4.5.3}$$

and

$$\sup_\alpha \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}) < \infty. \tag{4.5.4}$$

Then, for all $\varepsilon > 0$ there exists a sequence $\alpha(\varepsilon) \to \infty$, such that

$$\liminf_{\varepsilon\to 0} \liminf_{\alpha\to\infty} \mathcal{H}(x_\alpha, p^1_{\alpha,\varepsilon}) - \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}) \le 0. \tag{4.5.5}$$

Proof. We only prove the statement for $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$ found in Lemma 4.5.5. The proof in the context of Lemma 4.5.6 is analogous. The proof is given in two steps. We sketch the steps, before giving full proof.

Step 1: We will show that there are controls $\theta_{\alpha,\varepsilon}$ such that

$$\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) = \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}).$$

As a consequence we have

$$\begin{aligned}\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) - \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}) \leq \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \Lambda(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) \\ + \mathcal{I}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}). \quad (4.5.6)\end{aligned}$$

For establishing (4.5.5), it is sufficient to bound the differences in (4.5.6) by using Assumption 4.3.3(VII).

Step 2: We verify the conditions to apply the continuity estimate, Assumption 4.3.3 (VII) which then concludes the proof.

Proof of Step 1: Recall that $\mathcal{H}(x,p)$ is given by

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \left[\Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)\right].$$

Since $\Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \cdot) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \cdot) : \Theta \to \mathbb{R}$ is upper semi-continuous and bounded by Assumption 4.3.3 (IV) and 4.3.3(II), there exists an optimizer $\theta_{\alpha,\varepsilon} \in \Theta$ such that

$$\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) = \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}). \quad (4.5.7)$$

Choosing the same point in the supremum of the second term $\mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon})$, we obtain for all $\varepsilon > 0$ and $\alpha > 0$ the estimate

$$\begin{aligned}\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) - \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}) \leq \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \Lambda(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) \\ + \mathcal{I}(y_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}).\end{aligned}$$

Proof of Step 2: We will construct for each $\varepsilon > 0$ a sequence $\alpha = \alpha(\varepsilon) \to \infty$ such that the collection $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon})$ is fundamental for $\Lambda - \mathcal{I}$ in the sense of Definition 4.2.1. We thus need to verify for each $\varepsilon > 0$

(i)

$$\inf_\alpha \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) > -\infty,$$

(ii)

$$\sup_\alpha \Lambda(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) < \infty$$

(iii) The set of controls $\theta_{\alpha,\varepsilon}$ is relatively compact.

We will first establish (i) and (ii) for all α. Then, (iii) will follow from (i) and (ii) and Assumption 4.3.3(V).

By (4.5.3) and (4.5.7),

$$-\infty < \inf_{\alpha} \mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) = \inf_{\alpha} \Lambda(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon})$$

establishing (i).

By (4.5.4),

$$\sup_{\alpha} \Lambda(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) - \mathcal{I}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon}) < \sup_{\alpha} \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}) < \infty$$

implying (ii). □

Proof of Proposition 4.5.4. Proof of (a). Fix $\lambda > 0$ and $h_1, h_2 \in C_b(E)$. Let u be a viscosity subsolution and v be a viscosity supersolution of $f - \lambda H_\dagger f = h_1$ and $f - \lambda H_\ddagger f = h_2$ respectively. For any $\varepsilon > 0$ and any $\alpha > 0$, define the map $\Phi_{\alpha,\varepsilon} : E \times E \to \mathbb{R}$ by

$$\Phi_{\alpha,\varepsilon}(x,y) := \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \frac{\alpha}{2} d^2(x,y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y).$$

Let $\varepsilon > 0$. By Lemma 4.5.5, there is a compact set $K_\varepsilon \subseteq E$ and there exist points $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in K_\varepsilon$ such that

$$\Phi_{\alpha,\varepsilon}(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = \sup_{x,y\in E} \Phi_{\alpha,\varepsilon}(x,y), \tag{4.5.8}$$

and

$$\lim_{\alpha\to\infty} \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = 0.$$

For all α it follows that

$$\begin{aligned}
\sup_E (u - v) &= \lim_{\varepsilon\to 0} \sup_{x\in E} \frac{u(x)}{1-\varepsilon} - \frac{v(x)}{1+\varepsilon} \\
&\leq \liminf_{\varepsilon\to 0} \sup_{x,y\in E} \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \frac{\alpha}{2} d^2(x,y) \\
&\qquad\qquad - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) \\
&= \liminf_{\varepsilon\to 0} \frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \frac{\alpha}{2} d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \\
&\qquad\qquad - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\varepsilon}) \\
&\leq \liminf_{\varepsilon\to 0} \left[\frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} \right].
\end{aligned} \tag{4.5.9}$$

At this point, we want to use the sub- and supersolution properties of u and v. Define the test functions $\varphi_1^{\varepsilon,\alpha} \in \mathcal{D}(H_\dagger), \varphi_2^{\varepsilon,\alpha} \in \mathcal{D}(H_\ddagger)$ by

$$\varphi_1^{\varepsilon,\alpha}(x) := (1-\varepsilon)\left[\frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} + \frac{\alpha}{2}d^2(x, y_{\alpha,\varepsilon}) + \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) + \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\varepsilon})\right] + (1-\varepsilon)(x - x_{\alpha,\varepsilon})^2,$$

$$\varphi_2^{\varepsilon,\alpha}(y) := (1+\varepsilon)\left[\frac{u_1(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{\alpha}{2}d^2(x_{\alpha,\varepsilon}, y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y)\right] - (1+\varepsilon)(y - y_{\alpha,\varepsilon})^2.$$

Using (4.5.8), we find that $u - \varphi_1^{\varepsilon,\alpha}$ attains its supremum at $x = x_{\alpha,\varepsilon}$, and thus

$$\sup_E(u - \varphi_1^{\varepsilon,\alpha}) = (u - \varphi_1^{\varepsilon,\alpha})(x_{\alpha,\varepsilon}).$$

Denote $p^1_{\alpha,\varepsilon} := \alpha \mathrm{d}_x \frac{1}{2}d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$. By our addition of the penalization $(x - x_{\alpha,\varepsilon})^2$ to the test function, the point $x_{\alpha,\varepsilon}$ is in fact the unique optimizer, and we obtain from the subsolution inequality that

$$u(x_{\alpha,\varepsilon}) - \lambda\left[(1-\varepsilon)\mathcal{H}\left(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}\right) + \varepsilon C_\Upsilon\right] \leq h_1(x_{\alpha,\varepsilon}). \tag{4.5.10}$$

With a similar argument for u_2 and $\varphi_2^{\varepsilon,\alpha}$, we obtain by the supersolution inequality that

$$v(y_{\alpha,\varepsilon}) - \lambda\left[(1+\varepsilon)\mathcal{H}\left(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}\right) - \varepsilon C_\Upsilon\right] \geq h_2(y_{\alpha,\varepsilon}), \tag{4.5.11}$$

where $p^2_{\alpha,\varepsilon} := -\alpha \mathrm{d}_y \frac{1}{2}d^2(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$. With that, estimating further in (4.5.9) leads to

$$\sup_E(u - v) \leq \liminf_{\varepsilon\to 0}\liminf_{\alpha\to\infty}\left[\frac{h_1(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{h_2(y_{\alpha,\varepsilon})}{1+\varepsilon} + \frac{\varepsilon}{1-\varepsilon}C_\Upsilon + \frac{\varepsilon}{1+\varepsilon}C_\Upsilon + \lambda\left[\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) - \mathcal{H}(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon})\right]\right].$$

Note that, by the subsolution inequality (4.5.10),

$$-\infty < \frac{1}{\lambda}\inf_E(u - h_1) \leq (1-\varepsilon)\mathcal{H}(x_{\alpha,\varepsilon}, p^1_{\alpha,\varepsilon}) + \varepsilon C_\Upsilon,$$

and by the supersolution inequality (4.5.11),

$$(1+\varepsilon)\mathcal{H}\left(y_{\alpha,\varepsilon}, p^2_{\alpha,\varepsilon}\right) - \varepsilon C_\Upsilon \leq \frac{1}{\lambda}\sup_E(v - h_2) < \infty.$$

Thus, comparison principle follows from Proposition 4.5.7.

Proof of (b). Let u be a subsolution for $H_\dagger$, and v a supersolution for $H_\ddagger$. Let $T > 0$ be fixed.

For any $\beta > 0$, we have

$$\sup_{t\in[0,T],x} u(t,x) - v(t,x) \leq \sup_{t\in[0,T],x} u(t,x) - v(t,x) - \beta t + \beta T$$

We next incorporate our Lyapunov type functions

$$\sup_{t\in[0,T],x} u(t,x) - v(t,x) - \beta t + \beta T = \lim_{\varepsilon\downarrow 0} \sup_{t\in[0,T],x} \frac{u(t,x)}{1-\varepsilon} - \frac{v(t,x)}{1+\varepsilon} - \frac{2\varepsilon}{1-\varepsilon^2}\Upsilon(x) - \beta t + \beta T.$$

Thus, for any $\varepsilon > 0$, $\alpha, \gamma > 0$, we have

$$\begin{aligned} &\sup_{t\in[0,T],x} \frac{u(t,x)}{1-\varepsilon} - \frac{v(t,x)}{1+\varepsilon} - \frac{2\epsilon}{1-\epsilon^2}\Upsilon(x) - \beta t + \beta T \\ &\leq \sup_{s,t\in[0,T],x,y} \frac{u(t,x)}{1-\varepsilon} - \frac{v(s,y)}{1+\varepsilon} - \frac{\alpha}{2}d^2(x,y) - \frac{\gamma}{2}(s-t)^2 \\ &\qquad - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) - \frac{\beta}{2}(t+s) + \beta T. \end{aligned} \tag{4.5.12}$$

By Lemma 4.5.6, there are

$$(x_{\alpha,\gamma}, t_{\alpha,\gamma}, y_{\alpha,\gamma}, s_{\alpha,\gamma}) = (x_{\varepsilon,\beta,\alpha,\gamma}, t_{\varepsilon,\beta,\alpha,\gamma}, y_{\varepsilon,\beta,\alpha,\gamma}, s_{\varepsilon,\beta,\alpha,\gamma})$$

optimizing the supremum on the right-hand side and such that

$$\frac{\alpha}{2}d^2(x_{\alpha,\gamma}, y_{\alpha,\gamma}) \to 0 \quad \text{as } \alpha \to \infty.$$

First assume $t_{\alpha,\gamma}, s_{\alpha,\gamma} > 0$. We will aim for a contradiction.
Define the functions $f^{\dagger}_{\varepsilon,\beta,\alpha}(x) \in D(H_\dagger)$ and $f^{\ddagger}_{\varepsilon,\beta,\alpha}(y) \in D(H_\ddagger)$ by

$$\begin{aligned} f^\dagger(x) &= (1-\varepsilon)\left[\frac{v(s_{\alpha,\gamma}, y_{\alpha,\gamma})}{1+\varepsilon} + \frac{\alpha}{2}d^2(x, y_{\alpha,\gamma}) + \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\gamma}) - \beta T\right] \\ &\quad + \varepsilon\Upsilon(x), \\ f^\ddagger(y) &= (1+\varepsilon)\left[\frac{u(t_{\alpha,\gamma}, x_{\alpha,\gamma})}{1-\varepsilon} - \frac{\alpha}{2}d^2(x_{\alpha,\gamma}, y) - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\gamma}) + \beta T\right] \\ &\quad - \varepsilon\Upsilon(y). \end{aligned}$$

Observe that

$$\begin{aligned} u(t_{\alpha,\gamma}, x_{\alpha,\gamma}) - f^\dagger(x_{\alpha,\gamma}) - h_1(t_{\alpha,\gamma}) &= \sup_{t\in[0,T],x} u(t,x) - f^\dagger(x) - h_1(t), \\ v(s_{\alpha,\gamma}, y_{\alpha,\gamma}) - f^\ddagger(y_{\alpha,\gamma}) - h_2(s_{\alpha,\gamma}) &= \inf_{t\in[0,T],y} v(s,y) - f^\ddagger(y) - h_2(s), \end{aligned}$$

where $h_1(t) = (1-\varepsilon)(\frac{\beta}{2}(t+s_{\alpha,\gamma}) + \frac{\gamma}{2}(t-s_{\alpha,\gamma})^2)$ and $h_2(s) = (1+\varepsilon)(-\frac{\beta}{2}(t_{\alpha,\gamma}+s) - \frac{\gamma}{2}(t_{\alpha,\gamma}-s)^2)$.

Then, using the sub and supersolution inequalities, dividing them by $(1-\varepsilon)$ and $(1+\varepsilon)$ respectively, we find

$$\gamma(t_{\alpha,\gamma} - s_{\alpha,\gamma}) + \frac{\beta}{2} - \mathcal{H}(x_{\alpha,\gamma}, \alpha \mathrm{d}_x \frac{1}{2} d^2(\cdot, y_{\alpha,\gamma})(x_{\alpha,\gamma})) - \frac{\varepsilon}{1-\varepsilon} c_\Upsilon \leq 0,$$
$$\gamma(t_{\alpha,\gamma} - s_{\alpha,\gamma}) - \frac{\beta}{2} - \mathcal{H}(y_{\alpha,\gamma}, -\alpha \mathrm{d}_y \frac{1}{2} d^2(x_{\alpha,\gamma}, \cdot)(y_{\alpha,\gamma})) + \frac{\varepsilon}{1+\varepsilon} c_\Upsilon \geq 0.$$

Combining the two equations yields

$$\beta \leq \mathcal{H}(x_{\alpha,\gamma}, \alpha \mathrm{d}_x \frac{1}{2} d^2(\cdot, y_{\alpha,\gamma})(x_{\alpha,\gamma})) - \mathcal{H}(y_{\alpha,\gamma}, -\alpha \mathrm{d}_y \frac{1}{2} d^2(x_{\alpha,\gamma}, \cdot)(y_{\alpha,\gamma})) + \frac{2\varepsilon}{1-\varepsilon^2} c_\Upsilon$$

sending $\alpha \to \infty$ and $\varepsilon \to 0$, and using Proposition 4.5.7, we get a contradiction for small ε as $\beta > 0$. So it holds that for small ε, large α and all $\gamma > 0$, we have $t_{\alpha,\gamma} = 0$ or $s_{\alpha,\gamma} = 0$.

Proceeding from equation (4.5.12), we get

$$\begin{aligned}
&\sup_{t\in[0,T],x} u(t,x) - v(t,x) \\
&\quad \leq \sup_{t\in[0,T],x} u(t,x) - v(t,x) - \beta t + \beta T \\
&\quad \leq \lim_{\varepsilon\downarrow 0} \sup_{s,t\in[0,T],x,y} \frac{u(t,x)}{1-\varepsilon} - \frac{v(s,y)}{1+\varepsilon} - \frac{\alpha}{2} d^2(x,y) - \frac{\gamma}{2}(s-t)^2 \\
&\qquad\qquad - \frac{\varepsilon}{1-\varepsilon}\Upsilon(x) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y) - \frac{\beta}{2}(t+s) + \beta T \\
&\quad \leq \lim_{\varepsilon\downarrow 0} \lim_{\gamma\to\infty} \lim_{\alpha\to\infty} \frac{u(t_{\alpha,\gamma}, x_{\alpha,\gamma})}{1-\varepsilon} - \frac{v(s_{\alpha,\gamma}, y_{\alpha,\gamma})}{1+\varepsilon} \\
&\qquad\qquad + \frac{\varepsilon}{1-\varepsilon}\Upsilon(x_{\alpha,\gamma}) - \frac{\varepsilon}{1+\varepsilon}\Upsilon(y_{\alpha,\gamma}) - \frac{\beta}{2}(t_{\alpha,\gamma} + s_{\alpha,\gamma}) + \beta T \\
&\quad \leq \sup_x u(0,x) - v(0,x) + \beta T
\end{aligned}$$

where we used that u is upper semi-continuous, v is lower semi-continuous, and Lemma 4.5.6 (b)(ii) and the fact that $t_{\alpha,\gamma} = 0$ or $s_{\alpha,\gamma} = 0$. As $\beta > 0$ was arbitrary, we conclude.

□

4.6 EXAMPLES OF HAMILTONIANS

The purpose of this section is to showcase that the method introduced in this paper is versatile enough to capture interesting examples that could not be treated before. Before introducing two entirely new examples, we want to emphasize that our approach enables us to address a wider range of examples compared to previous works.

4.6.1 *Comparison with previous works*

First of all, we can address some classic Hamiltonians such as the control Hamiltonian given for example in [BC97, Theorem 3.1]. It is indeed easy to show that the following Hamiltonian

$$\mathcal{H}(x,p) = \sup_{a\in A} \left\{ -f(x,a)\cdot p - l(x,a) \right\},$$

with A a compact set, f a Lipschitz continuous function and l a non-negative continuous function that admits a modulus of continuity, verifies our Assumption 4.3.3.

As explained in the introduction, cases in which $\mathcal{I}$ is not bounded or not continuous and Λ is not Lipschitz or not coercive are not covered by classical works such as [BC97]. In [KS21], the authors instead treat these cases, but keeping the cost function $\mathcal{I}$ independent of p. We would like to emphasize that while many of our assumptions are implied by the assumptions in [KS21], it is not straightforward and clear whether our assumption (V) can be derived from the assumptions in [KS21] (in particular by their assumption ($\mathcal{I}3$)). The inclusion of momentum in the cost function adds an element of complexity, making it challenging to generalize the previous assumption that does not account for momentum. Nevertheless, our approach includes all the examples addressed in [KS21] and more. Indeed, it is possible to prove, by using Lemma 4.3.4 and Lemma 4.3.5, that the examples showed in [KS21] satisfy our assumptions.

Moreover, we are also able to consider some cases of Hamiltonians with unbounded (in terms of θ) internal Hamiltonian Λ. This leads to a situation where Assumption $(\Lambda 4)$ in [KS21] fails. In our case, it may work. The key distinction in our approach, enabling the success of this class of examples while it remained unaddressed in [KS21], lies in our treatment of Λ and $\mathcal{I}$ in an integrated whole $\Lambda - \mathcal{I}$ rather than separately and with two different types of assumptions. In this case, if Λ is not bounded in terms of θ, but the composite $\Lambda - \mathcal{I}$ is bounded, our assumptions hold. Consequently, in contrast to Assumption $(\Lambda 4)$ in [KS21], which solely considers the boundedness of Λ and subsequently fails in this scenario, our Assumption 4.3.3 (II) instead requires the boundedness of $\Lambda - \mathcal{I}$ and consequently holds.

Finally, we want to mention that, in contrast to [KS21], we are able to cover cases where the cost function $\mathcal{I}$ depends on momenta p. For instance, in Chapter 3 or in [Pop18], it is proved comparison principle with "ad hoc" proofs involving coercivity or Lipschitz estimates and optimization problems, for an Hamilton–Jacobi–Bellman equation with an Hamiltonian of the type as in (4.1.3). These examples does not fall into the cases of [KS21] due to the presence of p in $\mathcal{I}$. Indeed, assumptions as $(\mathcal{I}5)$ and $(\mathcal{I}4)$ in [KS21] are not satisfied and quite challenging to modify to incorporate the momenta p. Our assumptions are instead satisfied.

We also want to mention that in [FK06] it is possible to find some examples of Hamiltonians with cost function depending on p. We confidently assert that our results can cover all these examples.

We can conclude, then, that our work can be used for a large class of examples including examples that fall into previously treated theories, as the ones mentioned above, as well as those that have been addressed by means of "ad hoc" proofs, as the examples in Chapter 3 and [Pop18].

4.6.2 *Comparison principle for the Hamiltonian of Section 3.2*

We can now conclude the discussion started in Chapter 3.2. We saw that to prove large deviations for the process described in Chapter 3.2, it is necessary to prove the comparison principle for a Hamilton-Jacobi equation with Hamiltonian given by

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta}\left[\int_{F_M} V(y;x,p)\,d\theta(y) + \inf_{\varphi\in C^2(F_M)}\int_{F_M} e^{-\varphi}L_{x,p}e^{\varphi}\,d\theta\right], \tag{4.6.1}$$

with $\Theta = \mathcal{P}(F_M)$ and V and $L_{x,p}$ as in (3.2.3) and (3.2.4).

Note that the Hamiltonian is of the form as (4.3.1). We can then prove the comparison principle by using Theorem 4.3.1. This is the content of the following theorem.

Theorem 4.6.1 (Comparison principle)**.** *Consider $\mathcal{H} : E \times \mathbb{R}^l \to \mathbb{R}$ as in (4.6.1). Suppose Assumption 3.2.3 and Assumption 3.2.4. Then the comparison principle for*

$$f - \lambda\mathcal{H}(x,\nabla f(x)) = h \quad \textit{and} \quad \partial_t f - \mathcal{H}(x,\nabla f(x)) = 0$$

holds.

Proof. To prove the comparison principle we firstly mention that $\mathcal{H}(x,p)$ is of the form (4.3.1) with $\Theta = \mathcal{P}(F_M)$ and

$$\Lambda(x,p,\theta) = \int_{F_M} V(y;x,p)\,d\theta(y),$$

and

$$\mathcal{I}(x,p,\theta) = -\inf_{u\in C^2(F_M),\inf u>0}\int_{F_M}\frac{L_{x,p}u}{u}\,d\theta = -\inf_{\varphi\in C^2(F_M)}\int_{F_M} e^{-\varphi}L_{x,p}e^{\varphi}\,d\theta.$$

We can then apply Theorem 4.3.1 to show the comparison principle. In the following we verify Assumption 4.3.3.

(I) The function $p \mapsto \mathcal{H}(x,p)$ is convex. Moreover, note that $V_{x,0} = 0$. Hence,

$$\mathcal{H}(x,0) = -\inf_{\theta}\mathcal{I}(x,p,\theta) = 0,$$

being $\mathcal{I} \geq 0$ (see [DV75]) and there exists a measure $\theta^0_{x,p}$ such that $\mathcal{I}(x,p,\theta^0_{x,p}) = 0$, due to Assumption 3.2.3 (see [Kle14] Theorem 17.51).

(II) $\theta \mapsto \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta)$ is bounded for every p and x being a continuous function over a compact set.

(III) $\Upsilon(x) = \frac{1}{2}\log(1+\sum_{i=1}^{l} x_i^2)$ is a containment function since the functions

$$(r(x,y,\gamma))\left(e^{\frac{x}{x^2+1}\gamma x}-1\right),$$

$$e^{-\varphi} r(x,y,\gamma) e^{\frac{x}{1+x^2}\gamma x} e^{\varphi}$$

and

$$e^{-\varphi} r(x,y,\gamma) e^{\varphi}$$

are bounded for every $\gamma \in \Gamma$, every $\varphi \in C^2(F_M)$ and every $(x,y) \in E \times F_M$.

(IV) Let $(x,p) \in E \times \mathbb{R}^l$. The function $V(\cdot,x,p)$ is continuous and hence $(x,p,\theta) \mapsto \Lambda(x,p,\theta)$ is continuous. Moreover, $\mathcal{I}$ is lower semicontinuous, as the supremum over continuous functions. Then, $\mathcal{I}-\Lambda$ is lower semicontinuous and the first property of Definition 4.2.3 follows. We prove now that if $x_n \to x$ and $p_n \to p$ and for all $\theta \in \Theta$, there are θ_n such that $\theta_n \to \theta$ and

$$\limsup_n \mathcal{I}(x_n,p_n,\theta_n) \le \mathcal{I}(x,p,\theta). \tag{4.6.2}$$

Then, the Γ– convergence of $\mathcal{I}-\Lambda$ will follow from (4.6.2) and continuity of Λ. For any $m \in \mathbb{N}$, there exists $\varphi_m \in C^2(F_M)$ such that

$$\mathcal{I}(x,p,\theta) \le \int_{F_M} e^{-\varphi_m} L_{x,p} e^{\varphi_m}\, d\theta + \frac{1}{m}.$$

Then, taking into account the continuity of $L_{x,p}$ and choosing $\theta_n = \theta$ for every n, we get

$$\limsup_n \mathcal{I}(x_n,p_n,\theta_n) \le \int_{F_M} e^{-\varphi_m} L_{x,p} e^{\varphi_m}\, d\theta + \frac{1}{m}.$$

By letting m to infinity we obtain (4.6.2).

(V) As Θ is compact, any closed subset of Θ is compact.

(VI) As explained above, there exists a measure $\theta^0_{x,p}$ such that $\mathcal{I}(x,p,\theta^0_{x,p}) = 0$. Then, $\mathcal{I}(x,p,\theta^0_{x,p}) - \Lambda(x,p,\theta^0_{x,p}) \le -\Lambda(x,p,\theta^0_{x,p})$. Taking $g(x,p) = -\Lambda(x,p,\theta^0_{x,p})$, $\phi_g(x,p)$ is not empty, as $\theta^0_{x,p} \in \phi_g(x,p)$.

(VII) Let $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}, \theta_{\alpha,\varepsilon})$ be a fundamental sequence as in Definition 4.2.1. Set $p_{\alpha,\varepsilon} = \alpha(x_{\alpha,\varepsilon} - y_{\alpha,\varepsilon})$. We aim to show

$$\liminf_{\alpha\to\infty} (\Lambda-\mathcal{I})\,(x_{\alpha,\varepsilon},p_{\alpha,\varepsilon},\theta_{\alpha,\varepsilon}) - (\Lambda-\mathcal{I})\,(y_{\alpha,\varepsilon},p_{\alpha,\varepsilon},\theta_{\alpha,\varepsilon}) \le 0.$$

By the definition of Λ and $\mathcal{I}$ in (3.2.5), the difference above is of the type

$$\int_{F_M} \sum_{\gamma\in\Gamma_1\cup\Gamma_2} (r(x_{\alpha,\varepsilon},z,\gamma) - r(y_{\alpha,\varepsilon},z,\gamma))\left(e^{\langle p_{\alpha,\varepsilon},\gamma x\rangle}-1\right) d\theta + \tag{4.6.3}$$

$$\inf_{\varphi\in C^2(F_M)} \int_{F_M} e^{-\varphi}\left(\sum_{\gamma\in\Gamma_2} r(x_{\alpha,\varepsilon},z,\gamma) - r(y_{\alpha,\varepsilon},z,\gamma)\right) e^{\langle p_{\alpha,\varepsilon},\gamma x\rangle} e^{\varphi} +$$

$$e^{-\varphi}\left(\sum_{\gamma\in\Gamma_3} r(x_{\alpha,\varepsilon},z,\gamma) - r(y_{\alpha,\varepsilon},z,\gamma)\right) e^{\varphi}\, d\theta.$$

Note that if $r(x,y,\gamma)$ is constant in x, the difference above is zero. Hence, we only take into account the parameters γ such that r depends on x.
Moreover, by the upper bound (C3) in Definition 4.2.1, we find that there is some $\alpha(\varepsilon)$ such that

$$\sup_{\alpha\geq\alpha(\varepsilon)} \Lambda(y_{\alpha,\varepsilon},p_{\alpha,\varepsilon},\theta_{\alpha,\varepsilon}) - \mathcal{I}(y_{\alpha,\varepsilon},p_{\alpha,\varepsilon},\theta_{\alpha,\varepsilon}) < \infty. \tag{4.6.4}$$

If $\lim_\alpha r(y_{\alpha,\varepsilon},z,\gamma) > 0$ for all γ, we can conclude by the bound (4.6.4) that $e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}$ is bounded. Then, by property (C2) of Definition 4.2.1 and continuity of the rates, (4.6.3) converges to 0 for $\alpha\to\infty$.
Consider now all terms γ such that $r(y_{\alpha,\varepsilon},z,\gamma)\to 0$ as $\alpha\to\infty$. Firstly note that, by Assumption 3.2.4, (4.6.3) equal to

$$\begin{aligned}
&\int_{F_M}\sum_{\gamma\in\Gamma_1\cup\Gamma_2}(\phi_1^{z,\gamma}(x_{\alpha,\varepsilon})\phi_2^{z,\gamma}(x_{\alpha,\varepsilon})\\
&-\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon}))\left(e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}-1\right)\,d\theta\\
&+\inf_{\varphi\in C^2(F_M)}\int_{F_M} e^{-\varphi}\left(\sum_{\gamma\in\Gamma_2}(\phi_1^{z,\gamma}(x_{\alpha,\varepsilon})\phi_2^{z,\gamma}(x_{\alpha,\varepsilon})\right.\\
&\qquad\left.-\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon}))e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}\right)e^{\varphi}\\
&\qquad\qquad+e^{-\varphi}\left(\sum_{\gamma\in\Gamma_3} r(x_{\alpha,\varepsilon},z,\gamma)-r(y_{\alpha,\varepsilon},z,\gamma)\right)e^{\varphi}\,d\theta.
\end{aligned}\tag{4.6.5}$$

The last line converges to 0 for $\alpha\to\infty$ by the continuity of the rates.
If $\langle p_{\alpha,\varepsilon},\gamma_x\rangle<0$, $e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}$ is bounded and the first two lines also converge to 0 by continuity of the rates.
Consider the case $\langle p_{\alpha,\varepsilon},\gamma_x\rangle>0$.
By Assumption 3.2.4, $\phi_1(y_{\alpha,\varepsilon})>\phi_1(x_{\alpha,\varepsilon})\geq 0$ and $\phi_2(y_{\alpha,\varepsilon})>0$. Then, we can write the first two lines of (4.6.5) as

$$\begin{aligned}
&\int_{F_M}\sum_{\gamma\in\Gamma_1\cup\Gamma_2}\underbrace{\left(\frac{\phi_1^{z,\gamma}(x_{\alpha,\varepsilon})\phi_2^{z,\gamma}(x_{\alpha,\varepsilon})}{\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon})}-1\right)}_{(I)}\cdot\underbrace{(\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon}))\left(e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}-1\right)}_{(II)}\,d\theta\\
&+\inf_{\varphi\in C^2(F_M)}\int_{F_M} e^{-\varphi}\left(\sum_{\gamma\in\Gamma_2}\underbrace{\left(\frac{\phi_1^{z,\gamma}(x_{\alpha,\varepsilon})\phi_2^{z,\gamma}(x_{\alpha,\varepsilon})}{\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon})}-1\right)}_{(III)}\right.\\
&\qquad\qquad\left.\cdot\underbrace{(\phi_1^{z,\gamma}(y_{\alpha,\varepsilon})\phi_2^{z,\gamma}(y_{\alpha,\varepsilon}))\,e^{\langle p_{\alpha,\varepsilon},\gamma_x\rangle}}_{(IV)}\right)e^{\varphi}
\end{aligned}$$

For $\alpha \to \infty$, by Assumption 3.2.4 (2), (1) and (3) are negative and (2) and (4) are positive and bounded by (4.6.4). Then, for α big the second and third lines of (4.6.3) are bounded above from zero and this concludes the proof.

□

4.7 CONCLUSIONS AND FUTURE PERSPECTIVES

In this chapter, we explored the comparison principle for the first-order Hamilton-Jacobi equations, a fundamental tool in proving the uniqueness of viscosity solutions. We discussed its significance in ensuring well-posedness and examined the applicability of our results to different settings. However, models arising from biology, physics, and biochemistry often describe motion on curved surfaces, naturally leading to Hamilton-Jacobi equations on manifolds. Extending the analysis of this chapter to this geometric setting presents several challenges, particularly in adapting the doubling variables technique, which relies on a well-defined smooth distance function. Nevertheless, recent advances, such as those in [KRV19], suggest that this extension is feasible with appropriate modifications.

5 EXISTENCE OF VISCOSITY SOLUTIONS FOR FIRST-ORDER HAMILTON-JACOBI EQUATIONS VIA LYAPUNOV CONTROL

In this chapter, we complete the discussion started in Chapter 4 about the well-posedness in the viscosity sense for first-order Hamilton-Jacobi equations, by considering the existence of viscosity solutions for a stationary and time-dependent equation. Moreover, in this chapter we work on a general d-dimensional smooth manifold.

The results presented in this chapter are based on:

S.Della Corte and R.C. Kraaij, "Existence of viscosity solutions for Hamilton-Jacobi equations via Lyapunov control", Preprint, 2024.

5.1 INTRODUCTION

In this chapter, we present a novel perspective on the existence of viscosity solutions for both stationary and time-dependent first-order Hamilton-Jacobi equations on a d-dimensional smooth manifold $\mathcal{M}$. Let $\mathcal{H} : T^*\mathcal{M} \to \mathbb{R}$. The specific equations we address are:

$$u(x) - \lambda \mathcal{H}(x, \mathrm{d}u(x)) = h(x), \tag{5.1.1}$$

where $\lambda > 0$ and h is a bounded continuous function, and its evolutionary version on $\mathcal{M} \times [0, T]$:

$$\begin{cases} \partial_t u(x,t) + \lambda u(x,t) - \mathcal{H}(x, \mathrm{d}_x u(\cdot,t)(x)) = 0, & \text{if } t > 0, \\ u(x,0) = u_0(x) & \text{if } t = 0 \end{cases} \tag{5.1.2}$$

with $\lambda \geq 0$. Our candidate solutions, denoted as $R_{\lambda,h} : \mathcal{M} \to \mathbb{R}$ and $\mathbf{v}_\lambda : \mathcal{M} \times [0,\infty) \to \mathbb{R}$, are defined through the control problems:

$$R_{\lambda,h}(x) = \sup_{\gamma \in Adm, \gamma(0)=x} J_\lambda(\gamma), \tag{5.1.3}$$

where

$$J_\lambda(\gamma) = \int_0^\infty \lambda^{-1} e^{-\lambda^{-1} t} \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s)) \right) dt$$

and

$$\mathbf{v}_\lambda(x,t) = \sup_{\gamma \in Adm, \gamma(0)=x} W_\lambda(\gamma, t), \tag{5.1.4}$$

where

$$W_\lambda(\gamma, t) = \int_0^t -e^{-\lambda s}\mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds + e^{-\lambda t}u_0(\gamma(t)).$$

Here, $\mathcal{L} : T\mathcal{M} \to \mathbb{R}$ is

$$\mathcal{L}(x, v) = \sup_{p \in T_x^*\mathcal{M}} \langle p, v\rangle - \mathcal{H}(x, p),$$

that it the *convex conjugate* of $\mathcal{H}$ (or Legendre transform when $\mathcal{H}$ is a convex operator) and Adm a set of admissible curves.

The new perspective introduced in this work builds on recent comparison principle results, where an "upper" and "lower" bound of the Hamiltonian are established, as seen in [Tat92],[Tat94],[CL94] and the works by J. Feng and co-authors [Fen06; FK06; FMZ21; AF14]. We follow this tradition, particularly drawing on the more recent works [FK09],[DFL11], [KS21], and [DK24; DCK23], which utilize a specific approach defined in terms of a Lyapunov function. This Lyapunov function plays a crucial role in our analysis as well. In line with these works, we introduce two operators, $H_\dagger$ and $H_\ddagger$, defined using the Lyapunov function.

A Lyapunov function Υ is a function such that

1. its sublevel sets are compact;
2. $\sup_x \mathcal{H}(x, \mathrm{d}\Upsilon(x)) < \infty$.

Taking into account the intuition behind a Lyapunov function, our new operators act on test functions of the type

$$f_\dagger := (1-\varepsilon)f(x) + \varepsilon\Upsilon,$$
$$f_\ddagger := (1+\varepsilon)f(x) - \varepsilon\Upsilon,$$

where $\varepsilon \in (0,1)$, $f \in D(\mathcal{H}) \subseteq C_b(\mathcal{M})$ and Υ a Lyapunov type function. Then, the actions of $H_\dagger$ and $H_\ddagger$ will be respectively

$$H_\dagger f_\dagger(x) := (1-\varepsilon)\mathcal{H}(x, \mathrm{d}f(x)) + \varepsilon C_\Upsilon,$$
$$H_\ddagger f_\ddagger(x) := (1+\varepsilon)\mathcal{H}(x, \mathrm{d}f(x)) - \varepsilon C_\Upsilon,$$

with C_Υ that is morally $\sup_x \mathcal{H}(x, \mathrm{d}\Upsilon(x))$. The above definitions will be motivated in Section 5.5 in the case where $p \mapsto \mathcal{H}(x,p)$ is convex on $T_x^*\mathcal{M}$. We mention that in the preceding discussion on the Lyapunov function, the focus was on C^1 functions for the sake of clarity; however, in further sections of this paper, the analysis will be extended to continuous functions.

We then prove, in Theorems 5.3.3 and 5.3.4, that the upper and lower semi-continuous regularization of (5.1.3) and (5.1.4) are respectively a viscosity subsolution of the equations in terms of $H_\dagger$ and a viscosity supersolution of the equations in terms of $H_\ddagger$.

Our strategy is based on three main steps, each leveraging the use of the Lyapunov control and the regularization of candidate solutions. Below, we outline the steps to demonstrate that the upper semi-continuous regularization of $R_{\lambda,h}$, denoted as $(R_{\lambda,h})^*$, is a subsolution of (5.1.1). The same principles apply to $(\mathbf{v}_\lambda)^*$ and the supersolution proof.

1. **Optimizers construction.** The first step in proving that $(R_{\lambda,h})^*$ is a subsolution involves identifying, for every test function $f_\dagger$, a point x_0 that optimizes

$$\sup\{(R_{\lambda,h})^* - f_\dagger\}. \tag{5.1.5}$$

   Our test function $f_\dagger$ is defined in terms of the Lyapunov function Υ, making it a lower semi-continuous function with compact sublevel sets. This property allows us to identify a sequence of "almost optimizers" x_n for

$$\sup\{(R_{\lambda,h}) - f_\dagger\}$$

   that lie in a compact set. We can then extract a limit point that can serve as optimizer of (5.1.5).
2. **Containment of control paths.** We consider "almost optimizers" γ_n of the control problem (5.1.3) started from x_n found in the step above. By using the two properties of the Lyapunov function Υ, we prove that these sequences stay within compact sets.
3. **Subsolution property in terms of averages.** By using Dynamic Programming Principle and Young's inequality, we prove the subsolution inequality in the sequences of controls γ_n and in terms of averages on small intervals $[0, t_n]$.
   Here, for the supersolution proof, a slightly different analysis is required. Specifically, it is necessary to construct curves that optimize Young's inequality, as outlined in Assumption 5.3.5 (IV).
4. **Stability of the averages.** To show the final subsolution inequality in x_0, we need to prove that the averages considered in step (3) are "stable". Taking the limit $t_n \to 0$ and controlling the asymptotic integrability of these averages with Assumption 5.3.5 (V), we obtain the final inequality.

Here, we place our work within the broader framework of current methodologies for proving the existence of viscosity solutions.

Typically, existence of viscosity solutions proofs are based on two methods. The classical one, called Perron's method, was developed by Ishii [Ish87] and relies on the comparison principle for continuous viscosity solutions and on the existence of a subsolution and a supersolution. This is the case of e.g. [IL90], [CIL92] or more recently [CDL07].

The second method involves the use of the regularity of the Hamiltonian's coefficients. This approach includes several key steps. First, under regularity conditions and analogue of Assumption 5.3.5 (V), the regularity of the solution is established and the set of controls is shown to be complete. Next, optimizers are constructed and the sub-super solution properties are established in these optimizers. Unlike our approach, this method does not require a separate step to prove stability of the averages, as the regularity conditions and the completeness of the controls suffice to ensure the sub-super solution inequalities.

Our method distinguishes itself from the above approaches by relocating the role of Assumption 5.3.5 (V) from the initial step to the final step and by considering the "almost optimizers" and the regularization of the candidate solutions. We then gain the following benefits:

- We avoid a priori analysis of the regularity of the candidate solution;
- We can relax the usual completeness assumption on the set of controls;
- We can relax the traditional assumptions on the Hamiltonian, such as modulus of continuity or uniform coercivity, which are typically necessary in other methods to achieve the two points above.

Moreover, our work extends beyond the current literature by also relaxing the typical assumption of convexity of the Hamiltonian. In this way, we can also consider Hamilton-Jacobi-Isaacs equations with a Hamiltonian expressed as "sup-inf" or "inf-sup" of a convex operator, as detailed in Section 5.5.

Finally, even if our work is mostly inspired by [FK06, Chapter 8], it diverges also from it. Firstly, [FK06] prove that the set of controls is complete. Secondly, their proofs are based on showing properties of (5.1.3) such as the fact that it serves as a classical "left-inverse" of the equation, that it is a *pseudo-resolvent* and it is *contractive* (see Lemma 7.8 and Theorem 8.27 in [FK06]). A similar approach for (5.1.4) is so far lacking. For this reason, the strategy developed in [FK06] can not be used to establish the existence of viscosity solutions for the parabolic case (5.1.2). Our approach, instead, is applicable to both stationary and evolutionary case. Moreover, even if the use of the Lyapunov control is encoded in the strategy of [FK06] and in particular in their Conditions 8.9, 8.10 and 8.11, we make this approach more explicit by introducing the Lyapunov function directly to the domain of our Hamiltonians.

The rest of the chapter is structured as follows: in Section 5.2 we give the main definitions. Our two main results, namely Theorems 5.3.3 and 5.3.4, are stated in Section 5.3 followed by the assumptions needed to prove them. We prove the main results in Section 5.4. Finally, in Section 5.5 and Section 5.6 we explore the context of a convex Hamiltonian and Hamilton-Jacobi-Isaacs equations respectively.

5.2 GENERAL SETTING AND MAIN DEFINITIONS

In this section, we firstly give some notions and definitions used throughout the paper.

Throughout the paper, $\mathcal{M}$ will be the d-dimensional smooth manifold on which we base our Hamilton-Jacobi equations.

The *tangent space* of $\mathcal{M}$ at $x \in \mathcal{M}$ is denoted by $T_x\mathcal{M}$ while $T\mathcal{M} := \bigsqcup_{x\in\mathcal{M}} T_x\mathcal{M}$ is the *tangent bundle* on $\mathcal{M}$. We then denote by $T_x^*\mathcal{M}$ the *cotangent space* of $\mathcal{M}$, that is the dual space of the tangent space, and the correspondent *cotangent bundle* by $T^*\mathcal{M}$. We refer to e.g. [Tu10] for more details about smooth manifolds.

We will also make use of the following notions.

Definition 5.2.1 (Containment function)**.** We call $\Upsilon : \mathcal{M} \to [0, \infty)$ a *containment function* if

(a) $\inf_{x\in\mathcal{M}} \Upsilon(x) = 0$,
(b) for every $c \geq 0$ the set $\{y \,|\, \Upsilon(y) \leq c\}$ is compact.

Definition 5.2.2 (Convergence determining set)**.** Let $\mathcal{A} \subseteq C_b(\mathcal{M})$. We say that $\mathcal{A}$ is *convergence determining* if for all $x_n \in \mathcal{M}$ a sequence in $\mathcal{M}$ and $x_0 \in \mathcal{M}$ the following property holds:

$$\lim_n g(x_n) = g(x_0) \quad \forall g \in \mathcal{A} \qquad \Longrightarrow \qquad \lim_n x_n = x_0.$$

The candidate solutions will be defined through two control problems. Before presenting them, we need to define the set of possible curves on which we set the mentioned control problems.

Definition 5.2.3 (Control set)**.** We say that $Adm \subseteq AC([0,\infty),\mathcal{M})$ is a set of admissible curves if the following two properties hold:
(a) If $\{\gamma(t)\}_{t\geq 0} \in Adm$ and $\tau > 0$, then $\{\gamma(t+\tau)\}_{t\geq 0} \in Adm$;
(b) If $\gamma_1, \gamma_2 \in Adm$ and $\tau > 0$ and let γ be the curve defined as

$$\gamma(t) = \begin{cases} \gamma_1(t) & t \leq \tau \\ \gamma_2(t-\tau) & t > \tau. \end{cases}$$

Then, $\{\gamma(t)\}_{t\geq 0} \in Adm$.

Finally, throughout the whole manuscript we will call $\mathcal{L} : T\mathcal{M} \to [0,\infty]$ the *convex conjugate* of the Hamiltonian, i.e., the function

$$\mathcal{L}(x,v) := \sup_{p\in T^*_x\mathcal{M}} \left[\langle p, v\rangle - \mathcal{H}(x,p)\right]. \tag{5.2.1}$$

Remark 5.2.4. By Definition (5.2.1) of $\mathcal{L}$ it follows that for all $x \in \mathcal{M}$, $v \in T_x\mathcal{M}$ and $p \in T^*\mathcal{M}$, the *Fenchel–Young's inequality* holds, i.e.,

$$\mathcal{L}(x,v) + \mathcal{H}(x,p) \geq \langle p, v\rangle. \tag{5.2.2}$$

5.3 ASSUMPTIONS AND MAIN RESULTS

In this section, we give our main results, namely Theorems 5.3.3 and 5.3.4.

First of all, we need to define the Hamiltonians and their corresponding equations, for which we aim to demonstrate the existence of viscosity solutions. This is the content of the next section. Later, after the statements of our main results, we specify the assumptions needed to prove them and we comment them.

5.3.1 *The upper and lower Hamiltonians*

We will work with two sets of equations in terms of an upper and lower bound of the Hamiltonian $\mathbf{H}f(x) = \mathcal{H}(x, \mathrm{d}f(x))$.

We proceed by introducing $H_\dagger$ and $H_\ddagger$.

Consider Assumption 5.3.5 (III) and the constant C_Υ therein.

Definition 5.3.1 (The operators $H_\dagger$ and $H_\ddagger$)**.** For $f \in C_\ell^\infty(\mathcal{M})$ and $\varepsilon \in (0,1)$ set

$$f_\dagger^\varepsilon := (1-\varepsilon) f + \varepsilon \Upsilon$$
$$g_\dagger^\varepsilon(x) := (1-\varepsilon)\mathcal{H}(x, \mathrm{d}f(x)) + \varepsilon C_\Upsilon.$$

and set

$$H_\dagger := \left\{ (f_\dagger^\varepsilon, g_\dagger^\varepsilon) \,\middle|\, f \in C_\ell^\infty(\mathcal{M}), \varepsilon \in (0,1) \right\}.$$

For $f \in C_u^\infty(\mathcal{M})$ and $\varepsilon \in (0,1)$ set

$$f_\ddagger^\varepsilon := (1+\varepsilon) f - \varepsilon \Upsilon$$
$$g_\ddagger^\varepsilon(x) := (1+\varepsilon)\mathcal{H}(x, \mathrm{d}f(x)) - \varepsilon C_\Upsilon.$$

and set

$$H_\ddagger := \left\{ (f_\ddagger^\varepsilon, g_\ddagger^\varepsilon) \,\middle|\, f \in C_u^\infty(\mathcal{M}), \varepsilon \in (0,1) \right\}.$$

We will establish existence of viscosity solutions for the set of stationary Hamilton–Jacobi equations on a manifold $\mathcal{M}$,

$$\begin{aligned} u(x) - \lambda H_\dagger u(x) &= h_\dagger(x), \\ v(x) - \lambda H_\ddagger v(x) &= h_\ddagger(x); \end{aligned} \tag{5.3.1}$$

where $\lambda > 0$ and $h_\dagger$ and $h_\ddagger$ are two continuous bounded function, and for the set of evolutionary versions

$$\begin{aligned} \partial_t u(x,t) + \lambda u(x,t) - H_\dagger u(x,t) &= 0, \\ \partial_t v(x,t) + \lambda v(x,t) - H_\ddagger v(x,t) &= 0; \end{aligned} \tag{5.3.2}$$

with initial datum u_0 and $\lambda \geq 0$.

In Section 5.5 and Section 5.6, we will show the relationship between the Hamiltonian $\mathbf{H}$ and the operators $H_\dagger$ and $H_\ddagger$ in the scenarios of a convex Hamiltonian and the Hamilton-Jacobi-Isaacs case, respectively. This explanation will then justify the designation of "upper and lower Hamiltonians".

5.3.2 *Main results: existence of viscosity solutions*

We define now the candidate solutions $R_{\lambda,h} : \mathcal{M} \to \mathbb{R}$ and $\mathbf{v}_\lambda : \mathcal{M} \times [0,T] \to \mathbb{R}$ through the control problems

$$R_{\lambda,h}(x) = \sup_{\gamma \in Adm, \gamma(0)=x} J_\lambda(\gamma), \tag{5.3.3}$$

where $J_\lambda(\gamma) = \int_0^\infty \lambda^{-1} e^{-\lambda^{-1} t} \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot\gamma(s))\, ds \right) dt$ and

$$\mathbf{v}_\lambda(x,t) = \sup_{\gamma \in Adm, \gamma(0)=x} W_\lambda(\gamma, t), \tag{5.3.4}$$

where $W_\lambda(\gamma,t) = \int_0^t -e^{-\lambda s} \mathcal{L}(\gamma(s), \dot\gamma(s))\, ds + e^{-\lambda t} u_0(\gamma(t))$.

Remark 5.3.2. We will show later that, by Assumption (IV), there exists a path γ with cost zero. Then, we are allowed to assume that the class of γ considered in the above supremum have finite cost.

We give here the statements of the main results that state respectively the existence of viscosity subsolution and supersolution for the set of equations (5.3.1) and (5.3.2).

We give specifics and the assumptions needed in Section 5.3.3.

Theorem 5.3.3 (Viscosity solution for the stationary equation)**.** *Assume that Assumption 5.3.5 holds. For $\lambda > 0$ and $h \in C_b(\mathcal{M})$ define $R_{\lambda,h}$ as in* (5.3.3) *and let $(R_{\lambda,h})^*$ and $(R_{\lambda,h})_*$ be its upper semi-continuous regularization and the lower semi-continuous regularization respectively. Then, $(R_{\lambda,h})^*$ and $(R_{\lambda,h})_*$ are respectively a viscosity subsolution of $u - \lambda H_\dagger u = h$ and a viscosity supersolution of $u - \lambda H_\ddagger u = h$ with $H_\dagger$ and $H_\ddagger$ defined as in Definition 5.3.1.*

Theorem 5.3.4 (Viscosity solution for the evolutionary equation)**.** *Assume that Assumption 5.3.5 holds. For $T > 0$ and $\lambda \geq 0$ define $\mathbf{v}_\lambda(x,t) : \mathcal{M} \times [0,T] \to \mathbb{R}$ as* (5.3.4) *and let $(\mathbf{v}_\lambda)^*$ and $(\mathbf{v}_\lambda)_*$ be its upper semi-continuous regularization and the lower semi-continuous regularization in both component respectively. Then, $(\mathbf{v}_\lambda)^*$ and $(\mathbf{v}_\lambda)_*$ are respectively a viscosity subsolution of $\partial_t u + \lambda u - H_\dagger u = 0$ and supersolution of $\partial_t u + \lambda u - H_\ddagger u = 0$ with initial value $u(x,0) = u_0(x)$ and $H_\dagger$ and $H_\ddagger$ defined as in Definition 5.3.1.*

5.3.3 *Assumptions*

To prove our main results we will make use of the following assumptions on the Hamiltonian $\mathcal{H}$.

Assumption 5.3.5. Let $\mathcal{H} : T^*\mathcal{M} \to \mathbb{R}$ and call $\mathbf{H}f(x) := \mathcal{H}(x, \mathrm{d}f(x))$ and $D(\mathbf{H}) \subseteq C_b^1(\mathcal{M})$ its domain. The following properties hold.

(I) $\mathcal{H}(x,0) = 0$ for all $x \in \mathcal{M}$;

(II) The map $(x,p) \mapsto \mathcal{H}(x,p)$ is continuous in x and p;

(III) There exists a containment function Υ as in Definition 5.2.1. Moreover, there exists a constant C_Υ such that for all $\gamma \in Adm$ and $T > 0$ the following holds

$$\Upsilon(\gamma(T)) - \Upsilon(\gamma(0)) \leq \int_0^T \mathcal{L}(\gamma(t), \dot\gamma(t))\, dt + TC_\Upsilon.$$

(IV) For all $f \in D(\mathbf{H})$ $x \in \mathcal{M}$ and $T > 0$, there exists $\gamma \in Adm$ such that $\gamma(0) = x$ and

$$f(\gamma(T)) - f(\gamma(0)) \geq \int_0^T \mathcal{L}(\gamma(t), \dot\gamma(t)) + \mathcal{H}(\gamma(t), \mathrm{d}f(\gamma(t)))\, dt.$$

(V) For every compact set K and positive constant c,

$$\overline{\mathcal{H}}(K,c) := \sup_{|p|\leq c} \sup_{x\in K} \mathcal{H}(x,p) < \infty, \qquad \underline{\mathcal{H}}(K,c) := \inf_{|p|\leq c} \inf_{x\in K} \mathcal{H}(x,p) > -\infty.$$

(VI) The space $D(\mathbf{H})$ is convergence determining.

We will now clarify the assumptions above.

The first two assumptions are standard in the context of well–posedness of a Hamilton-Jacobi equation.

Assumptions (V) and (VI) are technical assumptions that imply that the set of curves living in a compact set and having finite cost is relatively compact (that is Condition 8.9.3 of [FK06]). We will use them to prove Proposition 5.4.5 and Proposition 5.4.6.

Assumption (IV) implies that for all $f \in D(\mathbf{H})$ there exists at least one curve $\gamma \in Adm$ such that the *Fenchel–Young's inequality* (5.2.2) applied to $x = \gamma(t)$, $v = \dot{\gamma}(t)$ and $p = \mathrm{d}f(\gamma(t))$ holds with the equality for all $t \in [0, T]$. This assumption is also given in [FK06] as Condition 8.11.

Moreover, using Assumption (IV) and with $f = 1$ and Assumption (I), it follows that for every $x_0 \in \mathcal{M}$ there exists a path γ starting at x_0 such that

$$\int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds = 0.$$

We want to mention that this assumption is only needed to prove the existence of a viscosity supersolution. In the case where $\mathcal{H}(x, \cdot)$ is convex, Assumption (IV) is equivalent to solve the differential inclusion

$$\dot{\gamma}(t) \in \partial_p \mathcal{H}(\gamma(t), \mathrm{d}f(\gamma(t))),$$

(see e.g. [CS04] or [Roc70]). We refer to [FK06, Sec. 8.6.3] for general method to prove the inequality in Assumption (IV).

Finally, Assumption (III) plays a crucial role in many parts of this work. First of all, defining test functions in terms of a containment function Υ, we can work with the definitions of viscosity sub/super-solutions that consider optimizer points and not sequences (see Lemma 1.2.6 in Section 1). Secondly, the containment function allows us to prove that curves starting in a compact set and having finite cost stay in a compact set. This is Condition 8.9.4 of [FK06] and it is the content of Lemma 5.4.4. Moreover, we want to highlight that the containment function Υ is not assumed to be in the domain of the Hamiltonian. This allows us to use also functions that are not in $C^1(\mathcal{M})$. For instance, $\Upsilon(x) = \frac{1}{2}\log(1 + |x|^2)$ is a function that typically works as a containment function in the context of well-posedness of Hamilton-Jacobi equations. We will see in Section 5.5, that in the context where $\mathcal{H}$ is convex and Υ is smooth, assuming that

$$\sup_x \mathcal{H}(x, \nabla\Upsilon(x)) < C_\Upsilon < \infty,$$

the inequality in Assumption (III) is the Fenchel–Young's inequality (5.2.2) for $x = \gamma(t)$, $v = \dot{\gamma}(t)$, $p = \mathrm{d}\Upsilon(x)$ and C_Υ as above.

5.4 PROOFS OF THE MAIN THEOREMS

In this section we give the proofs of Theorems 5.3.3 and 5.3.4. First, we need some results given in the following subsections.

5.4.1 *Dynamic Programming Principle*

We start with an important property of the two value functions, $R_{\lambda,h}$ and $\mathbf{v}_\lambda$ in (5.3.3) and (5.3.4), namely the *Dynamic Programming Principle*. The proof of the following results are standard (see for example [BC97]). We include them for completeness.

Proposition 5.4.1 (Dynamic Programming Principle)**.** *Consider $R_{\lambda,h}$ and $\mathbf{v}_\lambda$ defined as in* (5.3.3) *and* (5.3.4) *respectively. Then, we have the following*

(a) For all $x \in \mathcal{M}$, $\lambda > 0$ and all $T > 0$

$$R_{\lambda,h}(x) = \sup_{\gamma\in Adm,\gamma(0)=x} \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt + e^{-\lambda^{-1}T}R_{\lambda,h}(\gamma(T)). \quad \text{(DPP)}$$

(b) For all $x \in \mathcal{M}$, $\lambda \geq 0$ and $0 < \tau \leq t$,

$$\mathbf{v}_\lambda(x,t) = \sup_{\gamma\in Adm,\gamma(0)=x} \left\{\int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + e^{-\lambda\tau}\mathbf{v}_\lambda(\gamma(\tau),t-\tau)\right\}. \quad \text{(DPPt)}$$

Proof. The proofs of the two properties are both based on integral change of variables that are possible by the definition of Adm which involves piece-wise connectable curves.

Proof of (a). We call $u_T(x)$ the right-hand side of (DPP).

Proof of $R_{\lambda,h}(x) \leq u_T(x)$. If $u_T(x) = +\infty$, there is nothing to prove. Assume that $u_T(x) < +\infty$. We start out with taking any γ in the definition of $R_{\lambda,h}(x)$ and upper bound it in terms of $u_T(x)$. Thus $\gamma \in Adm$ with $\gamma(0) = x$, then,

$$\begin{aligned}
&\int_0^\infty \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt \\
&= \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt \\
&\qquad + \int_T^\infty \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt \\
&= \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt \\
&\qquad + \int_0^\infty \left(h(\gamma(t+T)) - \int_0^{t+T} \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}(t+T)}\,dt \\
&= \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt \\
&\qquad + e^{-\lambda^{-1}T}\int_0^\infty \left(h(\tilde\gamma(t)) - \int_0^t \mathcal{L}(\tilde\gamma(s),\dot{\tilde\gamma}(s))\,ds\right)\lambda^{-1}e^{-\lambda^{-1}t}\,dt,
\end{aligned}$$

where $\tilde{\gamma}(t) = \gamma(t+T)$ that, by Definition 5.2.3 (a) of Adm, is inside the set Adm. Then, taking the supremum over Adm we obtain that $R_{\lambda,h}(x) \leq u_T(x)$.

Proof of $R_{\lambda,h}(x) \geq u_T(x)$. If $R_{\lambda,h}(x) = +\infty$ there is nothing to prove. Thus assume $R_{\lambda,h}(x) < +\infty$ and consider $\varepsilon > 0$, $\gamma \in Adm$ with $\gamma(0) = x$ and $\tilde{\gamma} \in Adm$ such that $\tilde{\gamma}(0) = \gamma(T)$ and

$$R_{\lambda,h}(\gamma(T)) \leq \int_0^\infty \left(h(\tilde{\gamma}(t)) - \int_0^t \mathcal{L}(\tilde{\gamma}(s), \dot{\tilde{\gamma}}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt + \varepsilon.$$

Define now

$$\bar{\gamma}(t) = \begin{cases} \gamma(t) & \text{if } 0 \leq t \leq T; \\ \tilde{\gamma}(t-T) & \text{if } T \leq t. \end{cases}$$

Then, by Definition 5.2.3 (b) of Adm, $\bar{\gamma} \in Adm$ with $\bar{\gamma}(0) = x$, so that

$$\begin{aligned} R_{\lambda,h}(x) &\geq \int_0^\infty \left(h(\bar{\gamma}(t)) - \int_0^t \mathcal{L}(\bar{\gamma}(s), \dot{\bar{\gamma}}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &= \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &\quad + \int_T^\infty \left(h(\tilde{\gamma}(t-T)) - \int_0^t \mathcal{L}(\tilde{\gamma}(s-T), \dot{\tilde{\gamma}}(s-T))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &= \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &\quad + e^{-\lambda^{-1} T} \int_0^\infty \left(h(\tilde{\gamma}(t)) - \int_0^t \mathcal{L}(\tilde{\gamma}(s), \dot{\tilde{\gamma}}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &\geq \int_0^T \left(h(\gamma(t)) - \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds \right) \lambda^{-1} e^{-\lambda^{-1} t}\, dt \\ &\quad + e^{-\lambda^{-1} T} R_{\lambda,h}(\gamma(T)) - \varepsilon. \end{aligned}$$

Due to the arbitrariness of ε, we obtain that $R_{\lambda,h}(x) \geq u_T(x)$ and this conclude the proof.

Proof of (b). We call $v_\tau(x,t)$ the right-hand side of (DPPt).

Proof of $\mathbf{v}_\lambda(x,t) \leq v_\tau(x,t)$. For $t = \tau$, (DPPt) is the definition of $\mathbf{v}_\lambda$ (5.3.4). Suppose $t > \tau$. If $v_\tau(x,t) = +\infty$, there is nothing to prove. Assume that $v_\tau(x,t) < +\infty$. We start out with taking any γ in the definition of $\mathbf{v}_\lambda(x,t)$ and upper bound it in terms of $v_\tau(x,t)$. Thus $\gamma \in Adm$ with $\gamma(0) = x$, then,

$$\begin{aligned}
&\int_0^t -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + e^{-\lambda t}u_0(\gamma(s))\\
&= \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + \int_\tau^t -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + e^{-\lambda t}u_0(\gamma(t))\\
&= \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + \int_0^{t-\tau} -e^{-\lambda s-\lambda\tau}\mathcal{L}(\gamma(s+\tau),\dot\gamma(s+\tau))\,ds\\
&\quad + e^{-\lambda\tau}e^{-\lambda(t-\tau)}u_0(\gamma(t))\\
&= \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds\\
&\quad + e^{-\lambda\tau}\left(\int_0^{t-\tau} -e^{-\lambda s}\mathcal{L}(\tilde\gamma(s),\dot{\tilde\gamma}(s))\,ds + e^{-\lambda(t-\tau)}u_0(\tilde\gamma(t-\tau))\right),
\end{aligned}$$

with $\tilde\gamma(t) = \gamma(t+\tau)$ that, by Definition 5.2.3 (a) of Adm, is inside the set Adm. Then, taking the supremum over Adm we obtain the inequality $\mathbf{v}_\lambda(x,t) \le v_\tau(x,t)$.

Proof of $\mathbf{v}_\lambda(x,t) \ge v_\tau(x,t)$. If $\mathbf{v}_\lambda(x,t) = +\infty$ there is nothing to prove. Thus assume $\mathbf{v}_\lambda(x,t) < \infty$ and consider $\varepsilon > 0$, $\gamma \in Adm$ such that $\gamma(0) = x$ and $\tilde\gamma \in Adm$ such that $\tilde\gamma(0) = \gamma(\tau)$ and

$$\mathbf{v}_\lambda(\gamma(\tau), t-\tau) \le \int_0^{t-\tau} -e^{-\lambda s}\mathcal{L}(\tilde\gamma(s),\dot{\tilde\gamma}(s)\,ds + e^{-\lambda(t-\tau)}u_0(\tilde\gamma(t-\tau)) + \varepsilon.$$

Define

$$\bar\gamma(s) = \begin{cases}\gamma(s) & \text{if } s \le \tau;\\ \tilde\gamma(s-\tau) & \text{if } t > \tau.\end{cases}$$

Then, by Definition 5.2.3 (b) of Adm, $\bar\gamma \in Adm$ and $\bar\gamma(0) = x$, so that

$$\begin{aligned}
\mathbf{v}_\lambda(x,t) &\ge \int_0^t -e^{-\lambda s}\mathcal{L}(\bar\gamma(s),\dot{\bar\gamma}(s))\,ds + e^{-\lambda t}u_0(\bar\gamma(t))\\
&= \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + \int_\tau^t -e^{-\lambda s}\mathcal{L}(\tilde\gamma(s-\tau),\dot{\tilde\gamma}(s-\tau))\,ds\\
&\qquad + e^{-\lambda t}u_0(\tilde\gamma(t-\tau))\\
&= \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s)\,ds + e^{-\lambda\tau}\int_0^{t-\tau} -e^{-\lambda s}\mathcal{L}(\tilde\gamma(s),\dot{\tilde\gamma}(s))\,ds\\
&\qquad + e^{-\lambda\tau}e^{-\lambda(t-\tau)}u_0(\tilde\gamma(t-\tau))\\
&\ge \int_0^\tau -e^{-\lambda s}\mathcal{L}(\gamma(s),\dot\gamma(s))\,ds + e^{-\lambda\tau}\mathbf{v}_\lambda(\gamma(\tau), t-\tau) - \varepsilon.
\end{aligned}$$

Due to the arbitrariness of ε we obtain that $\mathbf{v}_\lambda(x,t) \ge v_\tau(x,t)$ and this conclude the proof. □

5.4.2 *Properties of semi-continuous functions*

The following two propositions will be used for $R_{\lambda,h}$ and $\mathbf{v}_\lambda$ respectively. We only prove the first one as the second one follows similarly.

Proposition 5.4.2. *Let $\phi : \mathcal{M} \to \mathbb{R}$ be a bounded function on $\mathcal{M}$ and $f : \mathcal{M} \to \mathbb{R}$ a lower semi-continuous function with compact sublevel sets. Define ϕ^* the upper semi-continuous regularization of ϕ. Then, there exists a converging sequence $x_n \to x_0$ such that the following properties hold.*

(a) $\phi(x_n) - f(x_n) \geq \sup(\phi - f) - \frac{1}{n}$,
(b) $\phi(x_0) - f(x_0) = \sup(\phi - f) = \sup(\phi^* - f)$.
(c) $\lim_n \phi(x_n) = \phi^*(x_0)$.

Proof. (a) For every $n \geq 1$, there exists x_n such that

$$\phi(x_n) - f(x_n) \geq \sup(\phi - f) - \frac{1}{n}. \tag{5.4.2}$$

We prove that the sequence $\{x_n\}_{n\geq 1}$ is contained in a sublevel set of f, and, therefore, in a compact set. By (5.4.2), for all $x \in \mathcal{M}$,

$$\phi(x_n) - f(x_n) \geq \phi(x) - f(x) - \frac{1}{n}.$$

Let $\tilde{x}$ be a point in a sublevel set of f of constant C. We get

$$f(x_n) \leq \phi(x_n) - \phi(\tilde{x}) + f(\tilde{x}) + \frac{1}{n} \leq 2\|\phi\| + C + \frac{1}{n}.$$

Then, for n large the right-hand side of the above inequality is bounded from above by a constant M. We can conclude that, for n large, $x_n \in \{x \in \mathcal{M} : f(x) \leq M\}$. By going to a converging subsequence, we conclude the proof of the firt point.

(b) Let x_0 be the limit of the sequence x_n. By the upper semi continuity of $\phi^* - f$, we have

$$\phi^*(x_0) - f(x_0) \geq \limsup_n (\phi^*(x_n) - f(x_n)) = \sup(\phi^* - f).$$

(c) First note that by (a) and (b), we have that

$$\lim_n \phi(x_n) - f(x_n) = \phi^*(x_0) - f(x_0). \tag{5.4.3}$$

Moreover, by the definition of ϕ^* as the upper semicontinuous regularization of ϕ, it holds that $\limsup_n \phi(x_n) \leq \phi^*(x_0)$. We show that $\liminf_n \phi(x_n) \geq \phi^*(x_0)$. Suppose by contradiction that $\liminf_n \phi(x_n) = \underline{\phi} < \phi^*(x_0)$. Then, consider a subsequence x_{n_m} such that $\limsup_n \phi(x_{n_m}) = \underline{\phi}$. Then,

$$\begin{aligned}\limsup_n \phi(x_{n_m}) - f(x_{n_m}) &\leq \limsup_n \phi(x_{n_m}) - \liminf_n f(x_{n_m}) \\ &\leq \underline{\phi} - f(x_0) < \phi^*(x_0) - f(x_0),\end{aligned}$$

that is a contradiction to (5.4.3). This concludes the proof of (c).

□

Proposition 5.4.3. *Let $\phi : \mathcal{M} \times [0,T] \to \mathbb{R}$ be a bounded function, $f : \mathcal{M} \to \mathbb{R}$ a lower semi-continuous function with compact sublevel sets and $h : [0,T] \to \mathbb{R}$ a C^1 function. Define ϕ^* the upper semi-continuous regularization of ϕ in both variables. Then, the following properties hold.*

(a) $\exists\,(x_n, t_n)$ *almost optimizing $\phi - f - h$ with an error of $\frac{1}{n}$, that is*

$$\phi(x_n, t_n) - f(x_n) - h(t_n) \geq \sup(\phi - f - h) - \frac{1}{n},$$

and such that (x_n, t_n) has a converging subsequence, still denoted (x_n, t_n).

(b) *The limit point (x_0, t_0) of the sequence (x_n, t_n) is optimal for $\phi - f - h$ and $\phi^* - f - h$.*

(c) $\lim_n \phi(x_n, t_n) = \phi^*(x_0, t_0)$.

5.4.3 *Properties of the controls set*

In this subsection we prove some properties of the controls $\gamma \in Adm$. In particular, we will prove that curves starting in a compact and having finite cost stay in a compact set. Additionally, sequences composed of these type of curves will uniformly converge.

We want to emphasize that the assumption of the existence of the containment function, i.e. Assumption 5.3.5 (III), plays a crucial role here. Indeed, the property given by the lemma below is usually assumed in an optimal control context (see for example Condition 8.9.4 in [FK06]). In the following, we are able to prove it by using the compact sublevel sets of the containment function.

Lemma 5.4.4. *Let $T > 0$ and K_0 a compact in $\mathcal{M}$. Let $\gamma \in Adm$ such that $\gamma(0) \in K_0$. If there exists a constant $M = M(T, K_0)$ such that*

$$\int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t))\, dt < M,$$

then, there exists a compact K such that $\gamma(t) \in K$ for all $t \leq T$.

Proof. Firstly, recall the containment function Υ and the constant C_Υ given in Assumption (III). Then,

$$\begin{aligned} \Upsilon(\gamma(T)) &\leq \Upsilon(\gamma(0)) + C_\Upsilon T + \int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t))\, dt \\ &\leq \sup_{K_0} \Upsilon + C_\Upsilon T + M := \bar{M}. \end{aligned}$$

Then, the result follows with $K = \{x \in E \,:\, \Upsilon(x) \leq \bar{M}\}$ and by the property of Υ of having compact sublevel sets. □

We show now that sequences of curves lying in a compact set and having finite cost are uniformly convergent. More precisely, we have the following proposition.

Proposition 5.4.5. *Let $T > 0$ and $K \subseteq \mathcal{M}$ a compact set. Let $\gamma_n \in Adm$ a sequence of admissible curves such that $\gamma_n(t) \in K$, for every n and $t \leq T$. Let $T_n \in [0, T]$ such that $T_n \downarrow 0$. Let $x_n := \gamma_n(0)$ converge to $x_0 \in E$. If,*

$$\sup_n \sup_{t \leq T_n} \frac{1}{t} \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds < \infty, \tag{5.4.4}$$

then,

$$\lim_n \gamma_n(t_n) = x_0,$$

for all t_n vanishing sequence faster then T_n.

Proof. We will show the convergence by proving that for every function $g \in D(\mathbf{H})$ it holds that

$$\lim_n g(\gamma_n(t_n)) = g(x_0).$$

Then, the result will follow by Assumption 5.3.5 (VI).

Let $g \in D(\mathbf{H})$. First of all, we show that for all $n \geq 1$

$$|g(\gamma_n(t_n)) - g(\gamma_n(0))| \leq t_n \cdot M,$$

with $M > 0$. To this aim, note that by the Frenchel–Young's inequality (5.2.2) applied to $x = \gamma(t), v = \dot{\gamma}(t)$ and $p = \mathrm{d}g(\gamma(t))$,

$$\begin{aligned} |g(\gamma_n(t_n)) - g(\gamma_n(0))| &\leq \int_0^{t_n} \langle \mathrm{d}g(\gamma_n(s)), \dot{\gamma}_n(s) \rangle\, ds \\ &\leq \int_0^{t_n} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds + \int_0^{t_n} \mathcal{H}(\gamma_n(s), \mathrm{d}g(\gamma_n(s)))\, ds \\ &\leq M \cdot t_n, \end{aligned}$$

where the last bound follows from the assumption on $\mathcal{L}$, the continuity of $\mathcal{H}$ and the fact that $\gamma_n(s)$ lies on a compact set for all $s \in [0, T]$. Then, by triangular inequality it follows that

$$\begin{aligned} |g(\gamma_n(t_n)) - g(x_0)| &\leq |g(\gamma_n(t_n)) - g(\gamma_n(0))| + |g(\gamma_n(0)) - g(x_0)| \\ &\leq t_n \cdot M + |g(\gamma_n(0)) - g(x_0)|. \end{aligned}$$

Sending $n \to \infty$ and using the continuity of g, the claim follows. That concludes the proof. □

Finally, in order to apply the proposition above, we will need to show condition (5.4.4). The following proposition, give us a property that implies condition (5.4.4).

Proposition 5.4.6. *Let $T > 0$, K a compact set and $C_1, C_2 \geq 0$. Let $\gamma_n \in Adm$ a sequence of admissible curves such that $\gamma_n(t) \in K$, for every n and $t \leq T$. Let $T_n \in [0,T]$ such that $T_n \downarrow 0$. Moreover, let $f \in D(\mathbf{H})$ such that the following holds for every n and $t \leq T_n$*

$$\int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \leq C_1 \int_0^t \langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle + C_2 t. \tag{5.4.5}$$

Then,

$$\sup_n \sup_{t \leq T_n} \frac{1}{t} \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds < \infty.$$

Proof. Let $t \leq T_n$. Let $\psi_{f,K}$ be a function as in Lemma 5.4.7. Then,

$$\begin{aligned}\int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds &\leq C_1 \int_0^t \langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle\, ds + C_2 t \\ &\leq C_1 \int_0^t \psi_{f,K}(\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s)))\, ds + C_2 t.\end{aligned} \tag{5.4.6}$$

Moreover, by the fact that $\psi_{f,K}$ is non decreasing and the fact that $\frac{\psi_{f,K}(r)}{r}$ converges to 0 for $r \to \infty$, there exist $0 < m < 1$ and $r^* \geq 1$ such that $\frac{\psi_{f,K}(r)}{r} \leq m$ for every $r \geq r^*$. We get

$$\begin{aligned}&\int_0^t \psi_{f,K}(\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s)))\, ds \\ &\leq \int_{\{s \in [0,t] \,:\, \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s)) \geq r^*\}} \frac{\psi_{f,K}\left(\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\right)}{\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \\ &\quad + \int_{\{s \in [0,t] \,:\, \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s)) \leq r^*\}} \psi_{f,K}(\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))) \\ &\leq \int_0^t m\, \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds + \int_0^t \psi_{f,K}(r^*)\, ds \\ &\leq \int_0^t m\, \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds + t\psi_{f,K}(r^*).\end{aligned} \tag{5.4.7}$$

Combining (5.4.6) and (5.4.7) leads to

$$\sup_n \sup_{t \leq T_n} \left(\frac{1}{t} \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \right) \leq M,$$

for some $M > 0$, establishing the claim. □

The following technical lemma is inspired by Lemma 10.21 in [FK06].

Lemma 5.4.7. *For every $f \in D(\mathbf{H})$ and compact set $K \subseteq \mathcal{M}$ there exists a right continuous, non decreasing function $\psi_{f,K} : [0, \infty) \to [0, \infty)$ such that*

(a) $\lim_{r\to\infty} \frac{\psi_{f,K}(r)}{r} = 0$;
(b) $|\langle \mathrm{d}f(x), q\rangle| \leq \psi_{f,K}(\mathcal{L}(x,q))$ *for all* $x \in K, q \in T_x\mathcal{M}$.

Proof. First recall that by Assumption (V),

$$\overline{H}(K,c) := \sup_{|p|\leq c} \sup_{x\in K} H(x,p) < \infty \qquad \text{for all } c > 0.$$

Using the definition of $\mathcal{L}$ we obtain that

$$\frac{\mathcal{L}(x,v)}{|v|} \geq \sup_{|p|=c} \frac{\langle p, v\rangle}{|v|} - \frac{\overline{H}(K,c)}{|v|} = c - \frac{\overline{H}(K,c)}{|v|}.$$

It follows that

$$\lim_{N\to\infty} \inf_{x\in K} \inf_{|v|=N} \frac{\mathcal{L}(x,v)}{|v|} = +\infty.$$

Define

$$\varphi(s) = s \inf_{x\in K} \inf_{|v|\geq s} \frac{\mathcal{L}(x,v)}{|v|}.$$

Then φ is strictly increasing and $r^{-1}\varphi(r) \to \infty$ for $r \to \infty$. Moreover, for every $f \in D(\mathbf{H})$ and a compact set K there exists a constant $C_{f,K} > 0$ such that

$$|\langle \mathrm{d}f(x), q\rangle| \leq C_{f,K}|q| \qquad \text{for all } x \in K.$$

We define

$$\psi_{f,K}(r) := C_{f,K}\varphi^{-1}(r).$$

Then, $\psi_{f,K}$ is such that $r^{-1}\psi_{f,K}(r) \to 0$ for $r \to \infty$ and since

$$\varphi\left(C_{f,K}^{-1}|\langle \mathrm{d}f(x), q\rangle|\right) \leq \varphi(|q|) \leq \mathcal{L}(x,q),$$

we can conclude that $|\langle \mathrm{d}f(x), q\rangle| \leq \psi_{f,K}(\mathcal{L}(x,q))$. □

5.4.4 *Proofs of Theorems 5.3.3 and 5.3.4*

Proof of Theorem 5.3.3. The subsolution property. Let $f^\varepsilon_\dagger$ as in Definition 5.3.1. By Proposition 5.4.2, with $\phi = R_{\lambda,h}$ and $f = f^\varepsilon_\dagger$, there exists a sequence x_n in a compact set K_0 that is converging to a point x_0 and such that

$$R_{\lambda,h}(x_n) - f^\varepsilon_\dagger(x_n) \geq \sup(R_{\lambda,h} - f^\varepsilon_\dagger) - \frac{1}{n^2}, \tag{5.4.8}$$

$$(R_{\lambda,h})^*(x_0) - f^\varepsilon_\dagger(x_0) = \sup((R_{\lambda,h})^* - f^\varepsilon_\dagger).$$

and

$$\lim_n R_{\lambda,h}(x_n) = (R_{\lambda,h})^*(x_0). \tag{5.4.9}$$

It thus suffices to establish that

$$(R_{\lambda,h})^*(x_0) - \lambda g^\varepsilon_\dagger(x_0) - h(x_0) \leq 0. \tag{5.4.10}$$

Let $\gamma_n \in Adm$ such that $\gamma_n(0) = x_n$ and almost optimizing (DPP) at pag. 117, that is

$$\begin{aligned} & R_{\lambda,h}(x_n) - e^{-\lambda^{-1}\frac{1}{n}} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \\ & \leq \int_0^{1/n} \left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds\right] \lambda^{-1} e^{-\lambda^{-1}t}\, dt + \frac{1}{n^2}. \end{aligned} \tag{5.4.11}$$

Moreover, as pointed out in Remark 5.3.2, we can assume that $\int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds < \infty$, for all $t \leq \frac{1}{n}$.

Then, by the fact that $\gamma_n(0) \in K_0$ and Lemma 5.4.4 applied with $T = \frac{1}{n}$, there exists a compact K such that $\gamma_n(t) \in K$ for all $t \leq \frac{1}{n}$.

Rearranging (5.4.8), using the fundamental theorem of calculus, and Assumption (III), we find

$$\begin{aligned} & R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) - R_{\lambda,h}(x_n) \leq f^\varepsilon_\dagger\left(\gamma_n\left(\frac{1}{n}\right)\right) - f^\varepsilon_\dagger(x_n) + \frac{1}{n^2} \\ & = (1-\varepsilon)\left(f\left(\gamma_n\left(\frac{1}{n}\right)\right) - f(\gamma_n(x_n))\right) + \varepsilon\left(\Upsilon\left(\gamma_n\left(\frac{1}{n}\right)\right) - \Upsilon(\gamma_n(x_n))\right) + \frac{1}{n^2} \\ & \leq (1-\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot{\gamma}_n(s)\rangle\, ds + \varepsilon\int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds + \varepsilon\frac{1}{n}C_\Upsilon + \frac{1}{n^2}. \end{aligned} \tag{5.4.12}$$

Then, combining (5.4.11) and (5.4.12) leads to

$$\begin{aligned} & -(1-\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot{\gamma}_n(s)\rangle\, ds - \varepsilon\int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \\ & \qquad - \varepsilon\frac{1}{n}C_\Upsilon - \frac{1}{n^2} + \left(1 - e^{-\lambda^{-1}\frac{1}{n}}\right) R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \\ & \leq R_{\lambda,h}(x_n) - e^{-\lambda^{-1}\frac{1}{n}} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \\ & \leq \int_0^{1/n} \left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds\right] \lambda^{-1} e^{-\lambda^{-1}t}\, dt + \frac{1}{n^2}. \end{aligned}$$

Finally, rearranging terms on the first and third line and dividing by $\frac{1}{n}$ yields,

$$\begin{aligned}0 \leq n\left(e^{-\lambda^{-1}1/n}-1\right)R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right)+n\int_0^{1/n}\lambda^{-1}e^{-\lambda^{-1}t}h\left(\gamma_n(t)\right)\,dt\\ -n\int_0^{1/n}\left(\lambda^{-1}e^{-\lambda^{-1}t}\int_0^t\mathcal{L}(\gamma_n(s),\dot{\gamma}_n(s))\,ds\right)dt\\ +n(1-\varepsilon)\int_0^{1/n}\langle \mathrm{d}f(\gamma_n(s)),\dot{\gamma}_n(s)\rangle\,ds\\ +n\varepsilon\int_0^{1/n}\mathcal{L}(\gamma_n(s),\dot{\gamma}_n(s))\,ds+\varepsilon C_\Upsilon+O\left(\frac{1}{n}\right).\end{aligned}$$

Using integration by parts in the integral involving $\mathcal{L}$, we lead to

$$0 \leq n\left(e^{-\lambda^{-1}1/n}-1\right)R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \tag{5.4.13a}$$

$$+n\int_0^{1/n}\lambda^{-1}e^{-\lambda^{-1}1/n}h\left(\gamma_n(t)\right)\,dt \tag{5.4.13b}$$

$$+n(1-\varepsilon)\int_0^{1/n}\langle \mathrm{d}f(\gamma_n(t)),\dot{\gamma}_n(t)\rangle-\mathcal{L}(\gamma_n(t),\dot{\gamma}_n(t))\,dt+\varepsilon C_\Upsilon \tag{5.4.13c}$$

$$-n\int_0^{1/n}\left(e^{-\lambda^{-1}t}-1\right)\mathcal{L}(\gamma_n(t),\dot{\gamma}_n(t))\,dt+O\left(\frac{1}{n}\right). \tag{5.4.13d}$$

We show now that taking the limit in (5.4.13) as $n\to\infty$ leads to inequality (5.4.10). We consider the limit in (5.4.13a), (5.4.13b), (5.4.13c), and (5.4.13d) separately.

Key to the analysis of the above limiting behavior is to first establish

$$\sup_n \sup_{t\leq\frac{1}{n}}\frac{1}{t}\int_0^t\mathcal{L}(\gamma_n(s),\dot{\gamma}_n(s))\,ds<\infty, \tag{5.4.14}$$

and

$$\gamma_n\left(\frac{1}{n}\right)\to x_0,$$

where the second is a direct consequence of the first due to Proposition 5.4.5. To do so, we use Proposition 5.4.6 with $T_n=\frac{1}{n}$ for which we prove condition (5.4.5).

Putting all terms in (5.4.13) involving $\mathcal{L}$ on the left-hand side, we obtain

$$\begin{aligned}n\int_0^{1/n}\left(e^{-\lambda^{-1}t}-\varepsilon\right)\mathcal{L}(\gamma_n(t),\dot{\gamma}_n(t))\,dt \leq n(1-\varepsilon)\int_0^{1/n}\langle \mathrm{d}f(\gamma_n(t)),\dot{\gamma}_n(t)\rangle\,dt+\varepsilon C_\Upsilon\\ +n\left(e^{-\lambda^{-1}/n}-1\right)R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right)+n\int_0^{1/n}\lambda^{-1}e^{-\lambda^{-1}t}h\left(\gamma_n(t)\right)\,dt+O\left(\frac{1}{n}\right)\end{aligned}$$

Using that $\|R_{\lambda,h}\| \leq \|h\|$ leads to

$$n \int_0^{1/n} (e^{-\lambda^{-1}s} - \varepsilon)\mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \leq n \int_0^{1/n} (1-\varepsilon)\langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle\, ds + \varepsilon C_{\Upsilon} + 2\|h\| + O\left(\frac{1}{n}\right),$$

that is,

$$n \int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds \leq n \int_0^{1/n} \frac{(1-\varepsilon)}{(e^{-\lambda^{-1}\delta} - \varepsilon)}\langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle\, ds + \frac{\varepsilon}{(e^{-\lambda^{-1}\delta} - \varepsilon)} C_{\Upsilon} + \frac{2}{(e^{-\lambda^{-1}\delta} - \varepsilon)}\|h\| + O\left(\frac{1}{n}\right),$$

with $\delta \in (0, 1/n)$ small enough. We can conclude that γ_n verifies condition (5.4.5) of Proposition 5.4.6 implying the bound (5.4.14). Before exploring the limits in (5.4.13), note that by (5.4.14) and Proposition 5.4.5, we obtain that $\gamma_n\left(\frac{1}{n}\right) \to x_0$ for $n \to \infty$.

Limsup of (5.4.13a)**:** As the pre-factor in (5.4.13a) tends to $-\lambda$, we cannot directly use that $\gamma_n(1/n) \to x_n$ in combination with the upper semi-continuous regularization of $R_{\lambda,h}$. Instead, we will argue using (5.4.9). We set this up via (5.4.11). Note first that

$$\limsup_n n\left(e^{-\lambda^{-1}/n} - 1\right) R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) = -\lambda^{-1} \liminf_n e^{-\lambda^{-1}/n} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right). \tag{5.4.15}$$

By (5.4.11) and (5.4.9) we have

$$\begin{aligned}
&\liminf_n e^{-\lambda^{-1}/n} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \\
&\geq \liminf_n R_{\lambda,h}(x_n) - \int_0^{1/n}\left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds\right] \lambda^{-1} e^{-\lambda^{-1}t}\, dt - \frac{1}{n^2} \\
&= (R_{\lambda,h})^*(x_0) - \limsup_n \int_0^{1/n}\left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds\right] \lambda^{-1} e^{-\lambda^{-1}t}\, dt.
\end{aligned}$$

As the integral terms vanish because $\gamma_n(t) \in K$ and $\mathcal{L} \geq 0$, we conclude that

$$\limsup_n n\left(e^{-\lambda^{-1}/n} - 1\right) R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) = -\lambda^{-1}(R_{\lambda,h})^*(x_0).$$

Limsup of (5.4.13b)**:** By the convergence of $\gamma_n(1/n)$ to x_0, the continuity of h and the dominated convergence theorem, we get that

$$\limsup_n n \int_0^{1/n} \lambda^{-1} e^{-\lambda^{-1}t} h\left(\gamma_n(t)\right)\, dt = \lambda^{-1} h(x_0). \tag{5.4.16}$$

Limsup of (5.4.13c)**:** By using Fenchel–Young's inequality (5.2.2) (page 113) for $\langle \mathrm{d}f(\gamma_n(t)), \dot{\gamma}_n(t)\rangle$ we get,

$$\begin{aligned} &= \limsup_n n \int_0^{1/n} (1-\varepsilon)\left(\langle \mathrm{d}f(\gamma_n(t)), \dot{\gamma}_n(t)\rangle - \mathcal{L}(\gamma_n(t), \dot{\gamma}_n(t))\right)\, dt + \varepsilon C_{\mathrm{Y}} \qquad (5.4.17) \\ &\leq \limsup_n n \int_0^{1/n} (1-\varepsilon)\mathcal{H}(\gamma_n(t), \mathrm{d}f(\gamma_n(t)))\, dt + \varepsilon C_{\mathrm{Y}} \\ &= g_\dagger^\varepsilon(x_0), \end{aligned}$$

where in the last equality we used the dominated convergence theorem and the convergence of $\gamma_n(1/n) \to x_0$.

Limit of (5.4.13d)**:** By (5.4.14),

$$\lim_n n \int_0^{1/n} (e^{-\lambda^{-1}t} - 1)\mathcal{L}(\gamma_n(t), \dot{\gamma}_n(t))\, dt + O\left(\frac{1}{n}\right) = 0. \qquad (5.4.18)$$

Then, combining all the limits (5.4.15), (5.4.16), (5.4.17) and (5.4.18) in (5.4.13), we can conclude that

$$0 \leq -\lambda^{-1}(R_{\lambda,h})^*(x_0) + g_\dagger^\varepsilon(x_0) + \lambda^{-1}h(x_0).$$

that concludes the first part of the proof.

The supersolution property. Let $f_\ddagger^\varepsilon$ be as in Definition 5.3.1. By Proposition 5.4.2, with $\phi = -R_{\lambda,h}$ and $f = -f_\ddagger^\varepsilon$, there exists a sequence x_n in a compact set K_0 converging to a point x^0 and such that

$$\begin{aligned} &R_{\lambda,h}(x_n) - f_\ddagger^\varepsilon(x_n) \leq \inf_x (R_{\lambda,h} - f_\ddagger^\varepsilon) + \frac{1}{n^2}, \qquad (5.4.19) \\ &(R_{\lambda,h})_*(x^0) - f_\ddagger^\varepsilon(x^0) = \inf_x ((R_{\lambda,h})_* - f_\ddagger^\varepsilon), \end{aligned}$$

and

$$\lim_n R_{\lambda,h}(x_n) = (R_{\lambda,h})_*(x_0). \qquad (5.4.20)$$

It thus suffices to establish that

$$(R_{\lambda,h})_*(x^0) - \lambda g_\ddagger^\varepsilon(x^0) - h(x^0) \geq 0. \qquad (5.4.21)$$

By Assumption (IV), there exists $\gamma_n \in Adm$ with $\gamma_n(0) = x_n$ and such that

$$\int_0^{1/n} \langle \mathrm{d}f(\gamma_n(t)), \dot{\gamma}_n(t)\rangle\, dt = \int_0^{1/n} \mathcal{H}(\gamma_n(t), \mathrm{d}f(\gamma_n(t))) + \mathcal{L}(\gamma_n(t), \dot{\gamma}_n(t))\, dt. \qquad (5.4.22)$$

Moreover, by the fact that $\gamma_n(0) \in K_0$ and Lemma 5.4.4 applied with $T = \frac{1}{n}$, there exists a compact K such that $\gamma_n(t) \in K$ for all $t \leq \frac{1}{n}$.

By (5.4.19) and Assumption (III),

$$R_{\lambda,h}(x_n) - R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \leq f_{\ddagger}^{\varepsilon}(x_n) - f_{\ddagger}^{\varepsilon}\left(\gamma_n\left(\frac{1}{n}\right)\right) + \frac{1}{n^2}$$
$$= (1+\varepsilon)\left(f(\gamma_n(x_n)) - f\left(\gamma_n\left(\frac{1}{n}\right)\right)\right) - \varepsilon\left(\Upsilon(\gamma_n(x_n)) - \Upsilon\left(\gamma_n\left(\frac{1}{n}\right)\right)\right) + \frac{1}{n^2}$$
$$\leq -(1+\varepsilon)\int_0^{1/n}\langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle\, ds + \varepsilon\int_0^{1/n}\mathcal{L}(\gamma_n(s),\dot\gamma_n(s))\, ds + \varepsilon\frac{1}{n}C_\Upsilon + \frac{1}{n^2}.$$

Moreover, by (DPP) at pag. 117

$$R_{\lambda,h}(x_n) - e^{-\lambda^{-1}\frac{1}{n}}R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \tag{5.4.23}$$
$$\geq \int_0^{1/n}\left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s),\dot\gamma_n(s))\, ds\right]\lambda^{-1}e^{-\lambda^{-1}t}\, dt.$$

Then,

$$\int_0^{1/n}\left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s),\dot\gamma_n(s))\, ds\right]\lambda^{-1}e^{-\lambda^{-1}t}\, dt$$
$$\leq R_{\lambda,h}(x_n) - R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) + (1-e^{-\lambda^{-1}\frac{1}{n}})R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right)$$
$$\leq -(1+\varepsilon)\int_0^{1/n}\langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle\, ds + \varepsilon\int_0^{1/n}\mathcal{L}(\gamma_n(s),\dot\gamma_n(s))\, ds$$
$$+\varepsilon\frac{1}{n}C_\Upsilon + (1-e^{-\lambda^{-1}\frac{1}{n}})R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) + \frac{1}{n^2}.$$

Dividing by $\frac{1}{n}$ and using integration by parts in the integral involving $\mathcal{L}$ yields,

$$0 \leq -n\left(e^{-\lambda^{-1}\frac{1}{n}} - 1\right)R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \tag{5.4.24a}$$
$$-n\int_0^{1/n}\lambda^{-1}e^{-\lambda^{-1}\frac{1}{n}}h\left(\gamma_n(t)\right)\, dt \tag{5.4.24b}$$
$$-n(1+\varepsilon)\int_0^{1/n}\langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle - \mathcal{L}(\gamma_n(t),\dot\gamma_n(t))\, dt + \varepsilon C_\Upsilon \tag{5.4.24c}$$
$$+n\int_0^{1/n}\left(e^{-\lambda^{-1}t} - 1\right)\mathcal{L}(\gamma_n(t),\dot\gamma_n(t))\, dt + O\left(\frac{1}{n}\right). \tag{5.4.24d}$$

We show now that taking the limit in (5.4.24) as $n \to \infty$ leads to inequality (5.4.21). We analyze (5.4.13a),(5.4.24b),(5.4.24c), and (5.4.24d) separately.

As in the subsolution case, a key step in the analysis is to establish first

$$\sup_n \sup_{t \leq \frac{1}{n}} \frac{1}{t} \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds < \infty, \tag{5.4.25}$$
$$\gamma_n\left(\frac{1}{n}\right) \to x_0.$$

where the second follows from the first using Proposition 5.4.5. To do so, we again aim to apply Proposition 5.4.6. We prove in the following condition (5.4.5). By (5.4.22),

$$\begin{aligned}\int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds &= \int_0^{1/n} \langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle - \mathcal{H}(\gamma_n(s), \mathrm{d}f(\gamma_n(s)))\, ds \\ &\leq \int_0^{1/n} \langle \mathrm{d}f(\gamma_n(s)), \dot{\gamma}_n(s)\rangle\, ds + C_2,\end{aligned}$$

where in the last inequality we used Assumption (V) by taking into account that $\gamma_n(t) \in K$ for all $t \leq 1/n$. This concludes the proof of (5.4.5). Proceeding as in the subsolution proof, we can conclude that $\gamma_n(\frac{1}{n}) \to x^0$ as $n \to \infty$.

Limsup of (5.4.24a): Working towards the use of (5.4.20), note that

$$\limsup_n -n\left(e^{-\lambda^{-1}/n} - 1\right) R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) = \lambda^{-1} \limsup_n e^{-\lambda^{-1}/n} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right)$$

and that by (5.4.23)

$$\begin{aligned}&\limsup_n e^{-\lambda^{-1}/n} R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \\ &\quad \leq \limsup_n \left(R_{\lambda,h}(x_n) - \int_0^{1/n} \left[h(\gamma_n(t)) - \int_0^t \mathcal{L}(\gamma_n(s), \dot{\gamma}_n(s))\, ds\right] \lambda^{-1} e^{-\lambda^{-1} t}\, dt\right).\end{aligned}$$

We thus find that

$$\limsup_n -n\left(e^{-\lambda^{-1}/n} - 1\right) R_{\lambda,h}\left(\gamma_n\left(\frac{1}{n}\right)\right) \leq \lambda^{-1} (R_{\lambda,h})_*(x_0) \tag{5.4.26}$$

by (5.4.20), the fact that $\gamma_n(t) \in K$ and (5.4.25).

Limit of (5.4.24b): By the convergence of $\gamma_n(1/n)$ to x_0, the continuity of h and the dominated convergence theorem, the limit in (5.4.24b) is

$$\lim_{n \to \infty} -n \int_0^{1/n} \lambda^{-1} e^{-\lambda^{-1} 1/n} h\left(\gamma_n(t)\right)\, dt = -\lambda^{-1} h(x^0). \tag{5.4.27}$$

Limit of (5.4.24c): Recall that γ_n is constructed such that (5.4.22) holds. Then,

$$\begin{aligned}&-n \int_0^{1/n} (1+\varepsilon)\left(\langle \mathrm{d}f(\gamma_n(t)), \dot{\gamma}_n(t)\rangle - \mathcal{L}(\gamma_n(t), \dot{\gamma}_n(t))\right)\, dt \\ &= -n \int_0^{1/n} (1+\varepsilon) \mathcal{H}(\gamma_n(t), \mathrm{d}f(\gamma_n(t)))\, dt.\end{aligned}$$

This yields

$$\begin{aligned}
&\limsup_n -n(1+\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle - \mathcal{L}(\gamma_n(t),\dot\gamma_n(t))\,dt + \varepsilon C_\Upsilon \qquad (5.4.28)\\
&\leq \limsup_n -n\int_0^{1/n}(1+\varepsilon)\mathcal{H}(\gamma_n(t),\mathrm{d}f(\gamma_n(t)))\,dt+\varepsilon C_\Upsilon\\
&= -g^\varepsilon_\ddagger(x^0).
\end{aligned}$$

Limit of (5.4.24d)**:** Note that by Assumption (I),

$$\mathcal{L}(x,v) \geq -\mathcal{H}(x,0) = 0. \qquad (5.4.29)$$

Then, (5.4.24d) is bounded above by 0.

Taking the limit for $n\to\infty$ in (5.4.24) and putting together (5.4.26), (5.4.27), (5.4.28) and (5.4.29) we obtain that

$$0 \leq -\lambda^{-1}(R_{\lambda,h})_*(x_0) - g^\varepsilon_\dagger(x_0) - \lambda^{-1}h(x_0),$$

that concludes the proof. □

Proof of Theorem 5.3.4. The proof follows the same line as in Theorem 5.3.3. For completeness we give the main steps.

The subsolution property. Let $f^\varepsilon_\dagger$ as in Definition 5.3.1. Applying Proposition 5.4.3 with $\phi = \mathbf{v}_\lambda$ and $f = f^\varepsilon_\dagger$ and $h \in C^1([0,T])$, there exists a sequence (x_n,t_n) in a compact set converging to a point (x_0,t_0) and such that

$$\begin{aligned}
&\mathbf{v}_\lambda(x_n,t_n) - f^\varepsilon_\dagger(x_n) - h(t_n) \geq \sup(\mathbf{v}_\lambda - f^\varepsilon_\dagger - h) - \frac{1}{n},\\
&(\mathbf{v}_\lambda)^*(x_0,t_0) - f^\varepsilon_\dagger(x_0) - h(t_0) = \sup(\mathbf{v}_\lambda - f^\varepsilon_\dagger - h), \qquad (5.4.30)
\end{aligned}$$

and

$$\lim_n \mathbf{v}_\lambda(x_n,t_n) = (\mathbf{v}_\lambda)^*(x_0,t_0).$$

It thus suffices to establish that

$$\begin{cases}\partial_t h(t_0) + \lambda(\mathbf{v}_\lambda)^*(x_0,t_0) - g^\varepsilon_\dagger(x_0) \leq 0 & \text{if } t_0>0;\\ [\partial_t h(t_0) - g^\varepsilon_\dagger(x_0)] \wedge [(\mathbf{v}_\lambda)^*(t_0,x_0) - u_0(x)] \leq 0 & \text{if } t_0 = 0.\end{cases} \qquad (5.4.31)$$

Let $\gamma_n \in \mathit{Adm}$ be such that $\gamma_n(0) = x_n$ and almost optimizing (DPPt) at page 117, that is

$$\mathbf{v}_\lambda(x_n,t_n) \leq \int_0^{1/n} -e^{-\lambda s}\mathcal{L}(\gamma_n(s),\dot\gamma_n(s))\,ds + e^{-\frac{\lambda}{n}}\mathbf{v}_\lambda(\gamma_n(1/n),t_n-1/n)+\frac{1}{n^2}. \qquad (5.4.32)$$

Rewriting (5.4.30), and afterwards using the fundamental theorem of calculus and Assumption (III), we find

$$\begin{aligned}
&\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) - \mathbf{v}_\lambda(x_n, t_n) \\
&\leq f^\varepsilon_\dagger(\gamma_n(1/n)) - f^\varepsilon_\dagger(x_n) + h(t_n - 1/n) - h(t_n) + \frac{1}{n^2} \\
&= (1-\varepsilon)\left(f\left(\gamma_n\left(\frac{1}{n}\right)\right) - f(\gamma_n(x_n))\right) + \varepsilon\left(\Upsilon\left(\gamma_n\left(\frac{1}{n}\right)\right) - \Upsilon(\gamma_n(x_n))\right) \\
&\quad + h(t_n - 1/n) - h(t_n) + \frac{1}{n^2} \\
&\qquad \leq (1-\varepsilon)\int_0^{1/n} \langle \mathrm{d}f(\gamma_n(s)), \dot\gamma_n(s)\rangle \, ds + \varepsilon \int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + \varepsilon \frac{1}{n} C_\Upsilon \\
&\qquad\quad + h(t_n - 1/n) - h(t_n) + \frac{1}{n^2}.
\end{aligned} \tag{5.4.33}$$

Then, combining (5.4.32) and (5.4.33), we obtain

$$\begin{aligned}
&-(1-\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle \, ds - \varepsilon\int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds - \varepsilon\frac{1}{n}C_\Upsilon \\
&\quad + h(t_n) - h(t_n - 1/n) - \frac{1}{n^2} + (1 - e^{-\lambda 1/n})\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) \\
&\leq \int_0^{1/n} -e^{-\lambda s}\mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + \frac{1}{n^2}.
\end{aligned}$$

Dividing by $\frac{1}{n}$ and rearranging yields,

$$\begin{aligned}
&0 \leq n(e^{-\lambda\frac{1}{n}} - 1)\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) \\
&\ + n(h(t_n - 1/n) - h(t_n)) \qquad \qquad (5.4.34a)\\
&\ + n(1-\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle - \mathcal{L}(\gamma_n(s), \dot\gamma_n(s) \, ds + \varepsilon C_\Upsilon \\
&\ + n\int_0^{1/n} (1 - e^{-\lambda s})\mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + O\left(\frac{1}{n}\right).
\end{aligned}$$

We obtain (5.4.31) by taking the limsup in the separate terms of (5.4.34). From this point onward, the proof follows that of the subsolution part of Theorem 5.3.3, with the straightforward modification to the limit in (5.4.34a).

The supersolution property. Let $f^\varepsilon_\ddagger$ be as in Definition 5.3.1. Applying Proposition 5.4.3 to $\phi = -\mathbf{v}_\lambda$, there exists a sequence (x_n, t_n) converging to a point (x^0, t^0) and such that

$$\mathbf{v}_\lambda(x_n, t_n) - f^\varepsilon_\ddagger(x_n) - h(t_n) \leq \inf(\mathbf{v}_\lambda - f^\varepsilon_\ddagger - h) + \frac{1}{n^2}, \tag{5.4.35}$$

$$(\mathbf{v}_\lambda)_*(x^0, t^0) - f^\varepsilon_\ddagger(x^0) - h(t^0) = \inf(\mathbf{v}_\lambda - f^\varepsilon_\ddagger - h),$$

and

$$\lim_n \mathbf{v}_\lambda(x_n, t_n) = (\mathbf{v}_\lambda)_*(x_0, t_0).$$

It thus suffices to establish that

$$\begin{cases} \partial_t h(t_0) + \lambda(\mathbf{v}_\lambda)_*(x^0, t^0) - g^\varepsilon_\ddagger(x^0) \geq 0 & \text{if } t^0 > 0; \\ [\partial_t h(t^0) - g^\varepsilon_\ddagger(x^0)] \vee [(\mathbf{v}_\lambda)_*(t^0, x^0) - u_0(x)] \geq 0 & \text{if } t^0 = 0. \end{cases} \tag{5.4.36}$$

By Assumption (IV), there exists $\gamma_n \in Adm$ with $\gamma_n(0) = x_n$ and such that

$$\int_0^{1/n} \langle \mathrm{d}f(\gamma_n(t)), \dot\gamma_n(t)\rangle \, dt = \int_0^{1/n} \mathcal{H}(\gamma_n(t), \mathrm{d}f(\gamma_n(t))) + \mathcal{L}(\gamma_n(t), \dot\gamma_n(t)) \, dt.$$

As $\gamma_n(0) \in K_0$, we obtain, by Lemma 5.4.4 and Remark 5.3.2, that there exists a compact K such that

$$\gamma_n(t) \in K \qquad \text{for all } t \leq \frac{1}{n}.$$

By (5.4.35), the fundamental theorem of calculus, and Assumption (III)

$$\begin{aligned} &\mathbf{v}_\lambda(x_n, t_n) - \mathbf{v}_\lambda\left(\gamma_n\left(\frac{1}{n}\right), t_n - 1/n\right) \\ &\leq f^\varepsilon_\ddagger(x_n) - f^\varepsilon_\ddagger\left(\gamma_n\left(\frac{1}{n}\right)\right) + h(t_n) - h(t_n - 1/n) + \frac{1}{n^2} \\ &= (1+\varepsilon)\left(f(\gamma_n(x_n)) - f\left(\gamma_n\left(\frac{1}{n}\right)\right)\right) - \varepsilon\left(\Upsilon(\gamma_n(x_n)) - \Upsilon\left(\gamma_n\left(\frac{1}{n}\right)\right)\right) + \frac{1}{n^2} \\ &\qquad + h(t_n) - h(t_n - 1/n) + \frac{1}{n^2} \\ &\leq -(1+\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle \, ds + \varepsilon \int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + \varepsilon C_\Upsilon \\ &\qquad + h(t_n) - h(t_n - 1/n) + \frac{1}{n^2}. \end{aligned} \tag{5.4.37}$$

Moreover, by (DPPt) at pag. 117

$$\mathbf{v}_\lambda(x_n, t_n) \geq \int_0^{1/n} -e^{-\lambda s}\mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + e^{-\lambda \frac{1}{n}} \mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n). \tag{5.4.38}$$

Then, combining (5.4.38) and (5.4.37), we obtain

$$\begin{aligned} &\int_0^{1/n} -e^{-\lambda s}\mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds \\ &\leq \mathbf{v}_\lambda(x_n, t_n) - \mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) + (1 - e^{-\lambda 1/n})\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) \\ &\leq -(1+\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\left(\gamma_n(s)\right), \dot\gamma_n(s)\rangle \, ds + \varepsilon \int_0^{1/n} \mathcal{L}(\gamma_n(s), \dot\gamma_n(s)) \, ds + \varepsilon \frac{1}{n} C_\Upsilon \\ &\qquad + h(t_n) - h(t_n - 1/n) + (1 - e^{-\lambda 1/n})\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) + \frac{1}{n^2}. \end{aligned}$$

Dividing by $\frac{1}{n}$ we get

$$\begin{aligned}
&0 \leq -n(e^{-\lambda 1/n} - 1)\mathbf{v}_\lambda(\gamma_n(1/n), t_n - 1/n) \\
&- n(h(t_n - 1/n) - h(t_n)) \\
&- n(1+\varepsilon)\int_0^{1/n} \langle \mathrm{d}f\,(\gamma_n(s))\,, \dot\gamma_n(s)\rangle - \mathcal{L}(\gamma_n(s), \dot\gamma_n(s)\,ds + \varepsilon C_\Upsilon \\
&- n\int_0^{1/n} (1 - e^{-\lambda s})\mathcal{L}(\gamma_n(s), \dot\gamma_n(s))\,ds + O\left(\frac{1}{n}\right).
\end{aligned}$$

We establish (5.4.36) by taking the limsup for $n \to \infty$ for the separate terms of (5.4.39).

From this point onward, the proof is analogous to that of the supersolution part of Theorem 5.3.3. □

5.5 CONVEX HAMILTONIANS

In this section, we consider Hamiltonians $\mathcal{H} : T^*\mathcal{M} \to \mathbb{R}$ such that the map $p \mapsto \mathcal{H}(x,p)$ is convex for all $x \in \mathcal{M}$. This is a typical assumption and includes cases such as the Hamilton-Jacobi-Bellman equations.

We give in the following the correspondent assumptions to Assumption 5.3.5 in this context.

First, note that when $\mathcal{H}$ is convex, the operator $\mathcal{L} : T\mathcal{M} \to [0,\infty)$ is its Legendre transform.

Assumption 5.5.1. Let $\mathcal{H} : T^*\mathcal{M} \to \mathbb{R}$ and call $\mathbf{H}f(x) := \mathcal{H}(x, \mathrm{d}f(x))$ and $D(\mathbf{H}) \subseteq C_b^1(\mathcal{M})$ its domain. The following properties hold.

- (I) $p \mapsto \mathcal{H}(x,p)$ is convex for all $x \in \mathcal{M}$;
- (II) $\mathcal{H}(x,0) = 0$ for all $x \in \mathcal{M}$;
- (III) The map $(x,p) \mapsto \mathcal{H}(x,p)$ is continuous in x and p;
- (IV) There exists a containment function in the sense of Definition 5.2.1 such that
  - (a) $\Upsilon \in C^1(\mathcal{M})$;
  - (b) There exists a constant C_Υ such that $\sup_x \mathcal{H}(x, \mathrm{d}\Upsilon(x)) < C_\Upsilon$.
- (V) Let $T > 0$. For all $f \in D(\mathbf{H})$ and $x_0 \in \mathcal{M}$, there exists $\gamma \in Adm$ such that $\gamma(0) = x_0$ and

$$\int_0^T \langle \mathrm{d}f(\gamma(t)), \dot\gamma(t)\rangle\,dt = \int_0^T \mathcal{L}(\gamma(t), \dot\gamma(t)) + \mathcal{H}(\gamma(t), \mathrm{d}f(\gamma(t)))\,dt.$$

- (VI) For every compact set K and positive constant c,

$$\bar{\mathcal{H}}(K,c) := \sup_{|p|\leq c}\sup_{x\in K} \mathcal{H}(x,p) < \infty.$$

- (VII) The space $D(\mathbf{H})$ is convergence determining.

We show in the following that Assumption 5.5.1 (IV) implies Assumption 5.3.5 (III).

Lemma 5.5.2. *Consider* $\Upsilon : \mathcal{M} \to [0, \infty)$ *as in Assumption 5.5.1 (IV). Then, for all* $\gamma \in Adm$ *and* $T > 0$ *the following inequality holds*

$$\int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t))\, dt + TC_\Upsilon \geq \Upsilon(\gamma(T)) - \Upsilon(\gamma(0)),$$

that is, Assumption 5.3.5 (III) holds.

Proof. The inequality follows immediately by the Frenchel–Young's inequality applied to $\langle \mathrm{d}\Upsilon(\gamma(t)), \dot{\gamma}(t)\rangle$ and that by assumption we have

$$\mathcal{H}(\gamma(t), \mathrm{d}\Upsilon(\gamma(t))) < C_\Upsilon.$$

□

We also mention that, in this case, Assumption (V) is equivalent to solve the differential inclusion

$$\dot{\gamma}(t) \in \partial_p \mathcal{H}(\gamma(t), \mathrm{d}f(\gamma(t))).$$

We refer to [Roc70] and [Dei92] for details.

When the Hamiltonian $\mathbf{H}$ is convex, the two operators $H_\dagger$ and $H_\ddagger$ are actually an upper and lower bound for the initial Hamiltonian. More precisely the three operators are linked each other by the following proposition whose proof is standard and can be found for example in [KS21]. We include it for completeness.

Proposition 5.5.3. *Fix* $\lambda > 0$ *and* $h \in C_b(\mathcal{M})$*.*

(a) Every subsolution to $f - \lambda \mathbf{H} f = h$ *is also a subsolution to* $f - \lambda H_\dagger f = h$*.*
(b) Every supersolution to $f - \lambda \mathbf{H} f = h$ *is also a supersolution to* $f - \lambda H_\ddagger f = h$*.*
(c) Every subsolution to $\partial_t f + \lambda f(x,t) - \mathbf{H} f = 0$ *is also a subsolution to* $\partial_t f + \lambda f(x,t) - H_\dagger f = 0$*.*
(d) Every supersolution to $\partial_t f + \lambda f(x,t) - \mathbf{H} f = 0$ *is also a supersolution to* $\partial_t f + \lambda f(x,t) - H_\ddagger f = 0$*.*

Proof. We only prove (a) as the other claims can be carried out analogously. Fix $\lambda > 0$ and $h \in C_b(\mathcal{M})$. Let u be a subsolution to $f - \lambda \mathbf{H} f = h$. We prove it is also a subsolution to $f - \lambda H_\dagger f = h$.

Fix $\varepsilon > 0$ and $f \in C_\ell^\infty(\mathcal{M})$ and let $(f_\dagger^\varepsilon, g_\dagger^\varepsilon) \in H_\dagger$ as in Definition 5.3.1. We will prove that there are $x_n \in \mathcal{M}$ such that

$$\lim_{n\to\infty} \left(u - f_\dagger^\varepsilon\right)(x_n) = \sup_{x \in \mathcal{M}} \left(u(x) - f_\dagger^\varepsilon(x)\right), \tag{5.5.1}$$

$$\limsup_{n\to\infty} \left[u(x_n) - \lambda g_\dagger^\varepsilon(x_n) - h(x_n)\right] \leq 0. \tag{5.5.2}$$

As the function $[u - (1-\varepsilon)f]$ is bounded from above and $\varepsilon\Upsilon$ has compact sublevel-sets, the sequence x_n along which the first limit is attained can be assumed to lie in the compact set

$$K := \left\{ x \,|\, \Upsilon(x) \leq \varepsilon^{-1} \sup_x \left(u(x) - (1-\varepsilon) f(x)\right) \right\}.$$

Set $M = \varepsilon^{-1} \sup_x (u(x) - (1-\varepsilon) f(x))$. Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a smooth increasing function such that

$$\gamma(r) = \begin{cases} r & \text{if } r \leq M, \\ M+1 & \text{if } r \geq M+2. \end{cases}$$

Denote by f_ε the function on $\mathcal{M}$ defined by

$$f_\varepsilon(x) := \gamma\left((1-\varepsilon) f(x) + \varepsilon \Upsilon(x)\right).$$

By construction f_ε is smooth and constant outside of a compact set and thus lies in $\mathcal{D}(H) = C_{cc}^\infty(\mathcal{M})$. As u is a viscosity subsolution for $f - \lambda \mathbf{H} f = h$ there exists a sequence $x_n \in K \subseteq \mathcal{M}$ (by our choice of K) with

$$\lim_n \left(u - f_\varepsilon\right)(x_n) = \sup_x \left(u(x) - f_\varepsilon(x)\right), \tag{5.5.3}$$

$$\limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \leq 0. \tag{5.5.4}$$

As f_ε equals $f_\dagger^\varepsilon$ on K, we have from (5.5.3) that also

$$\lim_n \left(u - f_\dagger^\varepsilon\right)(x_n) = \sup_{x \in \mathcal{M}} \left(u(x) - f_\dagger^\varepsilon(x)\right),$$

establishing (5.5.1). Convexity of $p \mapsto \mathcal{H}(x,p)$ yields for arbitrary points $x \in K$ the estimate

$$\begin{aligned} \mathbf{H} f_\varepsilon(x) &= \mathcal{H}(x, \mathrm{d} f_\varepsilon(x)) \\ &\leq (1-\varepsilon)\mathcal{H}(x, \mathrm{d} f(x)) + \varepsilon \mathcal{H}(x, \mathrm{d}\Upsilon(x)) \\ &\leq (1-\varepsilon)\mathcal{H}(x, \mathrm{d} f(x)) + \varepsilon C_\Upsilon = g_\dagger^\varepsilon(x). \end{aligned}$$

Combining this inequality with (5.5.4) yields

$$\limsup_n \left[u(x_n) - \lambda g_\dagger^\varepsilon(x_n) - h(x_n)\right] \leq \limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \leq 0,$$

establishing (5.5.2). This concludes the proof. □

Using the comparison principle proved in Chapter 4 to $(R_{\lambda,h})^*$ and $(R_{\lambda,h})_*$ for the stationary case and to $(\mathbf{v}_\lambda)^*$ and $(\mathbf{v}_\lambda)_*$ for the evolutionary case, we obtain the following corollary.

Corollary 5.5.4. *Let Assumption 5.5.1 hold. Then, $R_{\lambda,h}$ and $\mathbf{v}_\lambda$ are the unique solutions of the pairs* (5.3.1) *and* (5.3.2)*. Moreover, let $\mathcal{H} : T^*\mathcal{M} \to \mathbb{R}$ be as in Chapter 4. Then, if $u - \lambda \mathbf{H} u = h$ (resp. $\partial_t u + \lambda u - \mathbf{H} u = 0$) admits a solution, this solution is unique and it is equal to $R_{\lambda,h}$ (resp. $\mathbf{v}_\lambda$).*

Proof. The uniqueness follows from Theorem 4.3.1 of Chapter 4.

If $u - \lambda \mathbf{H} u = h$ admits a solution $\mathbf{u}$, this is a subsolution and a supersolution of respectively $u - \lambda H_\dagger u = h$ and $u - \lambda H_\ddagger u = h$ by Proposition 5.5.3. Then, by uniqueness, it has to be $u = R_{\lambda,h}$. The same holds for the evolutionary case. □

5.6 HAMILTON-JACOBI-ISAACS EQUATIONS

In this section we consider the two operators

$$\begin{aligned} \mathbf{H}_1 f(x) &= \mathcal{H}_1(x, \mathrm{d}f(x)) = \sup_{\theta_1 \in \Theta_1} \inf_{\theta_2 \in \Theta_2} \left\{ \mathbf{H}_{\theta_1 \theta_2} f - \mathcal{I}(x, \theta_1, \theta_2) \right\} \\ \mathbf{H}_2 f(x) &= \mathcal{H}_2(x, \mathrm{d}f(x)) = \inf_{\theta_2 \in \Theta_2} \sup_{\theta_1 \in \Theta_1} \left\{ \mathbf{H}_{\theta_1 \theta_2} f - \mathcal{I}(x, \theta_1, \theta_2) \right\}, \end{aligned}$$

with Θ_1, Θ_2 two compact sets, $\mathbf{H}_{\theta_1\theta_2} f = \mathcal{H}_{\theta_1\theta_2}(x, \mathrm{d}f(x))$ a convex map and $\mathcal{I} : \mathcal{M} \times \Theta_1 \times \Theta_2 \to [0, \infty]$.

In this case, the equation is called Hamilton-Jacobi-Isaacs equation and it is commonly used in for example the context of robust control problems involving two players with conflicting interests.

We will also assume the following condition, known as *Isaacs condition*, that corresponds to say that the optimal strategies for both players can be determined by solving a single Hamilton-Jacobi equation, rather than separate equations for each player.

Assumption 5.6.1 (Isaacs condition)**.** The following equality holds

$$\mathbf{H}_1 f = \mathbf{H}_2 f$$

for any $f \in D(\mathbf{H}_1) = D(\mathbf{H}_2)$.

We will then consider the Hamiltonian

$$\mathbf{H} f(x) := \mathbf{H}_1 f(x) = \mathbf{H}_2 f(x). \tag{5.6.1}$$

In the following, we provide the counterpart to Assumption 5.3.5 within this context.

Assumption 5.6.2. Let $\mathbf{H}(x) = \mathcal{H}(x, \mathrm{d}f(x))$ as in (5.6.1). The following properties hold.

(I) $\mathcal{H}_{\theta_1,\theta_2}(x, 0) = 0$ for all $x \in \mathcal{M}$ and θ_1, θ_2;

(II) The map $(x, p) \mapsto \mathcal{H}(x, p)$ is continuous;

(III) There exists a containment function in the sense of Definition 5.2.1 such that
   - (a) $\Upsilon \in C^1(\mathcal{M})$;
   - (b) There exists a constant C_Υ such that $\sup_x \mathcal{H}_{\theta_1\theta_2}(x, \mathrm{d}\Upsilon(x)) < C_\Upsilon$ for all θ_1, θ_2.

(IV) Let $T > 0$. For all $f \in D(\mathbf{H})$ and $x_0 \in \mathcal{M}$, there exists $\gamma \in Adm$ such that $\gamma(0) = x_0$ and

$$\int_0^T \langle \mathrm{d}f(\gamma(t)), \dot\gamma(t) \rangle \, dt = \int_0^T \mathcal{L}(\gamma(t), \dot\gamma(t)) + \mathcal{H}_{\theta_1\theta_2}(\gamma(t), \mathrm{d}f(\gamma(t))) \, dt$$

for all $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$.

(V) For every compact $K \subseteq \mathcal{M}$, all θ_1, θ_2 and positive constant c,

$$\overline{\mathcal{H}}_{\theta_1\theta_2}(K, c) := \sup_{|p| \le c} \sup_{x \in K} \mathcal{H}_{\theta_1\theta_2}(x, p) < \infty.$$

(VI) The space $\bigcap_{\theta_1,\theta_2} D(\mathbf{H}_{\theta_1\theta_2})$ is convergence determining.

We show in the following that Assumption 5.6.2 imply Assumption 5.3.5.

Lemma 5.6.3. *Assume Assumption 5.6.1 and Assumption 5.6.2. Then, Assumption 5.3.5 holds.*

Proof. The proofs of Assumption 5.3.5 (I), (II) and (V) are trivial.

Assumption 5.3.5 (III) follows as in the proof given in Lemma 5.5.2 by observing that

$$\mathcal{H}(\gamma(t), \mathrm{dY}(\gamma(t))) < \sup_{\theta_1} \mathcal{H}_{\theta_1\theta_2}(\gamma(t), \mathrm{dY}(\gamma(t))) < C_{\mathrm{Y}},$$

where we used that $\mathcal{I} \geq 0$.

The same strategy can be applied to prove Assumption 5.3.5 (V).

Finally, we prove Assumption 5.3.5 (IV). First of all, recall that the first inequality is simply the Fenchel-Young's inequality (5.2.2) and it is implied by the definition of $\mathcal{L}$. We only need to prove the opposite inequality.

Let $\gamma \in Adm$ be as in Assumption 5.6.2 (IV). Let $\theta_1^* \in \Theta_1$ be such that

$$\sup_{\theta_1} \inf_{\theta_2} \mathbf{H}_{\theta_1\theta_2} f - \mathcal{I}(x, \theta_1, \theta_2) = \inf_{\theta_2} \mathbf{H}_{\theta_1^*\theta_2} f - \mathcal{I}(x, \theta_1^*, \theta_2).$$

Then, we have

$$\begin{aligned}
&\int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t)) + \mathcal{H}(\gamma(t), \mathrm{d}f(\gamma(t)))\, dt \\
&= \int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t)) + \inf_{\theta_2} \mathbf{H}_{\theta_1^*\theta_2} f(\gamma(t)) - \mathcal{I}(x, \theta_1^*, \theta_2)\, dt \\
&\leq \int_0^T \mathcal{L}(\gamma(t), \dot{\gamma}(t)) + \mathbf{H}_{\theta_1^*\theta_2} f(\gamma(t)) \\
&\leq f(\gamma(T)) - f(\gamma(0)).
\end{aligned}$$

This concludes the proof. □

Remark 5.6.4. We want to point out that, even if the methods to prove the existence of the curve in Assumption 5.3.5 (IV) are typically challenging for non convex Hamiltonians, in this scenario it is sufficient to solve the differential inclusion in terms of the internal (and convex) Hamiltonian.

We conclude this section by showing the relation between the Hamiltonian (5.6.1) and $H_\dagger$ and $H_\ddagger$.

Proposition 5.6.5. *Let* $\mathbf{H}$ *be as in* (5.6.1)*. Fix* $\lambda > 0$ *and* $h \in C_b(\mathcal{M})$*.*

(a) Every subsolution to $f - \lambda \mathbf{H} f = h$ *is also a subsolution to* $f - \lambda H_\dagger f = h$*.*

(b) Every supersolution to $f - \lambda \mathbf{H} f = h$ *is also a supersolution to* $f - \lambda H_\ddagger f = h$*.*

(c) Every subsolution to $\partial_t f + \lambda f(x,t) - \mathbf{H} f = 0$ *is also a subsolution to* $\partial_t f + \lambda f(x,t) - H_\dagger f = 0$*.*

(d) Every supersolution to $\partial_t f + \lambda f(x,t) - \mathbf{H} f = 0$ is also a supersolution to $\partial_t f + \lambda f(x,t) - H_{\ddagger} f = 0$.

Proof. The proof follows the same line of the proof of Proposition 5.5.3.

Let u be a subsolution to $f - \lambda \mathbf{H} f = h$. We prove it is also a subsolution to $f - \lambda H_{\dagger} f = h$. Fix $\varepsilon > 0$ and $f \in C_\ell^\infty(\mathcal{M})$ and let $(f_\dagger^\varepsilon, g_\dagger^\varepsilon) \in H_\dagger$ as in Definition 5.3.1. We construct f_ε as in the proof of Proposition 5.5.3.

As u is a viscosity subsolution for $f - \lambda \mathbf{H} f = h$ there exists a sequence $x_n \in K \subseteq \mathcal{M}$ with

$$\begin{aligned} &\lim_n \left(u - f_\varepsilon\right)(x_n) = \sup_x \left(u(x) - f_\varepsilon(x)\right), \\ &\limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \leq 0. \end{aligned} \tag{5.6.2}$$

It follows, as in the proof of Proposition 5.5.3 that

$$\lim_n (u - f_\dagger^\varepsilon)(x_n) = \sup_{x \in \mathcal{M}} (u(x) - f_\dagger^\varepsilon(x)).$$

For any θ_1 Let $\theta_2^* = \theta_2^*(\theta_1)$ be optimal for the infimum

$$\inf_{\theta_2} \left\{ H_{\theta_1, \theta_2} f(x) - \mathcal{I}(x, \theta_1, \theta_2) \right\}.$$

Using convexity of H_{θ_1, θ_2^*} for any θ_1 and taking into account that $\mathcal{I} \geq (1-\varepsilon)\mathcal{I}$, since $\mathcal{I} \geq 0$, we have

$$\begin{aligned} \mathbf{H} f_\varepsilon &\leq \sup_{\theta_1} \mathbf{H}_{\theta_1 \theta_2^*} f_\varepsilon - \mathcal{I}(x, \theta_1, \theta_2^*) \\ &\leq \sup_{\theta_1} (1-\varepsilon) \mathbf{H}_{\theta_1 \theta_2^*} f + \varepsilon \mathbf{H}_{\theta_1 \theta_2^*} \Upsilon - \mathcal{I}(x, \theta_1, \theta_2^*) \\ &\leq (1-\varepsilon) \sup_{\theta_1} \mathbf{H}_{\theta_1 \theta_2^*} f + \varepsilon C_\Upsilon - \mathcal{I}(x, \theta_1, \theta_2) \\ &\leq (1-\varepsilon) \sup_{\theta_1} \{ \mathbf{H}_{\theta_1 \theta_2^*} f - \mathcal{I}(x, \theta_1, \theta_2) \} + \varepsilon C_\Upsilon \\ &= (1-\varepsilon) \mathbf{H} f + \varepsilon C_\Upsilon = g_\dagger(x). \end{aligned}$$

Combining this inequality with (5.6.2) yields

$$\limsup_n \left[u(x_n) - \lambda g_\dagger^\varepsilon(x_n) - h(x_n)\right] \leq \limsup_n \left[u(x_n) - \lambda \mathbf{H} f_\varepsilon(x_n) - h(x_n)\right] \leq 0.$$

This concludes the proof. □

5.7 CONCLUSIONS AND FUTURE PERSPECTIVES

In this chapter, we have addressed the existence of viscosity solutions for first-order Hamilton-Jacobi equations using Lyapunov control techniques. This work complements the analysis of the comparison principle developed in the previous chapter. With this chapter, we conclude our discussion on first-order Hamilton-Jacobi equations. The results obtained set the stage for the second-order case, which we explore in the next chapter.

We conclude with some interesting future directions:

- Our framework relies on the assumption that the Lagrangian $\mathcal{L}$ is positive, which follows from the condition $H(x,0)=0$. This assumption can be modified to allow for a controlled lower and upper bounds, for instance,

$$\exists C_1 \in \mathbb{R} \,:\, \forall \gamma \in Adm, \quad \forall T>0, \quad \int_0^T \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds \geq TC_1,$$
$$\exists C_2 \in \mathbb{R} \,:\, \forall x \in \mathcal{M}, \quad \exists \gamma \in Adm, \quad \gamma(0)=x,: \int_0^T \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds \leq TC_2.$$

  Exploring how this relaxation affects the existence theory could broaden the applicability of our results to more general control problems.
- The positivity assumption on $\mathcal{L}$ is crucial in our setting because we work with bounded solutions. A more general approach would involve unbounded solutions, where what is really required is that the solution grows at most as fast as the Lyapunov function Υ. More precisely, ensuring that $R_{\lambda,h}/\Upsilon$ remains bounded, rather than $R_{\lambda,h}$ itself, would allow for greater flexibility. This condition could be achieved through an assumption of the form:

$$\begin{aligned}
&\exists C_1 \in \mathbb{R} \,:\, \forall \gamma \in Adm, \quad \forall T>0,\\
&\qquad [\Upsilon(\gamma(T)) - \Upsilon(\gamma(0))] \leq TC_1 + \int_0^T \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds,\\
&\exists C_2 \in \mathbb{R} \,:\, \forall x \in \mathcal{M}, \quad \exists \gamma \in Adm, \quad \gamma(0)=x,:\\
&\qquad [\Upsilon(\gamma(T)) - \Upsilon(\gamma(0))] \leq TC_2 - \int_0^T \mathcal{L}(\gamma(s),\dot\gamma(s))\,ds.
\end{aligned}$$

- Another interesting direction is to directly analyze a system of Hamilton-Jacobi equations, proving the existence of a subsolution for one equation and a supersolution for the other. This approach would allow the assumptions to be split between the two Hamiltonians. For instance, instead of requiring the Hamiltonian H to be continuous, one could impose upper semicontinuity for H_1 and lower semicontinuity for H_2. This could lead to more refined results under weaker regularity assumptions.

6 SECOND ORDER HAMILTON-JACOBI EQUATIONS

In the previous chapters, we studied first-order Hamilton-Jacobi equations, focusing on their well-posedness and applications in large deviations theory. However, many real-world problems, particularly those involving stochastic dynamics, require us to consider additional effects, such as diffusion terms. This naturally leads to the study of second-order Hamilton-Jacobi equations, which will be the focus of this chapter.

In Chapter 1, we observed that the comparison principle for second-order Hamilton-Jacobi equations is more challenging compared to the one for first-order equations and requires additional techniques to deal with the second-order parts. In this chapter, we consider Hamilton-Jacobi, Hamilton-Jacobi-Bellman, and Isaacs equations and present a new proof of the comparison principle. We focus, in particular, on Hamiltonians with a partial integro-differential form that arise in many contexts, such as stock price movements and option pricing. Consider, for example, trading on the stock exchange, such as buying and selling stocks. To find the optimal policy, one has to consider a stochastic controlled problem where the control models the trading action. This leads to a Hamilton-Jacobi-Bellman equation with a Hamiltonian that is typically in an integro-differential form (see, e.g. [FS06], [CK15]).

The results presented in this chapter are based on:

S.Della Corte, F. Fuchs, R.C. Kraaij and M. Nendel, ”A comparison principle based on couplings of partial integro-differential operators”, Preprint, 2024.

6.1 INTRODUCTION

In this chapter, we provide a new perspective on comparison principles for viscosity solutions to the Hamilton–Jacobi equation

$$f - \lambda H f = h, \qquad \lambda > 0,\ h \in C_b(\mathbb{R}^q), \tag{6.1.1}$$

for Hamiltonians H of the type

$$\begin{aligned} Hf(x) = \langle b(x), \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x) D^2 f(x)\right) \\ + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\,\langle \mathbf{z}, \nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}) + \mathcal{H}(\nabla f(x)) \end{aligned} \tag{6.1.2}$$

and, more generally, for those in Bellman and Isaacs form

$$Hf(x) = \sup_{\theta \in \Theta} \{H_\theta f(x) - \mathcal{I}(x,\theta)\} \quad \text{and}$$
$$Hf(x) = \sup_{\theta_1 \in \Theta_1} \inf_{\theta_2 \in \Theta_2} \{H_{\theta_1,\theta_2} f(x) - \mathcal{I}(x,\theta_1,\theta_2)\},$$

with H_θ and H_{θ_1,θ_2} as in (6.1.2) but with θ and (θ_1, θ_2) dependent coefficients, respectively, and an appropriate cost functional $\mathcal{I}$.

Motivated by convex Hamiltonians, for which no unique classical or weak solutions exist in general, [CL83] introduced the notion of viscosity solutions. The seminal works [Lio84; CEL84; Ish84; Ish86; CIL87b] explore this framework for first-order equations.

Most modern comparison proofs for operators containing second-order terms are based on results of [JLS88; Jen88]. Using then recent advances for generalized differentials, [CI90] provided what is nowadays known as the Crandall–Ishii Lemma. An overview over uniqueness results for viscosity solutions to degenerate elliptic equations is given in the User's Guide [CIL92].

The treatment of non-local operators was initially motivated by problems in optimal control theory; see [Son86; Awa91; AT96] for early examples with non-local operators. The work [BI08] gives a non-local version of the Crandall–Ishii Lemma by adapting the original procedure in [CIL92], and [Hol16] extends these results to unbounded solutions. We also refer to [FGS17] for an overview of the Hilbertian setting, [Ble+23] for comparison principles for convex monotone semigroups on spaces of continuous functions, to [DKN21] for the classical well-posedness of convex Cauchy problems on L^p, to [HP21] for a comparison principle in the framework of G-Lévy processes, and to [Ber24] for a comparison principle for HJB equations on the set of probability measures.

Our approach and our main results, Theorem 6.3.1 and Corollary 6.3.2, innovate upon classical comparison principles in the following three ways:

(1) We reinterpret the classical doubling-of-variables method in the context of second-order equations by casting the Crandall–Ishii Lemma into a test function framework. This adaptation allows us to effectively handle non-local integral operators, such as generators of Lévy processes, in the same framework as second-order operators, paving the way for stability results.
(2) We translate the key estimate on the difference of Hamiltonians in terms of an adaptation of the probabilistic notion of couplings, providing a unified approach that applies to both continuous and discrete operators. We point out that [CIL87b] also discusses a coupling point of view, but only for first order operators.
(3) We strengthen the typical comparison principle using Lyapunov functionals from a sup-norm contractivity result to what we call the *strict comparison principle*, cf. Definition 6.2.2, which encodes continuity in the strict or sometimes also called mixed topology, cf. [Buc58; Sen72].

The results are illustrated in various examples in Section 6.4. To introduce the first two innovations, we heuristically trace back the classical doubling-of-variables procedure used to obtain comparison principles for first and second-order equations. For the sake of exposition, we focus on the θ-independent case.

Given a subsolution u and supersolution v to an equation of type (6.1.1) and, for $\alpha > 1$, optimizers (x_α, y_α) to

$$u(x_\alpha) - v(y_\alpha) - \frac{\alpha}{2} d^2(x_\alpha, y_\alpha) = \sup_{x,y\in\mathbb{R}^q} \left\{ u(x) - v(y) - \frac{\alpha}{2} d^2(x,y) \right\}, \tag{6.1.3}$$

one estimates

$$\sup_{x\in\mathbb{R}^q} u(x) - v(x) \leq h(x_\alpha) - h(y_\alpha) + \lambda \left[H\left(\frac{\alpha}{2} d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2} d^2(x_\alpha, \cdot)\right)(y_\alpha) \right].$$

Consequently, comparison then holds, if

$$\liminf_{\alpha\to\infty} H\left(\frac{\alpha}{2} d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2} d^2(x_\alpha, \cdot)\right)(y_\alpha) \leq 0. \tag{6.1.4}$$

The estimate (6.1.4), then translates into explicit conditions on H.

When H is, for example, of the form

$$Hf(x) = \langle b(x), \nabla f(x) \rangle + \frac{1}{2} |\nabla f(x)|^2,$$

the estimate (6.1.4) translates into

$$\begin{aligned} &H\left(\frac{\alpha}{2} d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2} d^2(x_\alpha, \cdot)\right)(y_\alpha) \\ &= \left[\langle b(x_\alpha), \alpha(x_\alpha - y_\alpha) \rangle + \frac{\alpha^2}{2} d^2(x_\alpha, y_\alpha) \right] - \left[\langle b(y_\alpha), \alpha(x_\alpha - y_\alpha) \rangle + \frac{\alpha^2}{2} d^2(x_\alpha, y_\alpha) \right] \\ &\leq \langle b(x_\alpha) - b(y_\alpha), \alpha(x_\alpha - y_\alpha) \rangle, \end{aligned}$$

which goes to 0 for $\alpha \to \infty$, if b is one-sided Lipschitz.

For second order operators, however, the same strategy fails since, considering, for example, the Laplacian $Hf(x) = \frac{1}{2}\Delta f(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$, we get

$$H\left(\frac{\alpha}{2} d^2(\cdot, y_\alpha)\right)(x_\alpha) - H\left(-\frac{\alpha}{2} d^2(x_\alpha, \cdot)\right)(y_\alpha) = 2\alpha,$$

which diverges as $\alpha \to \infty$.

The works [JLS88; Jen88] use the key insight that, while the first order-viscosity solution method explores the sequences of optimizers of (6.1.3) separately (fix y_α and vary x for the subsolution part and vice versa), for second order equations, one needs to treat the two sequences jointly. This insight was later formalized in [CI90] and as Theorem 3.2 in the User's Guide [CIL92], now known as the Crandall–Ishii Lemma. The lemma states for equations of type $Hf(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$ that, given $X_\alpha = D^2 u(x_\alpha)$ and $Y_\alpha = D^2 v(y_\alpha)$ or their appropriate generalizations, we have the estimate

$$\begin{pmatrix} X_\alpha & 0 \\ 0 & -Y_\alpha \end{pmatrix} \leq 3\alpha \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}.$$

Conjugating the matrices with

$$C := \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix}, \tag{6.1.5}$$

i.e. essentially using C to couple the subsolution and supersolution problems, we arrive at the desired estimate

$$\frac{1}{2}\operatorname{Tr}(X_\alpha) - \frac{1}{2}\operatorname{Tr}(Y_\alpha) = \frac{1}{4}\operatorname{Tr}\begin{pmatrix} X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \\ X_\alpha - Y_\alpha & X_\alpha - Y_\alpha \end{pmatrix} \leq 0. \tag{6.1.6}$$

We now briefly describe the three innovations (1)–(3).

Innovation 1: A test function framework. Examining the proof of the Crandall–Ishii Lemma, we can interpret the procedure as the construction of two test functions $\phi_\alpha, \psi_\alpha \in C^2(\mathbb{R}^q)$ that are squeezed between u and v on one-hand and $\frac{\alpha}{2}d^2$ on the other. To be more precise, we find $\phi_\alpha, \psi_\alpha \in C^2(\mathbb{R}^q)$ such that

$$u(x_\alpha) - \phi_\alpha(x_\alpha) = \sup_{x\in\mathbb{R}^q}\{u(x) - \phi_\alpha(x)\} \quad \text{and} \quad v(y_\alpha) - \psi_\alpha(y_\alpha) = \inf_{y\in\mathbb{R}^q}\{v(y) - \psi_\alpha(y)\},$$

and

$$\phi_\alpha(x_\alpha) - \psi_\alpha(y_\alpha) - \frac{\alpha}{2}d^2(x_\alpha, y_\alpha) = \sup_{x,y\in\mathbb{R}^q}\left\{u(x) - v(y) - \frac{\alpha}{2}d^2(x,y)\right\}. \tag{6.1.7}$$

As before, comparison now follows from the estimate

$$\liminf_{\alpha\to\infty} H\phi_\alpha(x_\alpha) - H\psi_\alpha(y_\alpha) \leq 0.$$

For the Laplacian $Hf(x) = \frac{1}{2}\operatorname{Tr}\left(D^2 f(x)\right)$, this translates to

$$H\phi_\alpha(x_\alpha) - H\psi_\alpha(y_\alpha) = \frac{1}{2}\operatorname{Tr}(D^2\phi_\alpha(x_\alpha)) - \frac{1}{2}\operatorname{Tr}(D^2\psi_\alpha(y_\alpha)). \tag{6.1.8}$$

At this point in proofs using the Crandall–Ishii Lemma, the estimate (6.1.6) is performed by conjugation with the matrix C in (6.1.5). We formalize this step by adapting the probabilistic notion of couplings, cf. [Lin92; Tho00; BK00], and identify the choice of the matrix C in (6.1.5) with the *synchronous coupling* (also called *co-monotone coupling*).

Innovation 2: The coupling approach. Indeed, given two Brownian motions starting in x and y, one can construct a coupling of the two by considering

$$(X(t), Y(t)) = (x + B(t), y + B(t)), \tag{6.1.9}$$

where $B(t)$ is a standard Brownian motion. The generator of the coupled process (6.1.9) is given by

$$\widehat{H}g(x,y) := \frac{1}{2}\left(\partial_x + \partial_y\right)^2 g(x,y) = \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} D^2 g(x,y)\right) = \frac{1}{2}\operatorname{Tr}\left(C D^2 g(x,y) C^T\right),$$

where we recover the matrix C of (6.1.5). Note that $\widehat{H}$ is indeed a coupling: For $f_1, f_2 \in C_b(\mathbb{R}^q)$ and $(f_1 \oplus f_2)(x,y) := f_1(x) + f_2(y)$, we have

$$\widehat{H}(f_1 \oplus f_2)(x,y) = Hf_1(x) + Hf_2(y).$$

Using the coupling $\widehat{H}$, we can now rewrite (6.1.8) as

$$\begin{aligned} H\phi_\alpha(x_\alpha) - H\psi_\alpha(y_\alpha) &= \widehat{H}\left(\phi_\alpha \oplus -\psi_\alpha\right)(x_\alpha, y_\alpha) \\ &\leq \widehat{H}\left(\frac{\alpha}{2}d^2\right)(x_\alpha, y_\alpha) = 0, \end{aligned} \tag{6.1.10}$$

where the first equality follows by the definition of a coupling, the inequality is based on the positive maximum principle with the optimizers from equation (6.1.7), and the final equality is due to the fact that the synchronous coupling controls distance growth.

A similar strategy can be used to treat a discretized version of the Brownian Motion by considering the generator $Hf(x) = \frac{1}{2}\left[f(x+1) - f(x)\right] + \frac{1}{2}\left[f(x-1) - f(x)\right]$ of a random walk: We synchronously couple the random walk with itself using the operator

$$\widehat{H}f(x,y) = \frac{1}{2}\left[f(x+1, y+1) - f(x,y)\right] + \frac{1}{2}\left[f(x-1, y-1) - f(x,y)\right].$$

The argument in (6.1.10) then works for the random walk exactly as it did for the Brownian motion.

This coupling approach is one of the main contributions of this paper, allowing for a unifying framework to show comparison for Hamilton–Jacobi equations with Hamiltonians of type (6.1.2) and their Bellman and Isaacs versions, cf. Theorem 6.3.1 and Corollary 6.3.2.

Innovation 3: The strict comparison principle. Our third innovation is on the final estimate that is obtained as the comparison principle. For a subsolution u to

$$f - \lambda Hf = h_1$$

and a supersolution v to

$$f - \lambda Hf = h_2$$

the comparison principle amounts to establishing that

$$\sup_{x \in \mathbb{R}^q} u(x) - v(x) \leq \sup_{x \in \mathbb{R}^q} h_1(x) - h_2(x).$$

The comparison principle, once established, thus implies sup-norm contractivity for the solution map $R(\lambda) : C_b(\mathbb{R}^q) \to C_b(\mathbb{R}^q)$, where $R(\lambda)h$ is the unique viscosity solution for the Hamilton–Jacobi equation (6.1.1).

It is well-known from examples, cf. [BC97; YZ99; CS04; FS06], that the map $R(\lambda)h$ takes the form of an exponentially discounted Markovian control problem. If the dynamics admits a Lyapunov function V, having compact sublevel sets and satisfying $HV \leq c$, then the controlled Markov processes satisfy tightness properties. More precisely, if the controlled process starts in a compact set K, one can find, for any time horizon $T > 0$ and $\varepsilon > 0$, a compact set $\widehat{K} \supseteq K$, given in terms of the sublevel sets of V such that, with probability $1 - \varepsilon$, the process remains in $\widehat{K}$ up to time T. Rewriting this in terms of an estimate on the solution map $R(\lambda)$, we then find

$$\sup_{x \in K} R(\lambda)h_1(x) - R(\lambda)h_2(x) \leq \varepsilon \, ||h_1 - h_2|| + \sup_{x \in \widehat{K}} h_1(x) - h_2(x). \tag{6.1.11}$$

Estimates of this type are indeed characterized by the strict topology, as was first established for linear functionals in [Sen72, Theorem 5.1] and for convex, monotone functionals in [Nen24, Corollary 2.10]. Note that in this paper, we do not establish convexity of $h \mapsto R(\lambda)h$, but want to point out that given a convex H, convexity of $R(\lambda)h$ is to be expected by performing a comparison principle in terms of three variables using variants of the, e.g., three dimensional Theorem 3.2 of [CIL92], see also the domination principle of Theorem 2.22 and Corollary 2.26 of [Hol16]. We leave this for future work.

Building upon the notion of Lyapunov functions, we will show that we can directly establish a variant of (6.1.11) for a subsolution u and a supersolution v. Given its motivation, we will call this estimate the *strict comparison principle*, see Definition 6.2.2 and the main result, Theorem 6.3.1, below.

The rest of the chapter is organized as follows: Section 6.2 introduces the notation and definitions. Section 6.3 introduces the framework by stating the necessary assumptions and formalizing the main results. In Section 6.4, we show how to apply our framework to operators of the form (6.1.2). Section 6.5 contains the construction of the required optimizing points and test functions. Finally, Section 6.6 contains the proof of the main theorems.

6.2 PRELIMINARIES AND GENERAL SETTING

6.2.1 *Notation and Preliminaries*

Throughout the chapter, let $q \in \mathbb{N}$ and $E = \mathbb{R}^q$.

Moreover, we write

$$C_+(E) := \{f \in C(E) \mid f \text{ has compact sub-level sets}\},$$
$$C_-(E) := \{f \in C(E) \mid f \text{ has compact super-level sets}\},$$
$$C_c(E) := \{f \in C(E) \mid f \text{ is constant outside of a compact set}\}.$$

We furthermore define the following intersections: $C_c^2(E) = C_c(E) \cap C^2(E)$,

$$C_+^2(E) := C_+(E) \cap C^2(E), \qquad C_-^2(E) := C_-(E) \cap C^2(E).$$

For $a, b \in \mathbb{R}$, we write $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. We denote the supremum norm by $||\cdot||$, that is

$$||f|| = \sup_{x \in E} |f(x)|,$$

for $f \in C_b(E)$, while, for $u \in C(E)$, we use the notation

$$\lceil u \rceil := \sup_{x \in E} u(x), \qquad \lfloor u \rfloor := \inf_{x \in E} u(x)$$

for a supremum or infimum over the entire space and

$$\lceil u \rceil_C := \sup_{x \in C} u(x), \qquad \lfloor u \rfloor_C := \inf_{x \in C} u(x)$$

for a supremum or infimum over a subset $C \subseteq E$.

We say that a function $\omega\colon [0, \infty) \to [0, \infty)$ is a *modulus of continuity*, if ω is upper semi-continuous with $\omega(0) = 0$. We say that a function $f \in C(E)$ *admits a modulus of continuity*, if, for every compact $K \subseteq E$, there exists a modulus of continuity $\omega_K\colon [0, \infty) \to [0, \infty)$ such that, for all $x, y \in K$, we have

$$|f(x) - f(y)| \leq \omega_K(d(x, y)).$$

A function $\phi\colon E \to \mathbb{R}$ is called *semi-convex* with constant $\kappa \in \mathbb{R}$ if for any $x_0 \in E$ the map

$$x \mapsto \phi(x) + \frac{\kappa}{2} d^2(x, x_0)$$

is convex. Moreover, ϕ is called *semi-concave* with constant $\kappa \in \mathbb{R}$ if $-\phi$ is semi-convex with constant $-\kappa$.

We say that a function $f \in C(E, \mathbb{R}^q)$ is *one-sided Lipschitz* if, for all $x, y \in E$ and some constant $C \in \mathbb{R}$, we have

$$\langle x - y, f(x) - f(y) \rangle \leq C d^2(x, y).$$

For any $z \in E$, let $s_z : E \to \mathbb{R}^q$ be the *shift map*

$$s_z(x) = x - z.$$

For any $z_1, z_2 \in E$, let

$$d_{z_1, z_2}(x, y) := d\left(s_{z_1}(x), s_{z_2}(y)\right).$$

Let $f_1, f_2 \in C(E)$. Then, we define the *direct sum* $f_1 \oplus f_2, f_1 \ominus f_2 \in C(E \times E)$ as

$$(f_1 \oplus f_2)(x_1, x_2) := f_1(x_1) + f_2(x_2) \quad \text{and} \quad (f_1 \ominus f_2)(x_1, x_2) := f_1(x_1) - f_2(x_2)$$

for all $x_1, x_2 \in E$. For two sets of functions $F_1, F_2 \subseteq C(E)$, we define

$$F_1 \oplus F_2 := \{f_1 \oplus f_2 \,|\, f_1 \in F_1, f_2 \in F_2\} \quad \text{and} \quad F_1 \ominus F_2 := \{f_1 \ominus f_2 \,|\, f_1 \in F_1, f_2 \in F_2\}.$$

6.2.2 *Operator notions*

We consider operators $H \subseteq C(E) \times C(E)$, where we identify H by its graph. As usual, the *domain* of H is given by

$$\mathcal{D}(H) := \{f \in C(E) \,|\, \exists g \in C(E) \colon (f,g) \in H\}.$$

Let $H_1, H_2 \subseteq C(E) \times C(E)$. We define

$$H_1 + H_2 := \{(f, g_1 + g_2) \,|\, (f, g_1) \in H_1, (f, g_2) \in H_2\},$$

which is an operator with domain

$$\mathcal{D}(H_1 + H_2) := \mathcal{D}(H_1) \cap \mathcal{D}(H_2).$$

We say that H is *linear on its domain* if, for any $f, g \in \mathcal{D}(H)$ and $a \in \mathbb{R}$ such that $af + g \in \mathcal{D}(H)$, we have

$$H(af + g) = aHf + Hg.$$

We will prove the comparison principle for the equation in terms of H by relating it to two equations in terms of two restrictions of H. To do so, we will need to be able to construct test functions in the domain of H from functions in the domain of the restrictions. In particular, we will need the following notion.

Definition 6.2.1 (Sequential Denseness)**.** Let $\mathcal{D} \subseteq C_b(E)$, $\mathcal{D}_+ \subseteq C_+(E)$, and $\mathcal{D}_- \subseteq C_-(E)$.

- We say that $\mathcal{D}$ is *upward sequentially dense* in $\mathcal{D}_+$ if, for any $f_\dagger \in \mathcal{D}_+$ and constant $a \in \mathbb{R}$, there exists a function $f_{\dagger,a} \in \mathcal{D}$ such that

$$\begin{cases} f_{\dagger,a}(x) = f_\dagger(x) & \text{if } f_\dagger(x) \leq a, \\ a < f_{\dagger,a}(x) \leq f_\dagger(x) & \text{if } f_\dagger(x) > a. \end{cases}$$

- We say that $\mathcal{D}$ is *downward sequentially dense* in $\mathcal{D}_-$ if, for any $f_\ddagger \in \mathcal{D}_-$ and constant $a \in \mathbb{R}$, there exists a function $f_{\ddagger,a} \in \mathcal{D}$ such that

$$\begin{cases} f_{\ddagger,a}(x) = f_\dagger(x) & \text{if } f_\ddagger(x) \geq a, \\ a > f_{\ddagger,a}(x) \geq f_\ddagger(x) & \text{if } f_\ddagger(x) < a. \end{cases}$$

6.2.3 *Viscosity solutions*

For $\lambda > 0$, consider $h_1 \in C_l(E)$ and $h_2 \in C_u(E)$ and two operators $H_1 \subseteq C_l(E) \times C(E)$ and $H_2 \subseteq C_u(E) \times C(E)$. We study the pair of equations

$$f - \lambda H_1 f \leq h_1, \tag{6.2.1}$$
$$f - \lambda H_2 f \geq h_2. \tag{6.2.2}$$

We already introduced the notion of viscosity solutions in Section 1.1.2. We also saw in the same section that associated with the definition of viscosity solutions, there is the comparison principle, which, for $h_1 = h_2$, implies uniqueness in the viscosity sense for solutions of the Hamilton–Jacobi equation $f - \lambda Hf = h$. We additionally introduce a new, stronger notion: the strict comparison principle. This name is inspired by the observation that the comparison principle implies contractivity in the sup-norm of the solution map. The strict comparison principle implies continuity in terms of the weaker strict topology, see e.g. [Sen72].

Definition 6.2.2. We say that the equations (6.2.1) and (6.2.2) satisfy

(a) the *comparison principle* if, for any subsolution u to (6.2.1) and any supersolution v to (6.2.2), we have

$$\sup_{x\in E} u(x) - v(x) \leq \sup_{x\in E} h_1(x) - h_2(x).$$

(b) the *strict comparison principle* if, for any subsolution u to (6.2.1), any supersolution v to (6.2.2), any compact set $K \subseteq E$ and $\varepsilon > 0$, there exist a compact set $\widehat{K} = \widehat{K}(K, \varepsilon, ||u||, ||v||)$ and a constant $C = C(u, v, K, h_1, h_2, \lambda)$ such that we have

$$\sup_{x\in K} u(x) - v(x) \leq \varepsilon C + \sup_{x\in \widehat{K}} h_1(x) - h_2(x).$$

Observe that the strict comparison principle implies the comparison principle. Indeed, by the strict comparison principle, for all $x_0 \in E$ and $\varepsilon > 0$, there exists a constant C, independent of ε, and a compact set $\widehat{K} \subseteq E$ such that

$$u(x_0) - v(x_0) \leq \varepsilon C + \sup_{x\in \widehat{K}} h_1(x) - h_2(x) \leq \varepsilon C + \sup_{x\in E} h_1(x) - h_2(x).$$

Letting $\varepsilon \downarrow 0$, we find that $u(x_0) - v(x_0) \leq \sup_{x\in E} h_1(x) - h_2(x)$. Taking the supremum over all $x_0 \in E$, the comparison princple follows.

6.2.4 *Notions for our framework*

One of the main innovations of this chapter is the use of a new approach to prove the comparison principle based on the notion of couplings of operators. In the following we give the main definitions underlying our new framework.

Definition 6.2.3 (Coupling)**.** Let $H \subseteq C(E) \times C(E)$ and $\widehat{H} \subseteq C(E^2) \times C(E^2)$ be linear on their respective domains. We say $\widehat{H}$ is a *coupling of* H if $\mathcal{D}(H) \oplus \mathcal{D}(H) \subseteq \mathcal{D}(\widehat{H})$ and, for any $f_1, f_2 \in \mathcal{D}(H)$, we have

$$\widehat{H}\,(f_1 \oplus f_2) = Hf_1 + Hf_2.$$

Definition 6.2.4 (Controlled growth)**.** Let $\widehat{H} \subseteq C(E^2) \times C(E^2)$. We say that $\widehat{H}$ has *controlled growth* if, for any $\alpha > 1$ and $z, z' \in E$, we have $\frac{\alpha}{2} d^2_{z,z'} \in \mathcal{D}(\widehat{H})$. In addition, for any compact set $K \subseteq E$, there exists a modulus of continuity $\omega_K : [0,\infty) \to [0,\infty)$ and $x, x', y, y' \in K$ such that

$$\widehat{H}\left(\frac{\alpha}{2} d^2_{x-y,\, x'-y'}\right)(x,x') \leq \omega_{\widehat{H},K}\Big(\alpha\left(d(x,y) + d(y,y') + d(y',x')\right)^2 + \left(d(x,y) + d(y,y') + d(y',x')\right)\Big).$$

Definition 6.2.5 (Controlled growth coupling)**.** Let $H \subseteq C(E) \times C(E)$ and $\widehat{H} \subseteq C(E^2) \times C(E^2)$ be linear on their respective domains. We say $\widehat{H}$ is a *controlled growth coupling of* H if the following properties are satisfied:

(a) $\widehat{H}$ satisfies the *maximum principle*, cf. Definition 2.2.11.
(b) $\widehat{H}$ is a *coupling* of H, cf. Definition 6.2.3.
(c) $\widehat{H}$ has *controlled growth*, cf. Definition 6.2.4.

We will split our Hamiltonian into a stochastic part that we can couple in the sense of the above definitions, and a deterministic part that we require to be a *convex semi-monotone* operator. Here we give the precise definitions.

Definition 6.2.6 (Local first-order operator)**.** We say that $H \subseteq C(E) \times C(E)$ is a *local first-order* operator if there exists a continuous map $\mathcal{B} : E \times \mathbb{R}^q \to \mathbb{R}$ such that, for any $f \in \mathcal{D}$, we have $Hf(x) = \mathcal{B}(x, \nabla f(x))$.

Definition 6.2.7 (Local semi-monotonicity)**.** Let $H \subseteq C(E) \times C(E)$ be local first-order for some $\mathcal{B}$, cf. Definition 6.2.6. We say that H is *locally semi-monotone* if, for any compact sets $K \subseteq E$, there exists a modulus of continuity $\omega_{\mathcal{B},K} : [0,\infty) \to [0,\infty)$ such that, for all $x, y \in K$ and $\alpha > 1$,

$$\mathcal{B}(x, \alpha(x-x')) - \mathcal{B}(y, \alpha(x-x')) \leq \omega_{\mathcal{B},K}\left(\alpha d^2(x,x') + d(x,x')\right).$$

Definition 6.2.8 (Convex semi-monotone operator)**.** We say that $H \subset C(E) \times C(E)$ is a *convex semi-monotone operator* if the following properties are satisfied:

(a) $H \subseteq C(E) \times C(E)$ is *locally semi-monotone* for some $\mathcal{B}$, cf. Definition 6.2.7.
(b) For all $x \in E$, the map $p \mapsto \mathcal{B}(x,p)$ is *convex*.

Finally, to work with Hamilton–Jacobi-Isaacs equations we need the following condition to be satisfied by our Hamiltonian.

Definition 6.2.9 (Isaacs' condition)**.** Let Θ_1 and Θ_2 be two compact, metric spaces. We say that a collection $\{H_{\theta_1,\theta_2}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2} \subseteq C(E) \times C(E)$ satisfies *Isaacs' condition* if, for all $f \in \bigcap_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2} \mathcal{D}(H_{\theta_1,\theta_2})$,

$$\sup_{\theta_1 \in \Theta_1} \inf_{\theta_2 \in \Theta_2} \left\{H_{\theta_1,\theta_2} f(x)\right\} = \inf_{\theta_2 \in \Theta_2} \sup_{\theta_1 \in \Theta_1} \left\{H_{\theta_1,\theta_2} f(x)\right\}.$$

Following typical comparison principle proofs, we will be perturbing the optimization problem

$$\sup_{x \in E} u(x) - v(x),$$

using a variant of the doubling of variables procedure to ensure that we can use the properties of sub- and supersolutions. Our perturbations consist of two components:

- We need a Lyapunov-type function, which ensures that we can work on compact sets, see Definition 6.2.10.
- We perform a variant of the Jensen perturbation to construct optimizers in which we can differentiate twice, see Definition 6.2.11.

Definition 6.2.10. We call $V : E \to [0, \infty)$ a *containment function* if

(a) $\inf_{y \in E} V(y) = 0$,
(b) V is semi-concave with semi-concavity constant κ_V,
(c) for every $c \geq 0$ the set $\{y \mid V(y) \leq c\}$ is compact.

Typically, the containment function is

$$V(x) = \log\left(1 + \frac{1}{2}x^2\right).$$

The next definition aims to produce optimizers for which we have twice differentiability via Jensen's Lemma, cf. Lemma A.3 of [CIL92]. The variant used here creates a unique, global optimizer from a local one using ξ and then shifts it slightly with ζ.The two sets of perturbations are based on the prototypical examples of lines, i.e.,

$$\zeta_{z,p}(x) = \langle p, x - z \rangle,$$

and parabolas, i.e.,

$$\xi_z(x) = \frac{1}{2} d^2(x, z),$$

both centered at some $z \in E$. We give them as pair to capture the idea that quadratic growth dominates linear growth, cf. Definition 6.2.11 (d) below, which is used in our variant of Jensen's Lemma in the Appendix, cf. Proposition 6.7.1.

Definition 6.2.11. We call collections of maps $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q} \subset C(E)$ and $\{\xi_z\}_{z \in E} \subset C^1(E)$ $\zeta_{z,p} : E \to \mathbb{R}$ and $\xi_z : E \to \mathbb{R}$ sets of *first* and *second order point penalizations*, respectively, if there exist constants $R > 0$ and $\kappa_\xi > 0$ such that for all $z \in E$:

(a) $\zeta_{z,p}$ is linear in terms of p around z:

$$\zeta_{z,p}(y) = \langle p, y - z \rangle$$

if $y \in B_R(z)$.

(b) The map ξ_z is semi-concave with constant κ_ξ.
(c) The map ξ_z is a penalization away from z:

$$\xi_z(z) = 0, \qquad \xi_z(y) > 0, \qquad \text{if } y \neq z.$$

(d) We have

$$\inf_{|p|\leq 1} \inf_{y \notin B_R(z)} \xi_z(y) + \zeta_{z,p}(y) > 0.$$

For any given $z_0, z_1 \in E$ and $p \in \mathbb{R}^q$, we consider the maps

$$\begin{aligned} \Xi^0(y) = \Xi^0_{z_0,p}(y) &:= \xi_{z_0}(y) + \zeta_{z_0,p}(y), \\ \Xi(y) = \Xi_{z_0,p,z_1}(y) &:= \xi_{z_0}(y) + \zeta_{z_0,p}(y) + \xi_{z_1}(y). \end{aligned}$$

6.3 ASSUMPTIONS AND MAIN RESULT

In this section, we present our main result, Theorem 6.3.1, and outline the fundamental assumptions underlying our analysis.

Theorem 6.3.1 states the strict comparison principle for operators in Hamilton–Jacobi-Isaacs (HJI) form, satisfying Isaacs' condition. Comparison principles for Hamilton–Jacobi (HJ) and Hamilton–Jacobi-Bellman (HJB) equations readily follow.

Heuristically, we assume that the base operator $\mathbb{H}$ can be split into a stochastic part $\mathbb{A}$, which we can couple, a semi-monotone deterministic term $\mathbb{B}$, and a cost functional $\mathcal{I}$, and that the action of the operator on the containment function V is controlled.

Theorem 6.3.1 (Strict comparison principle). *Let $\mathbb{H} \subseteq C(E) \times C(E)$ be given by*

$$\mathbb{H}f(x) = \sup_{\theta_1 \in \Theta_1} \inf_{\theta_2 \in \Theta_2} \left\{ \mathbb{A}_{\theta_1,\theta_2} f(x) + \mathbb{B}_{\theta_1,\theta_2} f(x) - \mathcal{I}(x, \theta_1, \theta_2) \right\}$$

with Θ_1 and Θ_2 compact, metric spaces and $\mathcal{I} : E \times \Theta_1 \times \Theta_2 \to (-\infty, \infty]$ a cost functional. Furthermore, consider a containment function V and penalization functions $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, $\{\zeta_z\}_{z \in E}$. Let $\mathbb{H}$ satisfy the technical Assumptions 6.3.4 and 6.3.5 below and assume that

(a) The collection of operators $\{\mathbb{A}_{\theta_1,\theta_2} + \mathbb{B}_{\theta_1,\theta_2} - \mathcal{I}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2}$ satisfies Isaacs' condition, cf. Definition 6.2.9.

(b) For all $\theta_1 \in \Theta_1, \theta_2 \in \Theta_2$, $\mathbb{A}_{\theta_1,\theta_2}$ is linear on its domain and admits a controlled growth coupling $\widehat{\mathbb{A}}_{\theta_1,\theta_2}$ as in Definition 6.2.5 with a modulus uniform in θ_1 and θ_2.

(c) For all $\theta_1 \in \Theta_1, \theta_2 \in \Theta_2$, $\mathbb{B}_{\theta_1,\theta_2}$ is a convex semi-monotone operator as in Definition 6.2.8 with a modulus uniform in θ_1 and θ_2.

(d) The cost functional $\mathcal{I}$ is lower semi-continuous in (x, θ_1, θ_2), upper semi-continuous in θ_2 for fixed (x, θ_1), and admits a modulus of continuity in x uniformly in (θ_1, θ_2).

(e) V is a Lyapunov function *for $\mathbb{H}$: $V \in \mathcal{D}(\mathbb{H})$ and*

$$c_V := \sup_{x \in E} \sup_{\theta_1 \in \Theta_1} \sup_{\theta_2 \in \Theta_2} \left\{ (\mathbb{A}_{\theta_1,\theta_2} + \mathbb{B}_{\theta_1,\theta_2}) V(x) - \mathcal{I}(x, \theta_1, \theta_2) \right\} < \infty. \tag{6.3.1}$$

Let $H := \{(f, g) \in \mathbb{H} \mid f \in C_b(E)\}$ be the restriction of $\mathbb{H}$ to $C_b(E)$ and consider

$$f - \lambda H f = h \tag{6.3.2}$$

for $\lambda > 0$ and $h \in C_b(E)$. Let u and v be a sub- and supersolution to (6.3.2) with h_1 and h_2 instead of h, respectively. Then for any compact set $K \subseteq E$ and $\varepsilon \in (0,1)$, we have

$$\sup_{x \in K} u(x) - v(x) \leq \varepsilon C_\varepsilon + \sup_{x \in \widehat{K}_\varepsilon} h_1(x) - h_2(x), \tag{6.3.3}$$

where $\widehat{K}_\varepsilon := \widehat{K}_\varepsilon(K,u,v)$ and $C_\varepsilon := C_\varepsilon(K,u,v,h_1,h_2)$ are given by

$$\widehat{K}_\varepsilon := \left\{ z \in E \,\middle|\, V(z) \leq \frac{||u|| + ||v||}{\varepsilon} + \lceil V \rceil_K \right\},$$

$$C_\varepsilon := \frac{2}{1-\varepsilon^2}\left(\lceil V \rceil_K + \lambda c_V\right) + \frac{1}{1-\varepsilon}||h_1|| + \frac{1}{1-\varepsilon}||h_2|| - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K .$$

In particular, the strict comparison principle holds for (6.3.2).

The proof of the above theorem is carried out in Sections 6.5 and 6.6. The next result follows by restricting the choice on Θ_1 and Θ_2.

Corollary 6.3.2.

(a) The strict comparison principle for HJB equations follows from Theorem 6.3.1 by taking Θ_2 to be a singleton.

(b) The strict comparison principle for HJ equations follows from Theorem 6.3.1 by taking both Θ_1 and Θ_2 to be singletons.

Remark 6.3.3. If H is the generator of a Markov process, the comparison principle implies uniqueness of the martingale problem using Theorem 3.7 in [CK15]. We also refer to [SV79; EK86] for details on the martingale problem.

6.3.1 *Regularity and compatibility assumptions*

In this section, we state the technical assumptions necessary for the proof the main theorem.

As we have a choice for the domain of our operator and only need functions with compact sub- and superlevel sets, we need to ensure that the domains of the restrictions are regular enough to perform our analysis. In particular, the action of the operator on test functions and their combinations with perturbations needs to be well-defined. Furthermore, we require that the domains are large enough to allow for approximations in the sense of Definition 6.2.1.

Assumption 6.3.4 (Regularity of $\mathbb{H}$)**.** Let $\mathbb{H} \subseteq C(E) \times C(E)$ be an operator with the following three restrictions

$$H := \{(f,g) \in \mathbb{H} \,|\, f \in C_b(E)\},$$

$$H_+ := \{(f,g) \in \mathbb{H} \,|\, f \in C_+(E)\},$$

$$H_- := \{(f,g) \in \mathbb{H} \,|\, f \in C_-(E)\},$$

satisfying

(a) $\mathbb{H}$ satisfies the maximum principle,
(b) $\mathcal{D}(H)$ is linear and $C_c^\infty(E) \subseteq \mathcal{D}(H) \subseteq C_b(E)$,
(c) $\mathcal{D}(H)$ is upward sequentially dense in $\mathcal{D}(H_+)$, as in Definition 6.2.1,
(d) $\mathcal{D}(H)$ is downward sequentially dense in $\mathcal{D}(H_-)$ as in Definition 6.2.1,
(e) $\mathcal{D}(H_+)$ is convex,
(f) for any $f \in \mathcal{D}(H)$ and $g \in \mathcal{D}(H_+)$ and $\delta \in (0,1)$ we have

$$(1-\delta)f + \delta g \in \mathcal{D}(H_+), \qquad (1+\delta)f - \delta g \in \mathcal{D}(H_-).$$

In our main theorem, Theorem 6.3.1, we assume that the Hamiltonian $\mathbb{H}$ has an Isaacs-type structure

$$\mathbb{H}f(x) = \sup_{\theta_1 \in \Theta_1} \inf_{\theta_2 \in \Theta_2} \left\{ \mathbb{A}_{\theta_1,\theta_2} f(x) + \mathbb{B}_{\theta_1,\theta_2} f(x) - \mathcal{I}(x, \theta_1, \theta_2) \right\}.$$

To ensure it is well-behaved, we need that the collections of operators $\{\mathbb{A}_{\theta_1,\theta_2}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2}$ and $\{\mathbb{B}_{\theta_1,\theta_2}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2}$ themselves are well-behaved as a functions of (θ_1, θ_2). We additionally assume that these collections behave well on the families of penalization functions introduced in Section 6.2.4.

Assumption 6.3.5 (Compatibility of $\mathbb{A}_{\theta_1,\theta_2}$ and $\mathbb{B}_{\theta_1,\theta_2}$)**.** Let Θ_1 and Θ_2 be compact, metric spaces. For $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, let $\mathbb{A}_{\theta_1,\theta_2}, \mathbb{B}_{\theta_1,\theta_2} \subseteq C(E) \times C(E)$. Consider a containment function V as in Definition 6.2.10 and penalization functions $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$ and $\{\zeta_z\}_{z \in E}$ as in Definition 6.2.11.

(a) Let the collection $\{\mathbb{A}_{\theta_1,\theta_2}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2}$ be *compatible* with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z \in E}$, i.e.,
  (1) we have

$$V \circ s_z \in \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2}), \qquad \Xi_{z_0,p,z_1} \circ s_z \in \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2})$$

  for any $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$ and $z \in \overline{B_1(0)}$,
  (2) the maps

$$(\theta_1, \theta_2, x, z_0, p, z_1, z) \mapsto \mathbb{A}_{\theta_1,\theta_2} \left(\Xi_{z_0,p,z_1} \circ s_z\right)(x),$$
$$(\theta_1, \theta_2, x, z) \mapsto \mathbb{A}_{\theta_1,\theta_2} \left(V \circ s_z\right)(x)$$

  are continuous,
  (3) the map

$$(\theta_1, \theta_2) \mapsto \mathbb{A}_{\theta_1,\theta_2} f(x)$$

  is continuous for any $x \in E$ and $f \in \bigcap_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2} \mathcal{D}(\mathbb{A}_{\theta_1,\theta_2})$.

(b) Let the collection $\{\mathbb{B}_{\theta_1,\theta_2}\}_{\theta_1 \in \Theta_1, \theta_2 \in \Theta_2}$ be *compatible* with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z \in E}$, i.e.,
  (1) we have

$$V \circ s_z \in \mathcal{D}(\mathbb{B}_{\theta_1,\theta_2}), \qquad \Xi_{z_0,p,z_1} \circ s_z \in \mathcal{D}(\mathbb{B}_{\theta_1,\theta_2})$$

  for any $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$ and $z \in \overline{B_1(0)}$,

(2) the maps

$$(\theta_1, \theta_2, x, z_0, p, z_1) \mapsto \mathbb{B}_{\theta_1,\theta_2}\Xi_{z_0,p,z_1}(x),$$
$$(\theta_1, \theta_2, x) \mapsto \mathbb{B}_{\theta_1,\theta_2}V(x)$$

are continuous,

(3) the map

$$(\theta_1, \theta_2) \mapsto \mathbb{B}_{\theta_1,\theta_2}f(x)$$

is continuous for any $x \in E$ and $f \in \bigcap_{\theta_1\in\Theta_1,\theta_2\in\Theta_2} \mathcal{D}(\mathbb{B}_{\theta_1,\theta_2})$.

6.4 APPLICATION TO PARTIAL INTEGRO-DIFFERENTIAL OPERATORS

In this section, we discuss the application of our framework to partial integro-differential operators of the type

$$\begin{aligned}\mathbb{H}f(x) &= \langle b(x), \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x)D^2f(x)\right) \\ &\quad + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\langle \mathbf{z}, \nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}) + \mathcal{H}(\nabla f(x)), \quad (6.4.1)\end{aligned}$$

which, for simplicity, we consider without Bellman or Isaacs structure. More precisely, we split $\mathbb{H}$ into $\mathbb{A} + \mathbb{B}$ with

$$\mathbb{A} = \frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x)D^2f(x)\right) + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\langle \mathbf{z}, \nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}),$$

and

$$\mathbb{B} = \langle b(x), \nabla f(x)\rangle + \mathcal{H}(\nabla f(x))$$

and specify conditions under which

- we can find a Lyapunov function V and construct a coupling that has controlled growth for $\mathbb{A}$,
- we can find a Lyapunov function V and establish local semi-monotonicity for $\mathbb{B}$,
- we can verify that $\mathbb{A}$ and $\mathbb{B}$ are compatible.

For the verification of compatibility, we need to choose V and families $\{\zeta_{z,p}\}_{z\in E,p\in\mathbb{R}^q}$ and $\{\xi_z\}_{z\in E}$. We introduce two related families: The first family is suitable for local operators; the second is constructed from the first by suitable cut-off procedures, thus making them suitable for integral operators.

We point out that the second family is suitable for local operators as well. However, this comes at the cost of minor complexity in their construction.

Definition 6.4.1 (Example containment function and penalizations)**.** Consider the containment function

$$V(x) = \log\left(1 + \frac{1}{2}x^2\right)$$

and the following two collections of penalization functions:

COLLECTION 1 The base penalizations are

$$\zeta_{z,p}(x) = \langle p, x - z \rangle ,$$
$$\xi_z(x) = \frac{1}{2} d^2(x,z).$$

COLLECTION 2 Let $R'' > R' > R$ with R as in Definition 6.2.11. Let $\bar{\ell} : [0,\infty) \to [0,\infty)$ be a smooth function satisfying $\bar{l}(r) = 1$ for $r < R'$ and $\bar{l}(r) = 0$ for $x > R''$. Let

$$\bar{\xi}_z(x) = (1 - \bar{\ell}(d(x,z)))(R'' + 1)^2 + \bar{\ell}(d(x,z))\frac{1}{2} d^2(x,z),$$
$$\bar{\zeta}_{p,z}(x) = \bar{\ell}(d(x,z)) \langle p, x - z \rangle .$$

As all considered functions are smooth, part (1) of the compatibility assumptions for $\mathbb{A}$ and $\mathbb{B}$, cf. Assumptions 6.3.5 (a) and (b), hold for every part of $\mathbb{H}$ except the integral term immediately.

To simplify the verification of our conditions, we have the following two observations.

Remark 6.4.2. Let $\mathbb{A}_1, \mathbb{A}_2 \subseteq C(E) \times C(E)$ be linear on their respective domains and compatible with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z\in E}$ and with associated controlled growth couplings $\widehat{\mathbb{A}}_1, \widehat{\mathbb{A}}_2 \subseteq C(E^2) \times C(E^2)$. Then, the operator $\mathbb{A} := \mathbb{A}_1 + \mathbb{A}_2$ is linear on its domain and compatible with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z\in E}$ and with associated controlled growth coupling $\widehat{\mathbb{A}} := \widehat{\mathbb{A}}_1 + \widehat{\mathbb{A}}_2$.

Remark 6.4.3. Let $\mathbb{B}_1, \mathbb{B}_2 \subseteq C(E) \times C(E)$ be compatible with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z\in E}$ and convex semi-monotone operators. Then $\mathbb{B} := \mathbb{B}_1 + \mathbb{B}_2$ is compatible with V, $\{\zeta_{z,p}\}_{z \in E, p \in \mathbb{R}^q}$, and $\{\zeta_z\}_{z\in E}$ and convex semi-monotone operator.

The rest of this section is organized as follows:

- In Section 6.4.1, we consider drift terms and convex first-order Hamiltonians;
- In Section 6.4.2, we consider diffusion operators;
- In Section 6.4.3, we consider integral operators.

6.4.1 *Deterministic Example: Drift terms and convex first-order Hamiltonians*

In this section, we consider the deterministic part of the operator (6.4.1).

Proposition 6.4.4. *Suppose that* $\mathbb{B}$ *is given by*

$$\mathbb{B} f(x) = \langle b(x), \nabla f(x) \rangle + \mathcal{H}(\nabla f(x))$$

with the drift term $x \mapsto b(x)$ *locally, one-sided Lipschitz with constant* $L_{b,K}$ *and* $||b(x)|| \leq \frac{c_b}{2}(1 + ||x||)$ *for some constant* $c_b > 0$*, and* $p \mapsto \mathcal{H}(p)$ *continuous and convex.*
Then, $\mathbb{B}$ *is compatible with both collections of Definition 6.4.1, cf. Assumption 6.3.5 (b), and convex semi-monotone. Furthermore,* $V = \log(1 + \frac{x^2}{2})$ *is a Lyapunov function:*

$$\sup_{x\in E} \mathbb{B} V(x) < \infty.$$

Proof. **Convex semi-monotonicity:**

Clearly, $\mathbb{B}$ is locally first-order with $\mathbb{B}f(x) = \langle b(x), \nabla f(x)\rangle + \mathcal{H}(\nabla f(x)) = \mathcal{B}(x, \nabla f(x))$. Additionally, for any compact set $K \subseteq E$ $\alpha > 0$, and $x, x' \in K$, we have

$$\begin{aligned}\mathcal{B}(x, \alpha(x-x')) - \mathcal{B}(y, \alpha(x-x')) &= \left\langle b(x), \alpha(x-x')\right\rangle + \mathcal{H}(\alpha(x-x')) \\ &\quad - \left\langle b(x'), \alpha(x-x')\right\rangle - \mathcal{H}(\alpha(x-x')) \\ &= \left\langle b(x) - b(x'), \alpha(x-x')\right\rangle \\ &\quad + \mathcal{H}(\alpha(x-x')) - \mathcal{H}(\alpha(x-x')) \\ &\leq \alpha L_{b,K} d^2(x,x'),\end{aligned}$$

establishing semi-monotonicity. As convexity of $p \mapsto \mathcal{B}(x,p)$ is immediate, we conclude that $\mathbb{B}$ is convex semi-monotone.

Lyapunov control: Using that $V(x) = \log\left(1 + \frac{x^2}{2}\right)$, $\nabla V(x) = \frac{2x}{2+|x|^2}$ is bounded as a function of x, b has linear growth, and that $\mathcal{H}$ is continuous, we find that

$$\sup_{x\in E} \mathbb{B}V(x) = \sup_{x\in E} \left\langle b(x), \frac{2x}{2+|x|^2}\right\rangle + \mathcal{H}\left(\frac{2x}{2+|x|^2}\right) < \infty.$$

Compatibility: We show the compatibility of $\mathbb{B}$, cf. Assumption 6.3.5 (b), by evaluation of the perturbation and containment function in the operator.

Using $\xi_z(x) = \frac{1}{2}d^2(x,z)$ and $\zeta_{z,p}(x) = \langle p, x-z\rangle$, we find for $z_0, z_1, z \in E$ and $p \in B_1(0)$

$$\begin{aligned}\mathbb{B}(\Xi_{z_0,p,z_1} \circ s_z)(x) &= \langle b(x), (x-z-z_0) + p + (x-z-z_1)\rangle \\ &\quad + \mathcal{H}\left((x-z-z_0) + p + (x-z-z_1)\right),\end{aligned}$$

which is continuous in (x, z_0, p, z_1, z) as b and $\mathcal{H}$ are continuous.

For $V(x) = \log\left(1 + \frac{1}{2}x^2\right)$ and $z \in E$, we find

$$\mathbb{B}(V \circ s_z)(x) = \left\langle b(x), \frac{2(x-z)}{2+|x-z|^2}\right\rangle + \mathcal{H}\left(\frac{2(x-z)}{2+|x-z|^2}\right),$$

which is continuous in (x,z) as b and $\mathcal{H}$ are continuous. Thus, $\mathbb{B}$ is compatible. □

6.4.2 *Stochastic Example: Diffusion operators*

In this section, we focus on diffusion operators of the form

$$\mathbb{A}f(x) = \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2 f(x)\right),$$

where $\Sigma(x)$ is a positive semi-definite matrix for each fixed $x \in E$.

Our main goal is to construct a controlled growth coupling for the operator $\mathbb{A}$. To illustrate the idea behind our approach, consider the simpler case of the Laplacian operator

$$\mathbb{A}_0 f(x) = \frac{1}{2}\operatorname{Tr}(D^2 f(x)),$$

which is the infinitesimal generator of Brownian motion. The well-known synchronous coupling of two Brownian motions started from x and x', respectively, is given by

$$(X(t), X'(t)) = (x + B(t), x' + B(t))$$

with $B(t)$ a standard Brownian motion, having generator

$$\widehat{\mathbb{A}}_0 g(x, x') = \frac{1}{2} \left(\partial_x + \partial_{x'}\right)^2 g(x, x'),$$

which satisfies $\widehat{\mathbb{A}}_0 d^2 = 0$. Aiming to generalize this, we rewrite

$$\widehat{\mathbb{A}}_0 g(x, x') = \mathrm{Tr}\left(CC^T D^2 g(x, x')\right) \quad \text{with} \quad C = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix}.$$

In general we obtain the following result.

Proposition 6.4.5. *Suppose that* $\mathbb{A}$ *is given by*

$$\mathbb{A} f(x) = \frac{1}{2} \mathrm{Tr}\left(\Sigma(x)\Sigma^T(x) D^2 f(x)\right)$$

with $\Sigma(x)$ *positive semi-definite for all* $x \in E$*,* $x \mapsto \Sigma(x)$ *locally Lipschitz with constant* $L_{\Sigma,K}$ *and* $||b(x)|| \leq \frac{c_\Sigma}{2}(1 + ||x||)$ *for some constant* $c_\Sigma > 0$*. Consider*

$$\widehat{\mathbb{A}} f(x, y) := \mathrm{Tr}\left(\widehat{\Sigma}^2(x, x') D^2 f(x, x')\right),$$

where

$$\widehat{\Sigma}^2(x, y) := \begin{pmatrix} \Sigma(x)\Sigma^T(x) & \Sigma(x')\Sigma^T(x) \\ \Sigma(x)\Sigma^T(x') & \Sigma(x')\Sigma^T(x') \end{pmatrix}.$$

Then, $\mathbb{A}$ *is compatible, cf. Assumption 6.3.5 (a), linear on its domain, and admitting the controlled growth coupling* $\widehat{\mathbb{A}}$*. Furthermore,* $V = \log(1 + \frac{x^2}{2})$ *is a Lyapunov function:*

$$\sup_x \mathbb{A} V(x) < \infty.$$

For the proof we make use of the following auxiliary lemma.

Lemma 6.4.6. *For each* $x \in E$*, let* $B(x)$ *be a positive semi-definite matrix and consider*

$$\mathbb{A} f(x) = \frac{1}{2} \mathrm{Tr}\left(B(x) D^2 f(x)\right).$$

For any $x, x' \in E$*, let* $\widehat{B}(x, x')$ *be a positive semi-definite matrix having block-structure*

$$\widehat{B}(x, x') = \begin{pmatrix} B(x) & B(x, x') \\ B(x, x')^T & B(x') \end{pmatrix}.$$

Define

$$\widehat{\mathbb{A}} f(x, x') := \frac{1}{2} \mathrm{Tr}\left(\widehat{B}(x, x') D^2 f(x, x')\right).$$

Then, $\widehat{\mathbb{A}}$ *is a coupling of* $\mathbb{A}$*.*

Proof.

$$\begin{aligned}\widehat{\mathbb{A}}(f_1 \oplus f_2)(x,y) &= \frac{1}{2}\operatorname{Tr}\left(\widehat{B}(x,y)D^2(f_1 \oplus f_2)(x,y)\right)\\ &= \frac{1}{2}\operatorname{Tr}\left(B(x)D^2 f(x)\right) + \frac{1}{2}\operatorname{Tr}\left(B(y)D^2 f(y)\right)\\ &= \mathbb{A}f_1 + \mathbb{A}f_2\end{aligned}$$

and it satisfies the maximum principle. □

Proof of Proposition 6.4.5. **Controlled growth coupling:** By Lemma 6.4.6, $\widehat{\mathbb{A}}$ is a coupling for $\mathbb{A}$. We thus verify that $\widehat{\mathbb{A}}$ has controlled growth. Consider $\alpha > 1$, $K \subseteq E$ a compact set, and $x, x', y, y' \in K$. Then,

$$\begin{aligned}\widehat{\mathbb{A}}\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x') &= \frac{1}{2}\operatorname{Tr}\left(\widehat{\Sigma}^2(x,x')D^2\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x')\right)\\ &= \frac{1}{2}\operatorname{Tr}\left(\widehat{\Sigma}^2(x,x')\left(\alpha\begin{pmatrix}\mathbb{1} & -\mathbb{1}\\ -\mathbb{1} & \mathbb{1}\end{pmatrix}\right)(x,x')\right)\\ &= \frac{\alpha}{2}\operatorname{Tr}((\Sigma^T(x)-\Sigma^T(y))(\Sigma(x)-\Sigma(y)))\\ &\le \alpha L^2_{\Sigma,K}d^2(x,x'),\end{aligned}$$

establishing controlled growth.

Lyapunov control: Using $V(x) = \log(1+\frac{x^2}{2})$ and the fact that Σ has linear growth, we find that

$$\sup_{x\in E}\mathbb{A}V(x) = \sup_{x\in E}\frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2V(x)\right) < \infty. \tag{6.4.2}$$

Compatibility: Using $\xi_z(x) = \frac{1}{2}d^2(x,z)$, $\zeta_{z,p}(x) = \langle p, x-z\rangle$ and $V(x) = \log(1+\frac{x^2}{2})$, we find for $z_0, z_1, z \in E$ and $p \in B_1(0)$

$$\begin{aligned}\mathbb{A}(\Xi\circ s_z)(x) &= 2\operatorname{Tr}(\Sigma(x)\Sigma^T(x)),\\ \mathbb{A}(V\circ s_z)(x) &= \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\Sigma^T(x)D^2(V\circ s_z)(x)\right),\end{aligned}$$

which, by an analogous calculation as in equation (6.4.2), is continuous in (x, z_0, p, z_1, z) and (x,z). Consequently, $\mathbb{A}$ is compatible. □

6.4.3 *Stochastic Example: Integral operators*

In this section, we cover examples of spatially inhomogeneous Lévy processes that have generators of the type

$$\mathbb{A}f(x) = \int\left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\langle\mathbf{z},\nabla f(x)\rangle\right]\mu_x(\mathrm{d}\mathbf{z}), \tag{6.4.3}$$

where $\chi_{B_1(0)}(\mathbf{z}) = l(|\mathbf{z}|)$ for some smooth non-decreasing function l satisfying $l = 1$ on a neighborhood of 0 and $l(r) = 0$ for $r \geq 1$.

We next specify the space from which we can take our jump measures μ_x. For this, we need to control the mass close to 0 as for large values of $\mathbf{z}$. The following function controls both:

$$W(\mathbf{z}) := \chi_{B_1(0)}(\mathbf{z})|\mathbf{z}|^2 + (1 - \chi_{B_1(0)}(\mathbf{z})) \log\left(1 + |\mathbf{z}|^2\right).$$

We take the family of jump measures $\{\mu_x\}_{x \in E}$ from the set of equivalence classes

$$\mathcal{M}_W(\mathbb{R}^q) := \mathcal{M}(\mathbb{R}^q)/\sim$$

with

$$\mathcal{M}(\mathbb{R}^q) := \left\{ \mu \in M(\mathbb{R}^q) \,\middle|\, \int W(\mathbf{z})\, \mu(\mathrm{d}\mathbf{z}) < \infty \right\},$$

where $M(\mathbb{R}^q)$ is the set of all Borel measures on $\mathbb{R}^q$ and where

$$\mu \sim \nu \qquad \text{if and only if} \qquad \mu|_{\mathbb{R}^q \setminus \{0\}} = \nu|_{\mathbb{R}^q \setminus \{0\}}.$$

We topologize the set $\mathcal{M}_W(\mathbb{R}^q)$ by the weak topology σ_W induced by the pairings

$$\mu \mapsto \int g(\mathbf{z})\mu(\mathrm{d}\mathbf{z}) \qquad \forall\, g \in C_W, \tag{6.4.4}$$

where

$$C_W := \left\{ g \in C(\mathbb{R}^q) \,\middle|\, g(0) = 0, \text{ and } \sup_{\mathbf{z} \neq 0} \frac{|g(\mathbf{z})|}{W(\mathbf{z})} < \infty. \right\}.$$

Below, we construct controlled growth couplings for operators of the type (6.4.3). To clarify the concepts, we consider the example of an uncompensated process, i.e., having an operator of the type

$$\mathbb{A} f(x) = \int f(x + \mathbf{z}) - f(x)\mu_x(\mathrm{d}\mathbf{z}).$$

Couplings for this type of operator are of the form

$$\widehat{\mathbb{A}} f(x, x') = \int \left[f(x + \mathbf{z}_1, y + \mathbf{z}_2) - f(x, x') \right] \pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathbf{z}_2), \tag{6.4.5}$$

where $\pi_{x,x'}$ couples μ_x and $\mu_{x'}$. In the following example, we illustrate the need of being able to couple jumps synchronously.

Example 6.4.7 (Random Walk). *Consider the simple random walk on* $\mathbb{R}$ *making jumps of size* 1*, i.e* $\mu_x = \mu = \delta_{-1} + \delta_1$ *leading to the operator*

$$\mathbb{A}f(x) = [f(x-1) + f(x+1) - 2f(x)] .$$

Well known couplings include walks with simultaneous jumps but independent directions, fully independent jumps, and synchronous jumps. The corresponding generators are given as in (6.4.5) *with jump measures*

$$\begin{aligned}
\pi^1 &:= \mu \otimes \mu, \\
\pi^2 &:= \delta_{(-1,0)} + \delta_{(1,0)} + \delta_{(0,-1)} + \delta_{(0,1)}, \\
\pi^3 &:= \delta_{(-1,-1)} + \delta_{(1,1)},
\end{aligned}$$

respectively. This leads to the operators

$$\begin{aligned}
\widehat{\mathbb{A}}^1 f(x,x') &= f(x-1,x'-1) + f(x-1,x'+1) \\
&\qquad + f(x+1,x'-1) + f(x+1,x'+1) - 4f(x,x'), \\
\widehat{\mathbb{A}}^2 f(x,x') &= f(x-1,x') + f(x+1,x') - 2f(x,x') \\
&\qquad + f(x,x'-1) + f(x,x'+1) - 2f(x,x'), \\
\widehat{\mathbb{A}}^3 f(x,x') &= f(x-1,x'-1) + f(x+1,x'+1) - 2f(x,x').
\end{aligned}$$

Only for the final example, we see that $\widehat{\mathbb{A}}^3 d^2 \leq 0$*, pointing at the necessity of the alignment of jumps.*

Note that the third coupling above has different total mass, and we thus work outside the realm of the typical notion of couplings of probability measures. A second feature of coupling jump measures, not present in the example above, is that we can make one process jump, whereas the other does not.

We formalize this in the following definition.

Definition 6.4.8. Let $\mu, \nu \in \mathcal{M}_W(\mathbb{R}^q)$. We say that $\pi \in M(\mathbb{R}^q \times \mathbb{R}^q)$ is an *extended coupling of* μ *and* ν, if

$$\begin{aligned}
\pi\left((A \setminus \{0\}) \times \mathbb{R}^q\right) &= \mu(A \setminus \{0\}) \qquad \forall\, A \in \mathcal{B}(\mathbb{R}^q), \\
\pi\left(\mathbb{R}^q \times (B \setminus \{0\})\right) &= \nu(B \setminus \{0\}) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}^q).
\end{aligned}$$

Remark 6.4.9. A variant of this coupling was introduced in [FG10]. There mass can be moved to the boundary of a domain. In our context, this boundary is the point 0.

Definition 6.4.10. Let $x \mapsto \mu_x$ be a map from E into $\mathcal{M}(\mathbb{R}^q)$. Let $(x,x') \mapsto \pi_{x,x'}$ be a map from E^2 into $M(\mathbb{R}^q \times \mathbb{R}^q)$.

(a) We say that $(x,x') \mapsto \pi_{x,x'}$ is an extended coupling of $x \mapsto \mu_x$, if for all $x, x' \in E$, we have that $\pi_{x,x'}$ is an extended coupling of μ_x and μ'_x.

(b) We say that $(x, x') \mapsto \pi_{x,x'}$ is locally Lipschitz, if, for any compact set $K \subseteq E$, there exits a constant $L_{\pi,K}$ such that, for $x, x' \in K$, we have

$$\int d^2(\mathbf{z}_1, \mathbf{z}_2)\pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathrm{d}\mathbf{z}_2) \leq L_{\pi,K} d^2(x, x').$$

Remark 6.4.11. Note that conditions (12), (34), and (35) in [BI08] for μ and j correspond to our choice of $\mathcal{M}_W(\mathbb{R}^q)$ and locally Lipschitz extended coupling $\pi_{x,x'}$.

Remark 6.4.12. Let $\eta : \mathbb{R} \to \mathbb{R}$ be any locally Lipschitz map with local Lipschitz constants $L_{\eta,K}$. Set $\mu_x := \delta_{\eta(x)} \mathbb{1}_{\eta(x)\neq 0}(x)$ and $\pi_{x,x'} = \delta_{(\eta(x),\eta(x'))}$. Then, $(x, x') \mapsto \pi_{x,x'}$ is a locally Lipschitz coupling of $x \mapsto \mu_x$ with $L_{\pi,K} = L_{\eta,K}$.

The main proposition of this subsection below aims to show that integral operators of the form (6.4.3) can be treated analogous to the other examples above. We work with the second collection of penalization functions, cf. Definition 6.4.1, to avoid integrability issues.

Proposition 6.4.13. *Consider*

$$\mathbb{A}f(x) = \int \Big[f(x + \mathbf{z}) - f(x) - \chi_{B_1(0)} \langle \mathbf{z}, \nabla f(x)\rangle\Big] \mu_x(\mathrm{d}\mathbf{z}).$$

Suppose there exists a σ_W-continuous map $x \mapsto \mu_x$ in $\mathcal{M}_W(\mathbb{R}^q)$, cf. (6.4.4), and that there exists a locally Lipschitz extended coupling $(x, x') \mapsto \pi_{x,x'}$ of $x \mapsto \mu_x$ with Lipschitz constant $L_{\pi,K}$ and, for $\widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) := \chi_{B_1(0)}(\mathbf{z}_1)\chi_{B_1(0)}(\mathbf{z}_2)$, set

$$\widehat{\mathbb{A}}g(x, x') := \int \Big[g(x + \mathbf{z}_1, x' + \mathbf{z}_2) - g(x, x') \\ - \widehat{\chi}(\mathbf{z}_1, \mathbf{z}_2) \Big\langle (\mathbf{z}_1, \mathbf{z}_2)^T, \nabla g(x, x')\Big\rangle \Big]\pi_{x,x'}(\mathrm{d}\mathbf{z}_1, \mathrm{d}\mathbf{z}_2).$$

Assume furthermore that

$$\sup_{x\in E} \int \log\left(1 + \frac{\frac{1}{2}|\mathbf{z}|^2 + \langle x, \mathbf{z}\rangle}{1 + \frac{1}{2}|x|^2}\right) \mu_x(\mathrm{d}\mathbf{z}) < \infty.$$

Then, $\mathbb{A}$ is compatible, cf. Definition 6.3.5 (a), and linear on its domain admitting the controlled growth coupling $\widehat{\mathbb{A}}$. Furthermore, $V = \log(1 + \frac{x^2}{2})$ is a Lyapunov function:

$$\sup_{x\in E} \mathbb{A}V(x) < \infty.$$

Remark 6.4.14. Corresponding to Remark 6.3.3, we refer to [Bas88, Corollary 2.3] for a uniqueness result for a Lévy process martingale problem.

The proof of Proposition 6.4.13 is based on the following two auxiliary lemmas. In the first, we obtain bounds on the integrand of our operator acting on the Lyapunov function V. In the second, we compute the integrand of our Lévy type operator acting on the shifted squared metric. We prove these two lemmas following the proof of Proposition 6.4.13.

Lemma 6.4.15. *Fix $x, z \in E$.*

(a) For $\mathbf{z} \in \mathbb{R}^q$, we have

$$-\log\left(1 + \frac{1}{2}|x-z|^2\right) \le V \circ s_z(x+\mathbf{z}) - V \circ s_z(x)$$
$$\le \log\left(1 + \frac{\frac{1}{2}|\mathbf{z}|^2 + \langle x-z, \mathbf{z}\rangle}{1 + \frac{1}{2}(x-z)^2}\right)$$
$$\le \log\left(1 + |\mathbf{z}|^2\right).$$

(b) For $\mathbf{z} \in B_1(0)$, we have

$$|V \circ s_z(x+\mathbf{z}) - V \circ s_z(x) - \langle \mathbf{z}, \nabla(V \circ s_z)(x)\rangle| \le \frac{1}{2}|\mathbf{z}|^2.$$

Lemma 6.4.16. *We have*

$$\frac{1}{2}d^2_{x-y,x'-y'}(x+\mathbf{z}_1, x'+\mathbf{z}_2) - \frac{1}{2}d^2_{x-y,x'-y'}(x,x')$$
$$- \widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\left\langle \begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix}, \nabla\left(\frac{1}{2}d^2_{x-y,x'-y'}\right)(x,x')\right\rangle$$
$$\le \left(1 - \frac{1}{2}\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\right) d^2(\mathbf{z}_1,\mathbf{z}_2) + (1-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2))\frac{1}{2}d^2(y,y').$$

Proof of Proposition 6.4.13. **Controlled growth coupling**: As $(x,x') \mapsto \pi_{x,x'}$ is a locally Lipschitz extended coupling of $x \mapsto \mu_x$, cf. Definition 6.4.10, we have that $\widehat{\mathbb{A}}$ is a coupling. Thus, we need to verify the controlled growth property of $\widehat{\mathbb{A}}$.

Let $x, x', y, y' \in K$ for $K \subseteq E$ a compact set. Using Lemma 6.4.16, we then have

$$\widehat{\mathbb{A}}\left(\frac{\alpha}{2}d^2_{x-y,x'-y'}\right)(x,x') \le \frac{\alpha}{2}\int\left(1 - \frac{1}{2}\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\right)d^2(\mathbf{z}_1,\mathbf{z}_2)\pi_{x,x'}(\mathrm{d}\mathbf{z}_1,\mathrm{d}\mathbf{z}_2)$$
$$+ \frac{\alpha}{2}\int(1-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2))\frac{1}{2}d^2(y,y')\pi_{x,x'}(\mathrm{d}\mathbf{z}_1,\mathrm{d}\mathbf{z}_2)$$
$$\le \frac{\alpha}{2}L_{\pi,K}d^2(x,x') + \frac{\alpha}{4}c'_\pi d^2(y,y'),$$

where the second inequality is due to the local Lipschitz property of the map $(x,x') \mapsto \pi_{x,x'}$ and $c'_\pi > 0$ exists since, for every $x, x' \in E$, $\pi_{x,x'} \in M(\mathbb{R}^q \times \mathbb{R}^q)$. As such, $\mathbb{A}$ admits the controlled growth coupling $\widehat{\mathbb{A}}$.

Lyapunov control: Using Lemma 6.4.15, we find

$$\sup_{x\in E}\mathbb{A}V(x) \le \sup_{x\in E}\int(1-\chi_{B_1(0)})\log(1+|\mathbf{z}|^2) + \chi_{B_1(0)}|\mathbf{z}|^2\mu_x(\mathrm{d}\mathbf{z}) < \infty.$$

Compatibility:

We start by establishing the continuity of $(x,z) \mapsto \mathbb{A}(V \circ s_z)(x)$. Let (x_n, z_n) converge to (x,z). We aim to apply Lemma 6.7.7 with $\mathcal{X} = \mathbb{R}^q \setminus \{0\}$, $\nu_n = \mu_{x_n}$, and

$$\phi_n(\mathbf{z}) := V \circ s_{z_n}(x_n + \mathbf{z}) - V \circ s_{z_n}(x_n) - \chi_{B_1(0)} \langle \mathbf{z}, \nabla(V \circ s_{z_n})(x_n) \rangle ,$$
$$\phi_\infty(\mathbf{z}) := V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) - \chi_{B_1(0)} \langle \mathbf{z}, \nabla(V \circ s_z)(x) \rangle .$$

As ϕ_n is continuous, it remains to show that $\sup_{n \in \mathbb{N}} \sup_{\mathbf{z} \neq 0} \frac{|\phi_n(\mathbf{z})|}{W(\mathbf{z})} < \infty$. By Lemma 6.4.15, we can estimate

$$|\phi_n(\mathbf{z})| \leq \chi_{B_1(0)} \frac{1}{2} |\mathbf{z}|^2 + (1 - \chi_{B_1(0)}) \max \left\{ -\log\left(1 + \frac{1}{2}\left|x_n - z_n\right|^2\right), \log\left(1 + |\mathbf{z}|^2\right) \right\} .$$

Since (x_n, z_n) is convergent, hence bounded, we obtain the desired estimate. Continuity of $(x,z) \mapsto \mathbb{A}(V \circ s_z)(x)$ now follows by Lemma 6.7.7.

Using the particular form of Ξ_{z_0,p,z_1}, cf. Definition 6.4.1, one readily verifies that the map $(x, z_0, p, z_1, z) \mapsto \mathbb{A}\left(\Xi_{z_0,p,z_1} \circ s_z\right)(x)$ is continuous with an analogous argumentation. □

Proof of Lemma 6.4.15. Let $y = x - z$, then we can write

$$\begin{aligned} &V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) \\ &\qquad = \log\left(1 + \frac{1}{2}(y + \mathbf{z})^2\right) - \log\left(1 + \frac{1}{2}|y|^2\right) = \log\left(1 + \frac{\frac{1}{2}|\mathbf{z}|^2 + \langle y, \mathbf{z} \rangle}{1 + \frac{1}{2}|y|^2}\right). \end{aligned}$$

Applying Young's inequality to $\langle y, \mathbf{z} \rangle$ leads to the upper bound

$$V \circ s_z(x+\mathbf{z}) - V \circ s_z(x) \leq \log\left(1 + \frac{|\mathbf{z}|^2 + \frac{1}{2}|y|^2}{1 + \frac{1}{2}|y|^2}\right) = \log\left(2 + \frac{|\mathbf{z}|^2 - 1}{1 + \frac{1}{2}|y|^2}\right) \leq \log\left(1 + |\mathbf{z}|^2\right).$$

Using that the first term is positive, we obtain the lower bound

$$V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) \geq \log\left(1 - \frac{\frac{1}{2}|y|^2}{1 + \frac{1}{2}|y|^2}\right) = -\log\left(1 + \frac{1}{2}|y|^2\right).$$

This establishes (a). For the proof of (b), we apply Taylor's Theorem to obtain

$$\begin{aligned} |V \circ s_z(x + \mathbf{z}) - V \circ s_z(x) - \langle \nabla(V \circ s_z)(x), \mathbf{z} \rangle| &\leq \frac{1}{2}|\mathbf{z}|^2 \sup_{\mathbf{z} \in B_1(0)} \sup_{i,j} \left|\nabla^2_{i,j} V(y + \mathbf{z})\right| \\ &\leq \frac{1}{2}|\mathbf{z}|^2, \end{aligned}$$

which follows by a direct inspection of

$$\nabla^2_{i,j} V(x) = \frac{2\delta_{i,j}\left(1 + \frac{1}{2}|x|^2\right) - 2x_i x_j}{(1 + \frac{1}{2}|x|^2)^2}.$$

□

Proof of Lemma 6.4.16. Evaluating the shift maps, calculating the gradient of the squared Euclidean distance, and expanding the squares leads to

$$\begin{aligned}
&\frac{1}{2}d^2_{x-y,x'-y'}(x+\mathbf{z}_1,x'+\mathbf{z}_2)-\frac{1}{2}d^2_{x-y,x'-y'}(x,x')\\
&\qquad-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\left\langle\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix},\nabla\left(\frac{1}{2}d^2_{x-y,x'-y'}\right)(x,x')\right\rangle\\
&=\frac{1}{2}d^2(y+\mathbf{z}_1,y'+\mathbf{z}_2)-\frac{1}{2}d^2(y,y')-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\left\langle y-y',\mathbf{z}_1-\mathbf{z}_2\right\rangle\\
&=\frac{1}{2}d^2(\mathbf{z}_1,\mathbf{z}_2)+\left\langle y-y',\mathbf{z}_1-\mathbf{z}_2\right\rangle-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\left\langle y-y',\mathbf{z}_1-\mathbf{z}_2\right\rangle\\
&\leq\left(1-\frac{1}{2}\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2)\right)d^2(\mathbf{z}_1,\mathbf{z}_2)+(1-\widehat{\chi}(\mathbf{z}_1,\mathbf{z}_2))\frac{1}{2}d^2(y,y'),
\end{aligned}$$

where in the second equality we use properties of the Euclidean distance d and the final line is due to Young's inequality. □

6.5 CONSTRUCTION OF TEST FUNCTIONS

In classical proofs of comparison principles, the approach to estimate $\sup u-v$ for a subsolution u and supersolution v is variable doubling or quadruplication, cf. [BC97, Theorem 3.1] or [CIL92, introduction of Section 3]: For $\alpha>1$

$$\sup_{x\in E}u(x)-v(x)\leq\sup_{x,x'\in E}u(x)-v(x')-\frac{\alpha}{2}d^2(x,x'). \tag{6.5.1}$$

Letting $\alpha\to\infty$, forces optimizing points, if they exist, of the right-hand side together. In addition, by varying either of the two components, one obtains basic test functions in terms of $\frac{\alpha}{2}d^2$ for the use in the definition of the sub- and supersolution properties of u and v.

To ensure that optimizers in (6.5.1) exist, we will consider instead, for small $\varepsilon>0$, the following problem that includes the containment function V and upper bounds $\sup u-v$ up to a term of order ε:

$$\begin{aligned}
&\sup_{x\in E}\frac{1}{1-\varepsilon}u(x)-\frac{1}{1+\varepsilon}v(x)\\
&\quad\leq\sup_{x,x'\in E}\frac{1}{1-\varepsilon}u(x)-\frac{1}{1+\varepsilon}v(x')-\frac{\alpha}{2}d^2(x,x')-\frac{\varepsilon}{1-\varepsilon}V(x)-\frac{\varepsilon}{1+\varepsilon}V(x').
\end{aligned}\tag{6.5.2}$$

The particular form of the factors $1-\varepsilon$ and $1+\varepsilon$ is motivated by convexity based arguments, which will show up in the proofs of Proposition 6.6.3 and Theorem 6.3.1 below.

The procedure in (6.5.2) would be sufficient for a standard, first-order Hamilton–Jacobi equation. The test functions produced by this procedure, however, will not be sufficient to treat second-order or integral operators. This problem was considered in

[CIL92] and [BI08]. We will follow their approach by considering a quadruplication of variables, which we also phrase in terms of sup- and inf-convolutions. We then perform a Jensen-type perturbation.

As we aim to unify proofs for both integral and differential operators, we revisit the full proof and state our result in terms of test functions.

In Propositions 6.5.1 and 6.5.3 below, which can be considered to be an extended two-variable variant of the Crandall–Ishii construction [CIL92, Theorem 3.2], we start out by considering the optimization (6.5.2) in terms of the sup- and inf-convolution of u and v, respectively, effectively leading to a quadruplication problem, see (6.5.3) below.

We then perform the Jensen perturbation, see (6.5.4). The rest of the proposition deals with various properties of the optimizers in relation to u and v.

In Proposition 6.5.3, we carry out an additional layer of smoothing operations to obtain C^∞-test functions. Consequently, we can move away from the notion of solutions in terms of sub- and superjets, which is of paramount importance to effectively treat diffusive and jump-type processes in a common framework.

For readability, we express suprema and infima using $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$, respectively, as defined in Section 6.2.1.

Proposition 6.5.1 (Construction of optimizers)**.** *Let u be bounded and upper semi-continuous, v be bounded and lower semi-continuous, V be a containment function as in Definition 6.2.10, and $\{\zeta_{z,p}\}_{z\in E,p\in\mathbb{R}^q} \subset C(E)$ and $\{\xi_z\}_{z\in E} \subset C^1(E)$ be collections of functions as in Definition 6.2.11. Fix $\varepsilon \in (0,1)$ and $\varphi \in (0,1]$.*

Then, there exist compact sets $K_{\varepsilon,0} \subseteq K_\varepsilon \subseteq E$ and, for any $\alpha > 1$, three pairs of variables $(y_{\alpha,0}, y'_{\alpha,0})$, (y_α, y'_α), (x_α, x'_α) in E^2 and $p_\alpha, p'_\alpha \in B_{1/\alpha}(0)$ such that the following four sets of properties hold.

PROPERTIES OF $y_{\alpha,0}, y'_{\alpha,0}$

The variables $y_{\alpha,0}, y'_{\alpha,0}$ optimize $\lceil\Lambda_\alpha\rceil$, where

$$\Lambda_\alpha(y,y') := \frac{1}{1-\varepsilon}P^\alpha[u](y) - \frac{1}{1+\varepsilon}P_\alpha[v](y') - \frac{\alpha}{2}d^2(y,y') - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y') \quad (6.5.3)$$

and satisfy the following property

(a) $y_{\alpha,0}, y'_{\alpha,0} \in K_{\varepsilon,0}$.

PROPERTIES OF y_α, y'_α AND p_α, p'_α

The pair y_α, y'_α optimizes

$$\left\lceil \Lambda_\alpha - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0 - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0 \right\rceil \quad (6.5.4)$$

and uniquely optimizes

$$\left\lceil \Lambda_\alpha - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1 - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2 \right\rceil \quad (6.5.5)$$

where Λ_α *is as in* (6.5.3) *and*

$$\Xi_1^0(y) := \Xi^0_{y_{\alpha,0},p_\alpha}(y), \qquad \Xi_2^0(y') := \Xi^0_{y'_{\alpha,0},p'_\alpha}(y'),$$
$$\Xi_1(y) := \Xi_{y_{\alpha,0},p_\alpha,y_\alpha}(y), \qquad \Xi_2(y') := \Xi_{y'_{\alpha,0},p'_\alpha,y'_\alpha}(y')$$

as in Definition 6.2.11. *Moreover, the optimizers* y_α, y'_α *of* (6.5.4) *and* (6.5.5) *satisfy*

(b) *We have*

$$d(y_\alpha, y_{\alpha,0}) \leq \frac{1}{\alpha}, \qquad d(y'_\alpha, y'_{\alpha,0}) \leq \frac{1}{\alpha}.$$

(c) $P^\alpha[u]$ *and* $P_\alpha[v]$ *are twice differentiable in* y_α *and* y'_α*, respectively.*

PROPERTIES OF x_α, x'_α

The variables x_α, x'_α *optimize*

$$P^\alpha[u](y_\alpha) = u(x_\alpha) - \frac{\alpha}{2}d^2(x_\alpha, y_\alpha),$$
$$P_\alpha[v](y'_\alpha) = v(x'_\alpha) + \frac{\alpha}{2}d^2(x'_\alpha, y'_\alpha),$$

and satisfy

(d) x_α *and* x'_α *are the unique optimizers in the definition of* $P^\alpha[u](y_\alpha)$ *and* $P_\alpha[v](y'_\alpha)$*, respectively.*

(e) *We have that*

$$u(x_\alpha) - P^\alpha[u] \circ s_{x_\alpha - y_\alpha}(x_\alpha) = \lceil u - P^\alpha[u] \circ s_{x_\alpha - y_\alpha} \rceil,$$
$$v(x'_\alpha) - P_\alpha[v] \circ s_{x'_\alpha - y'_\alpha}(x'_\alpha) = \lfloor v - P_\alpha[v] \circ s_{x'_\alpha - y'_\alpha} \rfloor.$$

BEHAVIOUR AS $\alpha \to \infty$

(f) *We have* $\lim_{\alpha\to\infty} \alpha d^2(y_{\alpha,0}, y'_{\alpha,0}) = 0$.

(g) *We have*

$$\lim_{\alpha\to\infty} \alpha \left(d\left(x_\alpha, y_\alpha\right) + d\left(y_\alpha, y'_\alpha\right) + d\left(y'_\alpha, x'_\alpha\right)\right)^2 = 0.$$

(h) $x_\alpha, y_\alpha, y'_\alpha, x'_\alpha \in K_\varepsilon$.

In addition, the following estimate on $u - v$ *holds: For any compact set* $K \subseteq E$*, there is a compact set* $\widehat{K} = \widehat{K}(K, \varepsilon, u, v)$ *given by*

$$\widehat{K} := \left\{ z \in E \,\middle|\, V(z) \leq \frac{||u|| + ||v||}{\varepsilon} + \lceil V \rceil_K \right\},$$

such that

(i) *For any compact set* $K \subseteq E$*,*

$$\lceil u - v \rceil_K \leq \frac{1}{1-\varepsilon} u(x_\alpha) - \frac{1}{1+\varepsilon} v(x_\alpha) + \varepsilon \left(c_{\varepsilon,\varphi} + o(1)\right),$$

where

$$c_{\varepsilon,\varphi} := \frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V \rceil_K - \left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K ,$$

and $o(1)$ *is in terms of* $\alpha \to \infty$ *for fixed* ε *and* φ.

(j) *Any limit point of the sequence* $(x_\alpha, y_\alpha, y_{\alpha,0}, y'_{\alpha,0}, y'_\alpha, x'_\alpha)$ *as* $\alpha \to \infty$ *is of the form* (z,z,z,z,z,z) *with* $z \in \widehat{K}$.

Figure 9 visualizes the relation between the different optimizing points.

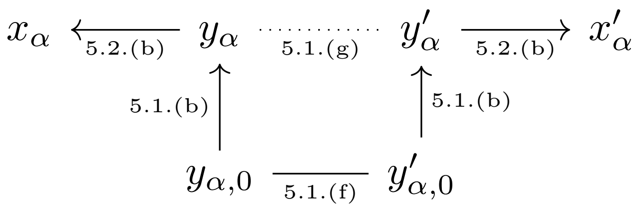

Figure 9: Relation between the optimizing points with a note which parts of the propositions give us distance control.

The proof of Proposition 6.5.1 uses various properties of sup- and inf-convolutions, which we gather in the next lemma. Its proof is relegated to Appendix 6.7.4.

Lemma 6.5.2. *Let* $u : E \to \mathbb{R}$ *be bounded and upper semi-continuous and* $v : E \to \mathbb{R}$ *be bounded and lower semi-continuous. For* $\alpha > 1$, *set*

$$P^\alpha[u](y) := \sup_{x \in E} \left\{ u(x) - \frac{\alpha}{2} d^2(x,y) \right\} = \left\lceil u - \frac{\alpha}{2} d^2(\cdot, y) \right\rceil, \tag{6.5.6}$$

$$P_\alpha[v](y) := \inf_{x \in E} \left\{ v(x) + \frac{\alpha}{2} d^2(x,y) \right\} = \left\lfloor u + \frac{\alpha}{2} d^2(\cdot, y) \right\rfloor. \tag{6.5.7}$$

Then,

(a) *we have* $||P^\alpha[u]|| \leq ||u||$ *and* $||P_\alpha[v]|| \leq ||v||$.

(b) *for any* $x, y \in E$ *such that*

$$P^\alpha[u](y) = u(x) - \frac{\alpha}{2} d^2(x,y),$$

we have $\frac{\alpha}{2} d^2(x,y) \leq u(x) - u(y)$. *Similarly, for any* $x, y \in E$ *with*

$$P_\alpha[v](y) = v(x) + \frac{\alpha}{2} d^2(x,y),$$

we have $\frac{\alpha}{2} d^2(x,y) \leq v(y) - v(x)$.

(c) $P^\alpha[u]$ *and* $-P_\alpha[v]$ *are decreasing in* α.

(d) $P^\alpha[u]$ *and* $-P_\alpha[v]$ *are semi-convex with semi-convexity constant* α. *As a consequence, both are locally Lipschitz continuous.*

(e) *if $P^\alpha[u]$ is differentiable at y_0, then there exists a unique optimizer x_0 in* (6.5.6) *such that*

$$P^\alpha[u](y_0) = u(x_0) - \frac{\alpha}{2} d^2(x_0, y_0)$$

and $DP^\alpha[u](y_0) = \alpha(x_0 - y_0)$. Similarly, if $P_\alpha[v]$ is differentiable at y_0, then there is a unique optimizer x_0 in (6.5.7) *such that*

$$P_\alpha[v](y_0) = v(x_0) + \frac{\alpha}{2} d^2(x_0, y_0)$$

and $DP_\alpha[v](y_0) = -\alpha(x_0 - y_0)$.

Proof of Proposition 6.5.1. **Proof of (a):** As u and v are bounded, by Lemma 6.5.2 (a), the same holds for $||P^\alpha[u]||$ and $||P_\alpha[v]||$. Using that V has compact sublevelsets, cf. Definition 6.2.10, the existence of optimizers $(y_{\alpha,0}, y'_{\alpha,0})$ for $\lceil \Lambda_\alpha \rceil$ follows.

The definition of Λ_α and the convolutions $P^\alpha[u]$ and $P_\alpha[v]$ imply that

$$\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}) \leq \frac{1}{1-\varepsilon}\lceil u \rceil - \frac{1}{1+\varepsilon}\lfloor v \rfloor - \lceil \Lambda_\alpha \rceil . \tag{6.5.8}$$

Comparing the optimizers for $\Lambda_\alpha(y, y')$ to, e.g., the suboptimial choice $(y, y') = (\hat{y}, \hat{y})$ satisfying $V(\hat{y}) = 0$, we find

$$\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}) \leq \frac{2}{1-\varepsilon}||u|| + \frac{2}{1+\varepsilon}||v|| .$$

From this estimate, we deduce that $(y_{\alpha,0}, y'_{\alpha,0}) \in K_{\varepsilon,0} \times K_{\varepsilon,0}$ with

$$K_{\varepsilon,0} := \left\{ y \in E \,\middle|\, V(y) \leq \varepsilon^{-1} C_\varepsilon(||u|| + ||v||) \right\}$$

for some constant $C_\varepsilon > 0$ satisfying $\lim_{\varepsilon \downarrow 0} C_\varepsilon = \frac{2}{1-\varphi}$, establishing (a).

Proof of (b) and (c): For the proof of these two statements, we first move from $\lceil \Lambda_\alpha \rceil$ to its perturbed version (6.5.4). To do so, we use Proposition 6.7.1. Note, that the function $(y, y') \mapsto \Lambda_\alpha(y, y')$ of (6.5.3) over which we optimize in $\lceil \Lambda_\alpha \rceil$ is semi-convex with semi-convexity constant

$$\kappa = \left(\frac{2}{1-\varepsilon^2} + \frac{1}{2} \right) \alpha + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\kappa_V > 1$$

for $\alpha > 1$. In addition, it is bounded from above and has optimizers $(y_{\alpha,0}, y'_{\alpha,0})$. We can thus apply Proposition 6.7.1 with

$$\eta = \frac{1}{\alpha}, \qquad \varepsilon_1 = \frac{\varepsilon}{1-\varepsilon}\varphi, \qquad \varepsilon_2 = \frac{\varepsilon}{1+\varepsilon}\varphi. \tag{6.5.9}$$

Consequently, it follows that there exist $p_\alpha, p'_\alpha \in B_{1/\alpha}(0)$ such that y_α, y'_α are optimizers of

$$\left\lceil \widehat{\Lambda}_\alpha \right\rceil = \widehat{\Lambda}_\alpha(y_\alpha, y'_\alpha),$$

where

$$\widehat{\Lambda}_\alpha(y,y') := \Lambda_\alpha(y,y') - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y) - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y') \tag{6.5.10}$$

with Ξ_1^0 and Ξ_2^0 as defined above. This establishes (6.5.4). An additional penalization around (y_α, y'_α) gives (6.5.5). A secondary outcome of Proposition 6.7.1 is that $\widehat{\Lambda}_\alpha$ is twice differentiable in the optimizing point (y_α, y'_α), establishing (c). Furthermore, the optimizers satisfy

$$d(y_\alpha, y_{\alpha,0}) < \eta, \qquad d(y'_\alpha, y'_{\alpha,0}) < \eta,$$

which, together with (6.5.9), yields

$$\max\left\{d(y_\alpha, y_{\alpha,0}), d(y'_\alpha, y_{\alpha,0})'\right\} \leq \frac{1}{\alpha},$$

establishing (b).

Proof of (d): This follows immediately from Lemma 6.5.2 (e).

Proof of (e): We only establish

$$u(x_\alpha) - P^\alpha[u]\circ s_{x_\alpha - y_\alpha}(x_\alpha) = \lceil u - P^\alpha[u]\circ s_{x_\alpha-y_\alpha}\rceil,$$

as the second equation follows similarly. Note that by definition of $P^\alpha[u]$, we have

$$P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x) \geq u(x) - \frac{\alpha}{2}d^2\big(x, s_{x_\alpha-y_\alpha}(x)\big).$$

On the other hand, by (d), we have

$$P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x_\alpha) = P^\alpha[u](y_\alpha) = u(x_\alpha) - \frac{\alpha}{2}d^2\left(x_\alpha, y_\alpha\right).$$

Combining the two statements yields, for any $x \in E$, that

$$\begin{aligned}
&u(x_\alpha) - P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x_\alpha)\\
&\quad = \frac{\alpha}{2}d^2\left(x_\alpha, y_\alpha\right) + P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x) - P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)\\
&\quad \geq u(x) - P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x) + \frac{\alpha}{2}\left(d^2\left(x_\alpha, y_\alpha\right) - d^2\left(x, s_{x_\alpha-y_\alpha}(x)\right)\right)\\
&\quad = u(x) - P^\alpha[u]\circ s_{x_\alpha-y_\alpha}(x)
\end{aligned}$$

as the shift map preserves distances. This establishes (e).

For the proof of the final five properties, we consider the limit $\alpha \to \infty$.

Proof of (f): Consider $\lceil\Lambda_\alpha\rceil$:

$$\begin{aligned}
\lceil\Lambda_\alpha\rceil = &\frac{1}{1-\varepsilon}P^\alpha[u](y_{\alpha,0}) - \frac{1}{1+\varepsilon}P_\alpha[v](y'_{\alpha,0}) - \frac{\alpha}{2}d^2(y_{\alpha,0}, y'_{\alpha,0})\\
&- \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0}).
\end{aligned}$$

Note, that $\lceil\Lambda_\alpha\rceil$ is decreasing in α, since $-\frac{\alpha}{2}d^2(y_{\alpha,0},y'_{\alpha,0})$, $P^\alpha[u]$, and $-P_\alpha[v]$ are decreasing in α by Lemma 6.5.2 (c). Note in addition that, by evaluating Λ_α in the particular choice $(y,y')=(\hat{y},\hat{y})$ as above, we have, by Lemma 6.5.2 (a), that

$$\lceil\Lambda_\alpha\rceil \geq \frac{1}{1-\varepsilon}P^\alpha[u](\hat{y}) - \frac{1}{1+\varepsilon}P_\alpha[v](\hat{y}) \geq \frac{1}{1-\varepsilon}\,||u|| - \frac{1}{1+\varepsilon}\,||v||\,,$$

which is lower bounded uniformly in α. It follows that the limit $\lim_{\alpha\to\infty}\sup\Lambda_\alpha$ exists. For any $\alpha>1$, we find

$$\begin{aligned}\lceil\Lambda_{\alpha/2}\rceil &\geq \frac{1}{1-\varepsilon}P^{\alpha/2}[u](y_{\alpha,0}) - \frac{1}{1+\varepsilon}P_{\alpha/2}[v](y'_{\alpha,0}) - \frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0})\\ &\quad - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0})\\ &\geq \frac{1}{1-\varepsilon}P^{\alpha}[u](y_{\alpha,0}) - \frac{1}{1+\varepsilon}P_{\alpha}[v](y'_{\alpha,0}) - \frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0})\\ &\quad - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_{\alpha,0}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_{\alpha,0})\\ &\geq \lceil\Lambda_\alpha\rceil + \frac{\alpha}{4}d^2(y_{\alpha,0},y'_{\alpha,0}),\end{aligned} \tag{6.5.11}$$

which implies that $\lim_{\alpha\to\infty}\alpha d^2(y_{\alpha,0},y'_{\alpha,0})=0$, as $\lceil\Lambda_\alpha\rceil$ and $\lceil\Lambda_{\alpha/2}\rceil$ converge to the same limit, establishing (f).

Proof of (g): We follow the same approach as in (6.5.11) but now expanding $P^\alpha[u](y_\alpha)$ and $P_\alpha[v](y'_\alpha)$ to obtain an optimization problem in terms of four variables.

$$\begin{aligned}\lceil\Lambda_{\alpha/2}\rceil &\geq \frac{1}{1-\varepsilon}P^{\alpha/2}[u](y_{\alpha}) - \frac{1}{1+\varepsilon}P_{\alpha/2}[v](y'_{\alpha}) - \frac{\alpha}{4}d^2(y_{\alpha},y'_{\alpha})\\ &\geq \frac{1}{1-\varepsilon}u(x_\alpha) - \frac{1}{1+\varepsilon}v(x'_\alpha) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha)\\ &\quad - \frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha) + d^2(y_\alpha,y'_\alpha) + \frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\ &= \left\lceil\widehat{\Lambda}_\alpha\right\rceil + \frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha) + d^2(y_\alpha,y'_\alpha) + \frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\ &\quad + \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha)\end{aligned}$$

by (6.5.10). It follows that

$$\begin{aligned}&\frac{\alpha}{4}\left(\frac{1}{1-\varepsilon}d^2(x_\alpha,y_\alpha) + d^2(y_\alpha,y'_\alpha) + \frac{1}{1+\varepsilon}d^2(y'_\alpha,x'_\alpha)\right)\\ &\qquad\qquad \leq \lceil\Lambda_{\alpha/2}\rceil - \left\lceil\widehat{\Lambda}_\alpha\right\rceil - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha) - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha).\end{aligned}$$

By (f), we obtain

$$\lim_{\alpha\to\infty}\lceil\Lambda_\alpha\rceil = \lim_{\alpha\to\infty}\left\lceil\widehat{\Lambda}_\alpha\right\rceil$$

and

$$\lim_{\alpha\to\infty} \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha) = 0.$$

Consequently, we have that

$$\lim_{\alpha\to\infty} \alpha\left(d^2(x_\alpha, y_\alpha) + d^2(y_\alpha, y'_\alpha) + d^2(y'_\alpha, x'_\alpha)\right) = 0.$$

From this, (g) follows using Young's inequality.

Proof of (h): (a), (b), (f), and (g) imply (h) by considering a bounded blow-up K_ε of $K_{\varepsilon,0}$.

Proof of (i): First, note that Corollary 6.7.2 and the definition of η in (6.5.9) yield

$$0 \leq -\frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha) - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha) \leq \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha} \tag{6.5.12}$$

and

$$\lceil\Lambda_\alpha\rceil \leq \left\lceil\widehat{\Lambda}_\alpha\right\rceil = \widehat{\Lambda}_\alpha(y_\alpha, y'_\alpha) \leq \lceil\Lambda_\alpha\rceil + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}. \tag{6.5.13}$$

Let $K \subseteq E$ be compact. We then obtain

$$\begin{aligned}
\lceil u - v\rceil_K &= \sup_{x\in K} u(x) - v(x) \\
&\leq \sup_{x\in K}\left\{u(x) - v(x) - \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left(V(x) - \lceil V\rceil_K\right)\right\} \\
&\leq \sup_{x\in K}\left\{\frac{1}{1-\varepsilon}u(x) - \frac{1}{1+\varepsilon}v(x) - \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)V(x)\right\} \\
&\qquad + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K \\
&\leq \sup_{x\in E}\left\{\frac{1}{1-\varepsilon}u(x) - \frac{1}{1+\varepsilon}v(x) - \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)V(x)\right\} \\
&\qquad + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K \\
&\leq \lceil\Lambda_\alpha\rceil + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \varepsilon\left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K.
\end{aligned} \tag{6.5.14}$$

Combining this estimate with the first inequality of (6.5.13), dropping non-positive terms, and then (6.5.12), leads to

$$\begin{aligned}
\lceil u - v\rceil_K &\leq \widehat{\Lambda}_\alpha(y_\alpha, y'_\alpha) \\
&\leq \frac{1}{1-\varepsilon}u(x_\alpha) - \frac{1}{1+\varepsilon}v(x'_\alpha) \\
&\qquad + \varepsilon\left(\varphi\frac{2}{1-\varepsilon^2}\frac{1}{\alpha} + \frac{2}{1-\varepsilon^2}(1-\varphi)\lceil V\rceil_K - \left\lfloor\frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v\right\rfloor_K\right),
\end{aligned}$$

which proves (i).

Proof of (j): We start by proving that any limiting point of

$$(x_\alpha, y_\alpha, y_{\alpha,0}, y'_{\alpha,0}, y'_\alpha, x'_\alpha)$$

as $\alpha \to \infty$ is of the form (z, z, z, z, z, z). We only prove $\lim_{\alpha\to\infty} d(x_\alpha, y_\alpha) = 0$, as the other limit follow analogously.

By (h), we find that, along subsequences, $(x_\alpha, y_\alpha) \to (x_0, y_0)$. Assume by contradiction that $x_0 \neq y_0$. Then, since αd^2 is increasing, we get that for all $\alpha_0 > 1$,

$$\liminf_{\alpha\to\infty} \alpha d^2(x_\alpha, y_\alpha) \geq \alpha_0 d^2(x_0, y_0).$$

We can conclude that $\alpha d^2(x_\alpha, y_\alpha) \to \infty$, contradicting (g).

We proceed to prove that any limiting point z lies in $\widehat{K}$. Similar to (6.5.8), but now also using (6.5.12) and the first inequality of (6.5.13), we find

$$\begin{aligned}
&\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha) \\
&\quad \leq \frac{1}{1-\varepsilon}\lceil u \rceil - \frac{1}{1+\varepsilon}\lfloor v \rfloor - \frac{\varepsilon}{1-\varepsilon}\varphi\Xi_1^0(y_\alpha) - \frac{\varepsilon}{1+\varepsilon}\varphi\Xi_2^0(y'_\alpha) - \left\lceil \widehat{\Lambda}_\alpha \right\rceil, \\
&\quad \leq \frac{1}{1-\varepsilon}\lceil u \rceil - \frac{1}{1+\varepsilon}\lfloor v \rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha} - \lceil \Lambda_\alpha \rceil,
\end{aligned}$$

Combining this with the upper bound on $-\lceil \Lambda_\alpha \rceil$ obtained from (6.5.14) leads to

$$\begin{aligned}
&\frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)V(y'_\alpha) \\
&\quad \leq \frac{1}{1-\varepsilon}\lceil u \rceil + \frac{1}{1+\varepsilon}\lfloor v \rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha} \\
&\qquad - \lceil u - v \rceil_K + \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\lceil V \rceil_K - \varepsilon\left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K.
\end{aligned}$$

This, in turn, yields

$$\begin{aligned}
&\frac{\varepsilon}{1-\varepsilon}(1-\varphi)\left(V(y_\alpha) - \lceil V \rceil_K\right) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\left(V(y'_\alpha) - \lceil V \rceil_K\right) \\
&\quad \leq \frac{1}{1-\varepsilon}\lceil u \rceil + \frac{1}{1+\varepsilon}\lfloor v \rfloor + \varphi\frac{2\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha} \\
&\qquad - \lceil u - v \rceil_K + \varepsilon\left\lfloor \frac{1}{1-\varepsilon}u - \frac{1}{1+\varepsilon}v \right\rfloor_K \\
&\quad \leq 2\left(||u|| + ||v||\right) + \varphi\frac{\varepsilon}{1-\varepsilon^2}\frac{1}{\alpha}.
\end{aligned}$$

The sequences $x_\alpha, y_\alpha, y'_\alpha, x'_\alpha$ have limit points $z \in K_\varepsilon$ as $\alpha \to \infty$ by (g) and (h). In combination with (b), we conclude that, for any such limiting point z,

$$\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left(V(z) - \lceil V \rceil_K\right) \leq 2\left(||u|| + ||v||\right),$$

establishing (j). □

The next proposition builds upon Proposition 6.5.1 to build a suitable collection of test functions for the use in the proof of the comparison principle. The sup- and inf-convolution $R^\alpha[u]$ and $R_\alpha[v]$ are not guaranteed to be smooth. However, they are twice differentiable in the relevant optimizing points.

Using the difference between Ξ_1^0 and Ξ_2^0 on one hand and Ξ_1 and Ξ_2 on the other, we are able to squeeze in a globally C^∞ function on the basis of Lemma 6.7.6, that can be used to replace $P^\alpha[u]$ and $P_\alpha[v]$. As an effect, we will approximate

$$\widehat{f}_\dagger \approx P^\alpha[u], \qquad f_\dagger \approx P^\alpha[u] \circ s_{x_\alpha - y_\alpha},$$
$$\widehat{f}_\ddagger \approx P_\alpha[v], \qquad f_\ddagger \approx P_\alpha[v] \circ s_{x'_\alpha - y'_\alpha},$$

which will be made rigorously in next proposition for fixed ε and α.

Proposition 6.5.3 (Test function construction). *Consider the setting of Proposition 6.5.1. Fix $\varepsilon \in (0,1)$, $\varphi \in (0,1]$, and $\alpha > 1$. Then, there are functions $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in C_c^\infty(E)$ such that*

$$f_1 = \widehat{f}_1 \circ s_{x_\alpha - y_\alpha}, \qquad f_2 = \widehat{f}_2 \circ s_{x'_\alpha - y'_\alpha}$$

and

$$\widehat{f}_\dagger := (1-\varepsilon)\widehat{f}_1 + \varepsilon(1-\varphi)V + \varepsilon\varphi\Xi_1, \qquad f_\dagger = \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha},$$
$$\widehat{f}_\ddagger := (1+\varepsilon)\widehat{f}_2 - \varepsilon(1-\varphi)V - \varepsilon\varphi\Xi_2, \qquad f_\ddagger = \widehat{f}_\ddagger \circ s_{x'_\alpha - y'_\alpha},$$

satisfying the following properties:

For $\widehat{f}_1, \widehat{f}_2$ and f_1, f_2, we have

(a) The pair (y_α, y'_α) is the unique optimizing pair of

$$\widehat{f}_1(y_\alpha) - \widehat{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha) = \left\lceil \widehat{f}_1 - \widehat{f}_2 - \frac{\alpha}{2} d^2 \right\rceil.$$

and the pair (x_α, x'_α) is the unique optimizing pair of

$$f_1(x_\alpha) - f_2(x'_\alpha) - \frac{\alpha}{2} d^2_{x_\alpha - y_\alpha,\, x'_\alpha - y'_\alpha}(x_\alpha, x'_\alpha) = \left\lceil f_1 - f_2 - \frac{\alpha}{2} d^2_{x_\alpha - y_\alpha,\, x'_\alpha - y'_\alpha} \right\rceil.$$

For $\widehat{f}_\dagger, \widehat{f}_\ddagger$, and $f_\dagger, f_\ddagger$ we have

(b) We have

$$P^\alpha[u](y) \leq \widehat{f}_\dagger(y),$$
$$P_\alpha[v](y') \geq \widehat{f}_\ddagger(y')$$

with equality in y_α and y'_α, respectively.

(c) We have that x_α, x'_α are the unique points such that

$$u(x_\alpha) - f_\dagger(x_\alpha) = \lceil u - f_\dagger \rceil,$$
$$v(x'_\alpha) - f_\ddagger(x'_\alpha) = \lceil v - f_\ddagger \rceil.$$

(d) *We have*

$$\begin{aligned} D\widehat{f}_\dagger(y_\alpha) &= Df_\dagger(x_\alpha) = \alpha(y_\alpha - x_\alpha), \\ D^2\widehat{f}_\dagger(y_\alpha) &= D^2 f_\dagger(x_\alpha), \\ D\widehat{f}_\ddagger(y'_\alpha) &= Df_\ddagger(x'_\alpha) = \alpha(x'_\alpha - y'_\alpha), \\ D^2\widehat{f}_\ddagger(y'_\alpha) &= D^2 f_\ddagger(x'_\alpha). \end{aligned}$$

As noted before the previous proposition, we aim to construct $\widehat{f}_\dagger \approx P^\alpha[u]$, but start out by first constructing $\widehat{f}_1 \in C_c^\infty(E)$, which, by re-arrangement, satisfies

$$\widehat{f}_1 \approx \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1(y)$$

and is constant outside of a compact set. As V has compact sublevel sets and other terms on the right-hand side are bounded from above, it suffices to first perform a smooth approximation and cut off the result. For the cut-off procedure, we use functions Ω_M^+ and Ω_M^-.

Definition 6.5.4 (Cut-off functions)**.** Let $M > 0$. We call a smooth increasing function $\Omega_M^+ : \mathbb{R} \to \mathbb{R}$ a *upper cut-off function at M*, if

$$\Omega_M^+(r) = \begin{cases} r & \text{if } r \leq M, \\ M+1 & \text{if } r \geq M+2. \end{cases}$$

We call Ω_M^- a *lower cut-off function* at M if $\Omega_M^-(r) = -\Omega_{-M}^+(-r)$.

Proof of Proposition 6.5.3. In this proof, we work in the context of Proposition 6.5.1 and will, correspondingly, follow its notation. We show the construction procedure for the test function f_1 used in the subsolution case only, as f_2 is constructed analogously. Denote

$$\begin{aligned} \Pi_1^0(y) &:= \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1^0(y), \\ \Pi_1(y) &:= \frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1(y). \end{aligned}$$

Note that we have $\Pi_1(y_\alpha) = \Pi_1^0(y_\alpha)$ and $\Pi_1(y) < \Pi_1^0(y)$ for all $y \in E \setminus \{y_\alpha\}$. By Lemma 6.7.6, we find a function $\mathfrak{f}_1 \in C^\infty(E)$ such that

$$\Pi_1(y) < \mathfrak{f}_1(y) < \Pi_1^0(y), \qquad y \neq y_\alpha.$$

The function $\mathfrak{f}_2$ is constructed analogously. By construction of $\mathfrak{f}_1$, $\mathfrak{f}_2$ and (6.5.4), (y_α, y'_α) is the unique optimizer of $\lceil \mathfrak{f}_1 - \mathfrak{f}_2 - \frac{\alpha}{2} d^2 \rceil$.

As our test functions need to be constant outside a compact set, we need to cut them off in an appropriate manner. However, we need to preserve their properties in the optimizer (y_α, y'_α). Taking these conditions into account, ensures that the cut-off procedure does not create new optimizers.

The above considerations lead to the cut-off procedure $\widehat{f_1} := \Omega^-_{M_1} \circ \mathfrak{f}_1$ and $\widehat{f_2} := \Omega^+_{M_2} \circ \mathfrak{f}_2$, with $\Omega^-_{M_1}$, $\Omega^+_{M_2}$ as in Definition 6.5.4 and the following choice of M_1 and M_2:
Pick $m_1, m_2 \in \mathbb{R}$ such that the level sets

$$\{y \in E \,|\, \mathfrak{f}_1(y) \geq m_1\}, \qquad \left\{y' \in E \,\middle|\, \mathfrak{f}_2(y') \leq m_1\right\}$$

are compact. Set

$$M_1 := \min\left\{m_1,\, \mathfrak{f}_1(y_\alpha) - \left(\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor\right) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)\right\},$$
$$M_2 := \max\left\{m_2,\, \mathfrak{f}_2(y'_\alpha) + \left(\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)\right) + \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)\right\}.$$

Using M_1 and M_2 as defined above, we find that (y_α, y'_α) is the unique optimizer of $\left\lceil \widehat{f_1} - \widehat{f_2} - \frac{\alpha}{2} d^2 \right\rceil$. To see this, denote

$$A_1 := \{y \in E \,|\, \mathfrak{f}_1(y) \geq M_1\} \quad \text{and} \quad A_2 := \left\{y' \in E \,\middle|\, \mathfrak{f}_2(y') \leq M_2\right\}.$$

Thus, for $i \in \{1, 2\}$, we find $\widehat{f_i} = \mathfrak{f}_i$ on A_i, whereas

$$\widehat{f_1}(y) < \mathfrak{f}_1(y_\alpha) - \left(\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor\right) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha),$$
$$\widehat{f_2}(y') > \mathfrak{f}_2(y'_\alpha) + \left(\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)\right) + \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)$$

if $y \notin A_1$ or $y' \notin A_2$, respectively.
As $\widehat{f_1} = \mathfrak{f}_1$ on A_1 and $\widehat{f_2} = \mathfrak{f}_2$ on A_2, it suffices to show that

$$\widehat{f_1}(y) - \widehat{f_2}(y') - \frac{\alpha}{2} d^2(y, y') < \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)$$

if $y \in A_1^c$ or $y \in A_2^c$. For the proof of this bound, we consider the following three separate cases.

Case $y \in A_1^c$ and $y' \in A_2$: We have

$$\begin{aligned}
\widehat{f_1}(y) - \widehat{f_2}(y') - \frac{\alpha}{2} d^2(y, y') &\leq \widehat{f_1}(y) - \widehat{f_2}(y') \\
&< \mathfrak{f}_1(y_\alpha) - \left(\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor\right) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha) - \mathfrak{f}_2(y') \\
&= \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha) - \left(\mathfrak{f}_2(y') - \lfloor \mathfrak{f}_2 \rfloor\right) \\
&\leq \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha).
\end{aligned}$$

Case $y \in A_1$ and $y' \in A_2^c$: Follows analogously to the case $y \in A_1^c$ and $y' \in A_2$.

Case $y \in A_1^c$ and $y' \in A_2^c$: We have

$$\begin{aligned}
\widehat{f}_1(y) - \widehat{f}_2(y') - \frac{\alpha}{2} d^2(y,y') \\
&\le \widehat{f}_1(y) - \widehat{f}_2(y') \\
&< \mathfrak{f}_1(y_\alpha) - \left(\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor\right) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha) \\
&\qquad - \left(\mathfrak{f}_2(y'_\alpha) + (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) + \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha)\right) \\
&\le \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - 2\frac{\alpha}{2} d^2(y_\alpha, y'_\alpha) - \left(\mathfrak{f}_2(y'_\alpha) - \lfloor \mathfrak{f}_2 \rfloor\right) - (\lceil \mathfrak{f}_1 \rceil - \mathfrak{f}_1(y_\alpha)) \\
&\le \mathfrak{f}_1(y_\alpha) - \mathfrak{f}_2(y'_\alpha) - \frac{\alpha}{2} d^2(y_\alpha, y'_\alpha).
\end{aligned}$$

We conclude that the pair (y_α, y'_α) is also the unique optimizer of $\left[\widehat{f}_1 - \widehat{f}_2 - \frac{\alpha}{2} d^2\right]$. Applying the shift maps $s_{x_\alpha - y_\alpha}$ and $s_{x'_\alpha - y'_\alpha}$, respectively, we find that (x_α, x'_α) uniquely optimize $\left[\mathfrak{f}_1 \circ s_{x_\alpha - y_\alpha} - \mathfrak{f}_2 \circ s_{x'_\alpha - y'_\alpha} - \frac{\alpha}{2} d^2_{x_\alpha - y_\alpha, x'_\alpha - y'_\alpha}\right]$. Additionally, as $M_1 \ge m_1$ and $M_2 \le m_2$, we have $\widehat{f}_1, \widehat{f}_2 \in C_c^\infty(E)$, establishing (a).

We next prove (b). As $r \le \Omega^-_{M_1}(r)$,

$$\frac{1}{1-\varepsilon} P^\alpha[u](y) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)V(y) - \frac{\varepsilon}{1-\varepsilon}\varphi \Xi_1(y) = \Pi_1(y) \le \Omega^-_{M_1} \circ \Pi_1(y) \le \widehat{f}_1(y),$$

which, after rearrangement of terms, implies (b).

We proceed with the proof of (c). By (b) and Proposition 6.5.1 (e),

$$\begin{aligned}
f_\dagger(x) - f_\dagger(x_\alpha) &= \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha}(x) - \widehat{f}_\dagger \circ s_{x_\alpha - y_\alpha}(x_\alpha) \\
&\ge \left(P^\alpha[u] \circ s_{x_\alpha - y_\alpha}\right)(x) - \left(P^\alpha[u] \circ s_{x_\alpha - y_\alpha}\right)(x_\alpha) \\
&\ge \left(u(x) - \frac{\alpha}{2} d^2(x, s_{x_\alpha - y_\alpha}(x)\right) - \left(u(x_\alpha) - \frac{\alpha}{2} d^2(x_\alpha, s_{x_\alpha - y_\alpha}(x_\alpha)\right) \\
&= u(x) - u(x_\alpha)
\end{aligned}$$

with equality uniquely realized at x_α, establishing (c).

We conclude with the proof of (d). First of all, note that the equality of first and second order derivatives for $f_\dagger$ and $\widehat{f}_\dagger$ as well as for $f_\ddagger$ and $\widehat{f}_\ddagger$ follows by the chain rule.

The expressions for $D\widehat{f}_\dagger(y_\alpha)$ and $D\widehat{f}_\ddagger(y'_\alpha)$ follow from (b) and Proposition 6.5.1 (c) and (d). □

6.6 PROOF OF THE STRICT COMPARISON PRINCIPLE

In this section, we prove Theorem 6.3.1. The proof is based on a variant of the variable quadruplication procedure on the basis of

$$\begin{aligned}
&\sup_{x \in E} \frac{1}{1-\varepsilon} u(x) - \frac{1}{1+\varepsilon} v(x) \\
&\leq \sup_{x,,y,y'x' \in E} \frac{1}{1-\varepsilon} u(x) - \frac{1}{1+\varepsilon} v(x') - \frac{\alpha}{2(1-\varepsilon)} d^2(x,y) - \frac{\alpha}{2} d^2(y,y') \\
&\qquad\qquad - \frac{\alpha}{2(1+\varepsilon)} d^2(y',x') - \frac{\varepsilon}{1+\varepsilon} V(x) - \frac{\varepsilon}{1+\varepsilon} V(x'),
\end{aligned}$$

which we have formalized in terms of test functions $f_\dagger, f_\ddagger$ in Propositions 6.5.1 and 6.5.3.

In a first step, we relate sub- and supersolutions for the Hamilton–Jacobi equation for H to those for H_+ and H_-: This will be carried out in Lemma 6.6.1. A second step is to show that $f_\dagger \in \mathcal{D}(H_+)$ and $f_\ddagger \in \mathcal{D}(H_-)$: This will be carried out in Lemma 6.6.2.

After establishing these technical points, we proceed to frame the comparison principle in terms of an estimate on

$$\frac{H_+ f_\dagger}{1-\varepsilon} - \frac{H_- f_\ddagger}{1+\varepsilon}. \tag{6.6.1}$$

This reduction will be carried out in Proposition 6.6.3, the statement of which is more involved than typically in the literature, but leads to the improved *strict* comparison principle. Its formulation and proof hinges on the use of V as a Lyapunov function.

The statements of Lemmas 6.6.1, 6.6.2, and Proposition 6.6.3 can be found in Section 6.6.1, their proofs in Section 6.6.2.

We finish in Section 6.6.3 by estimating (6.6.1) in two steps leading to our final result. We first establish in Lemma 6.6.4 that the pre-factors $(1-\varepsilon)^{-1}$ and $(1+\varepsilon)^{-1}$ work well with the combinations of functions that define $f_\dagger, f_\ddagger$ in Proposition 6.5.3. We conclude this section with the proof of Theorem 6.3.1, where we use this split, the coupling assumption on $\mathbb{A}$, the semi-monotonicity of $\mathbb{B}$, modulus of continuity control on $\mathcal{I}$ and, again, that V is a Lyapunov function to arrive at our final result.

6.6.1 *Comparison in terms of estimating the difference of Hamiltonians*

We start with connecting the notion of sub- and supersolutions for H to those for H_+ and H_-, respectively.

Lemma 6.6.1. *Let H and $\mathbb{H}$ satisfy Assumption 6.3.4. Then, for any $h \in C_b(E)$ and $\lambda > 0$, we have the following:*

(a) Any viscosity subsolution of $f - \lambda H f = h$ is also a viscosity subsolution of $f - \lambda H_+ f = h$.

(b) Any viscosity supersolution of $f - \lambda H f = h$ is also a viscosity supersolution of $f - \lambda H_- f = h$.

The proof follows in Section 6.6.2 below. In the next lemma we show that the test functions that we constructed in the previous section are in the domain of H_+ and H_-.

Lemma 6.6.2. *Let $\mathbb{H}$ be an operator satisfying Assumptions 6.3.4 and 6.3.5. Let $\widehat{f}_\dagger, f_\dagger$ and $\widehat{f}_\ddagger, f_\ddagger$ be as in Proposition 6.5.3. Then, $\widehat{f}_\dagger, f_\dagger \in \mathcal{D}(H_+)$ and $\widehat{f}_\ddagger, f_\ddagger \in \mathcal{D}(H_-)$.*

The proof of the lemma is outlined in Section 6.6.2 below. We next state our key proposition, which relates the strict comparison principle to an estimate on the difference of Hamiltonians.

Proposition 6.6.3. *Let $\mathbb{H} \subseteq C(E) \times C(E)$ satisfy Assumptions 6.3.4 and 6.3.5. Let $h_1, h_2 \in C_b(E)$, and $\lambda > 0$. Consider the equations*

$$f - \lambda H_+ f \leq h_1, \tag{6.6.2}$$
$$f - \lambda H_- f \geq h_2. \tag{6.6.3}$$

Let u and v by viscosity sub- and supersolutions to (6.6.2) and (6.6.3), respectively. For each $\varepsilon \in (0,1)$, $\varphi \in (0,1]$ and $\alpha > 1$, consider the construction of optimizers x_α, x'_α and test functions $f_\dagger, f_\ddagger$ as in Propositions 6.5.1 and 6.5.3.

Suppose there exists a map $\varepsilon \mapsto C^0_\varepsilon$, and for any $\varepsilon \in (0,1)$ a non-negative map $\varphi \mapsto C_{\varepsilon,\varphi}$ satisfying $\limsup_{\varepsilon \downarrow 0} C^0_\varepsilon < \infty$ and $\lim_{\varphi \downarrow 0} C_{\varepsilon,\varphi} = 0$ such that

$$\liminf_{\alpha \to \infty} \frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \leq \varepsilon \left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right). \tag{6.6.4}$$

Then, for any compact set $K \subseteq E$ and $\varepsilon \in (0,1)$,

$$\sup_{x \in K} u(x) - v(x) \leq \varepsilon C_\varepsilon + \sup_{x \in \widehat{K}} h_1(x) - h_2(x),$$

where $\widehat{K}_\varepsilon := \widehat{K}_\varepsilon(K,u,v)$ and $C_\varepsilon := C_\varepsilon(K,u,v,h_1,h_2)$ are given by

$$\widehat{K}_\varepsilon := \left\{ z \in E \,\middle|\, V(z) \leq \frac{||u|| + ||v||}{\varepsilon} + \lceil V \rceil_K \right\},$$
$$C_\varepsilon := \lambda C^0_\varepsilon + \frac{2}{1-\varepsilon^2} \lceil V \rceil_K + \frac{1}{1-\varepsilon} ||h_1|| + \frac{1}{1-\varepsilon} ||h_2|| - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K.$$

In particular, the strict comparison principle holds for (6.6.2) and (6.6.3).

6.6.2 *Proof of Lemmas 6.6.1, 6.6.2, and Proposition 6.6.3*

Proof of Lemma 6.6.1. We only prove the first statement, the second one follows analogously. Let u be a subsolution to $f - \lambda H f = h$ and let $(f,g) \in H_+$. Our claim thus follows if there exists x_0 satisfying

$$u(x_0) - f(x_0) = \lceil u - f \rceil, \tag{6.6.5}$$
$$u(x_0) - \lambda g(x_0) \leq h(x_0). \tag{6.6.6}$$

As u is upper semi-continous and bounded, and f has compact sublevel sets, the existence of x_0 satisfying (6.6.5) is immediate. We thus proceed with (6.6.6) using the sequential upward denseness of $\mathcal{D}(H)$ in $\mathcal{D}(H_+)$, cf. Assumption 6.3.4 (c). Set

$$a := f(x_0) + \lceil u \rceil - u(x_0), \qquad A := \{x \,|\, f(x) \leq a\}.$$

We can thus find $(f_a, g_a) \in H$ with f_a satisfying

$$\begin{cases} f_a(x) = f(x) & \text{if } x \in A, \\ a < f_a(x) \leq f(x) & \text{if } x \notin A. \end{cases}$$

We first establish that

$$u(x_0) - f_a(x_0) = \lceil u - f_a \rceil. \tag{6.6.7}$$

Using (6.6.5) and that $f = f_a$ on A, (6.6.7) follows by verifying that

$$u(x) - f_a(x) < u(x_0) - f(x_0), \qquad x \in A^c,$$

which follows from the definition of a:

$$\begin{aligned} u(x) - f(x) &< u(x) - a \\ &= u(x) - (f(x_0) + \lceil u \rceil - u(x_0)) \\ &= u(x_0) - f(x_0) - (\lceil u \rceil - u(x)) \\ &\leq u(x_0) - f(x_0). \end{aligned}$$

Thus, by (6.6.7), we can use the subsolution inequality for (f_a, g_a) in the point x_0. We obtain:

$$u(x_0) - \lambda g_a(x_0) \leq h(x_0). \tag{6.6.8}$$

Recalling that $f_a(x_0) = f(x_0)$ and $f_a \leq f$, we have

$$f_a(x_0) - f(x_0) = \lceil f_a - f \rceil.$$

Using the positive maximum principle for $\mathbb{H}$, cf. Assumption 6.3.4 (a), thus yields

$$g_a(x_0) \leq g(x_0). \tag{6.6.9}$$

Combining (6.6.8) and (6.6.9), leads to

$$u(x_0) - \lambda g(x_0) \leq u(x_0) - \lambda g_a(x_0) \leq h(x_0),$$

establishing (6.6.6) and consquently that u is a subsolution to $f - \lambda H_+ f = h$. □

Proof of Lemma 6.6.2. As $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in C_c^\infty(E)$, it follows by Assumption 6.3.4 (b) that $f_1, f_2, \widehat{f}_1, \widehat{f}_2 \in \mathcal{D}(H)$. By compatibility, cf. Assumption 6.3.5, we have $V \circ s_z, \Xi \circ s_z \in \mathcal{D}(\mathbb{H})$. By Assumption 6.3.4 (e) and the fact that V has compact sublevel sets, cf. Definition 6.2.10, we thus have $(1-\varphi)V \circ s_z + \varphi \Xi \circ s_z \in \mathcal{D}(H_+)$. Consequently, $\widehat{f}_\dagger, f_\dagger \in \mathcal{D}(H_+)$ and $\widehat{f}_\ddagger, f_\ddagger \in \mathcal{D}(H_-)$ by Assumption 6.3.4 (f). □

Proof of Proposition 6.6.3. Let u be a subsolution of $f - \lambda H_+ f = h_1$ and v a supersolution of $f - \lambda H_- f = h_2$. Consider the constructions in Propositions 6.5.1 and 6.5.3 for the subsolution u, supersolution v and $\varepsilon \in (0,1)$ and $\varphi \in (0,1]$.

By Lemma 6.6.2, we have $f_\dagger \in \mathcal{D}(H_+)$ and $f_\ddagger \in \mathcal{D}(H_-)$ and, by Proposition 6.5.3 (c), we find that (x_α, x'_α) are the unique optimizers in

$$\begin{aligned} u(x_\alpha) - f_\dagger(x_\alpha) &= \lceil u - f_\dagger \rceil, \\ v(x'_\alpha) - f_\ddagger(x'_\alpha) &= \lceil v - f_\ddagger \rceil, \end{aligned}$$

which, by the sub- and supersolution properties for H_+ and H_-, respectively, and Lemma 1.2.6, implies that

$$\begin{aligned} u(x_\alpha) - \lambda H_+ f_\dagger(x_\alpha) &\leq h_1(x_\alpha), \\ v(x'_\alpha) - \lambda H_- f_\ddagger(x'_\alpha) &\geq h_2(x'_\alpha). \end{aligned} \tag{6.6.10}$$

By Proposition 6.5.1 (i), we find

$$\lceil u - v \rceil_K \leq \frac{1}{1-\varepsilon} u(x_\alpha) - \frac{1}{1+\varepsilon} v(x'_\alpha) + \varepsilon \left(c_{\varepsilon,\varphi} + o(1) \right),$$

where

$$c_{\varepsilon,\varphi} := \frac{2}{1-\varepsilon^2}(1-\varphi) \lceil V \rceil_K - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K, \tag{6.6.11}$$

and $o(1)$ is in terms of $\alpha \to \infty$. Using (6.6.10), we estimate

$$\begin{aligned} \lceil u - v \rceil_K &\leq \frac{1}{1-\varepsilon} u(x_\alpha) - \frac{1}{1+\varepsilon} v(x'_\alpha) + \varepsilon \left(c_{\varepsilon,\varphi} + o(1) \right) \\ &\leq \frac{1}{1-\varepsilon} h_1(x_\alpha) - \frac{1}{1+\varepsilon} h_2(x'_\alpha) + \lambda \left[\frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \right] \\ &\qquad + \varepsilon \left(c_{\varepsilon,\varphi} + o(1) \right) \\ &\leq h_1(x_\alpha) - h_2(x'_\alpha) + \lambda \left[\frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \right] \\ &\qquad + \frac{\varepsilon}{1-\varepsilon} ||h_1|| + \frac{\varepsilon}{1+\varepsilon} ||h_2|| + \varepsilon \left(c_{\varepsilon,\varphi} + o(1) \right). \end{aligned}$$

We next expand $c_{\varepsilon,\varphi}$ from (6.6.11). Furthermore, taking $\liminf_{\alpha \to \infty}$ on the right-hand side, using Proposition 6.5.1 (j) to treat the difference $h_1 - h_2$, and (6.6.4) to treat the difference of Hamiltonians, we find

$$\begin{aligned} \lceil u - v \rceil_K &\leq \lceil h_1 - h_2 \rceil_{\widehat{K}} + \lambda \left(\varepsilon C_0 + C_{\varepsilon,\varphi} \right) + \frac{\varepsilon}{1-\varepsilon} ||h_1|| + \frac{\varepsilon}{1+\varepsilon} ||h_2|| \\ &\qquad + \varepsilon \left(\frac{2}{1-\varepsilon^2}(1-\varphi) \lceil V \rceil_K - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K \right). \end{aligned}$$

As $\varphi \in (0,1]$ was arbitrary, we can take the limit for $\varphi \downarrow 0$, which leads to

$$\begin{aligned}\lceil u - v \rceil_K &\leq \lceil h_1 - h_2 \rceil_{\widehat{K}} \\ &+ \varepsilon \left(\lambda C_\varepsilon^0 + \frac{2}{1-\varepsilon^2} \lceil V \rceil_K + \frac{1}{1-\varepsilon} ||h_1|| + \frac{1}{1+\varepsilon} ||h_2|| - \left\lfloor \frac{1}{1-\varepsilon} u - \frac{1}{1+\varepsilon} v \right\rfloor_K \right),\end{aligned}$$

establishing the claim. □

6.6.3 *Proof of Theorem 6.3.1*

We start with an auxiliary lemma that provides a detailed decomposition of the operators $\mathbb{A}$ and $\mathbb{B}$ evaluated in the test functions.

Lemma 6.6.4. *Let $\mathbb{A}$ and $\mathbb{B}$ both satisfy Assumption 6.3.4 and Assumption 6.3.5 (a) and (b), respectively. Fix $z_0, z_1 \in \mathbb{R}^q$ and $p \in \mathbb{R}^q$. Let $\Xi = \Xi_{z_0,p,z_1}$ as in Definition 6.2.11 and, for $\widehat{f} \in C_c^\infty(E)$, $\varepsilon \in (0,1)$, and $\varphi \in (0,1]$, set*

$$\begin{aligned}\widehat{f_\dagger} &:= (1-\varepsilon)\widehat{f} + \varepsilon(1-\varphi)V + \varepsilon\varphi\Xi, \\ \widehat{f_\ddagger} &:= (1+\varepsilon)\widehat{f} - \varepsilon(1-\varphi)V - \varepsilon\varphi\Xi.\end{aligned}$$

For $z \in E$, set $f_\dagger = \widehat{f_\dagger} \circ s_z$, and $f_\ddagger = \widehat{f_\ddagger} \circ s_z$. Then, the following statements hold:

(a) $f_\dagger \in \mathcal{D}(A_+)$ and $f_\ddagger \in \mathcal{D}(A_-)$. Suppose furthermore that $\mathbb{A}$ is linear on its domain, then

$$\begin{aligned}\frac{A_+ f_\dagger}{1-\varepsilon} &= A(\widehat{f} \circ s_z) + \frac{\varepsilon}{1-\varepsilon}(1-\varphi)A_+(V \circ s_z) + \frac{\varepsilon}{1-\varepsilon}\varphi\mathbb{A}(\Xi \circ s_z), \qquad (6.6.12) \\ \frac{A_- f_\ddagger}{1+\varepsilon} &= A(\widehat{f} \circ s_z) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)A_+(V \circ s_z) - \frac{\varepsilon}{1+\varepsilon}\varphi\mathbb{A}(\Xi \circ s_z),\end{aligned}$$

(b) $f_\dagger, \widehat{f_\dagger} \in \mathcal{D}(B_+)$ and $f_\ddagger, \widehat{f_\ddagger} \in \mathcal{D}(B_-)$. Suppose furthermore that $\mathbb{B}$ is convex, then for any x, y such that $z = x - y$, we have

$$\begin{aligned}\frac{B_+ f_\dagger}{1-\varepsilon}(x) &\leq \frac{1}{1-\varepsilon}\left(B_+ f_\dagger(x) - B_+\widehat{f_\dagger}(y)\right) + B\widehat{f}(y) \qquad (6.6.13) \\ &\qquad + \frac{\varepsilon}{1-\varepsilon}(1-\varphi)B_+V(y) + \frac{\varepsilon}{1-\varepsilon}\varphi B_+\Xi(y), \\ \frac{B_- f_\ddagger}{1+\varepsilon}(x) &\geq \frac{1}{1+\varepsilon}\left(B_- f_\ddagger(x) - B_-\widehat{f_\ddagger}(y)\right) + B\widehat{f}(y) \\ &\qquad - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)B_-V(y) - \frac{\varepsilon}{1+\varepsilon}\varphi B_-\Xi(y).\end{aligned}$$

Proof. The domain statements $f_\dagger \in \mathcal{D}(A_+)$, $f_\ddagger \in \mathcal{D}(A_-)$, $f_\dagger, \widehat{f_\dagger} \in \mathcal{D}(B_+)$ and $f_\ddagger, \widehat{f_\ddagger} \in \mathcal{D}(B_-)$ follow by Lemma 6.6.2. The four statements in (6.6.12) and (6.6.13) follow from linearity of A_+ and convexity of B_+. □

Proof of Theorem 6.3.1. To prove inequality (6.3.3), and consequently the strong comparison principle for the Hamilton–Jacobi equation in terms of H, it suffices by Lemma 6.6.1 and Proposition 6.6.3 to establish (6.6.4), which we repeat for readability:

$$\liminf_{\alpha\to\infty} \frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \leq \varepsilon\left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right). \tag{6.6.14}$$

Let $\theta^*_{1,\alpha} \in \Theta_1$ be such that

$$\begin{aligned} H_+ f_\dagger(x_\alpha) &= \sup_{\theta_1\in\Theta_1}\inf_{\theta_2\in\Theta_2}\left\{\mathbb{A}_{\theta_1,\theta_2} f_\dagger(x_\alpha) + \mathbb{B}_{\theta_1,\theta_2} f_\dagger(x_\alpha) - \mathcal{I}(x_\alpha,\theta_1,\theta_2)\right\} \\ &= \inf_{\theta_2\in\Theta_2}\left\{\mathbb{A}_{\theta^*_{1,\alpha},\theta_2} f_\dagger(x_\alpha) + \mathbb{B}_{\theta^*_{1,\alpha},\theta_2} f_\dagger(x_\alpha) - \mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta_2)\right\}. \end{aligned}$$

Such optimizer exists by the compactness of Θ_1 and the lower semi-continuity of $\mathcal{I}$ in θ_1 assumed in (d). By Isaacs' condition (a), we can write

$$H_- f_\ddagger(x'_\alpha) = \inf_{\theta_2\in\Theta_2}\sup_{\theta_1\in\Theta_1}\left\{\mathbb{A}_{\theta_1,\theta_2} f_\ddagger(x'_\alpha) + \mathbb{B}_{\theta_1,\theta_2} f_\ddagger(x_\alpha) - \mathcal{I}(x'_\alpha,\theta_1,\theta_2)\right\}.$$

Then, by compactness of Θ_2 and the upper semi-continuity of $\mathcal{I}$ in θ_2 assumed in (d), we can find $\theta^*_{2,\alpha} \in \Theta_2$ such that

$$H_- f_\ddagger(x'_\alpha) = \sup_{\theta_1\in\Theta_1}\left\{\mathbb{A}_{\theta_1,\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha) + \mathbb{B}_{\theta_1,\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha) - \mathcal{I}(x'_\alpha,\theta_1,\theta^*_{2,\alpha})\right\}.$$

Consequently, we can estimate

$$\begin{aligned} &\frac{1}{1-\varepsilon} H_+ f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon} H_- f_\ddagger(x'_\alpha) \\ &\leq \underbrace{\left[\frac{1}{1-\varepsilon}\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon}\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha)\right]}_{(1)} \\ &\quad + \underbrace{\left[\frac{1}{1-\varepsilon}\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\dagger(x_\alpha) - \frac{1}{1+\varepsilon}\mathbb{B}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha)\right]}_{(2)} \\ &\quad + \underbrace{\left[\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right]}_{(3)}. \end{aligned}$$

We treat (1), (2), and (3) separately. Note, that due the compactness of Θ_1 and Θ_2, the sequences of optimizers $\theta^*_{1,\alpha}$ and $\theta^*_{2,\alpha}$ converge to some θ^*_1 and θ^*_2, respectively.

Estimate (1)**:** Using the expansions of $A_+ f_\dagger$ and $A_- f_\ddagger$ obtained in Lemma 6.6.4 we find

$$\begin{aligned}
&\frac{\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{\mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_\ddagger(x'_\alpha)}{1+\varepsilon} \\
&= \frac{A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+} f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},-} f_\ddagger(x'_\alpha)}{1+\varepsilon} \\
&\leq A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_1(x_\alpha) - A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_2(x'_\alpha) \\
&\quad + \frac{\varepsilon}{1-\varepsilon}(1-\varphi) A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+} \left(V \circ s_{x_\alpha - y_\alpha}\right)(x_\alpha) \\
&\quad + \frac{\varepsilon}{1+\varepsilon}(1-\varphi) A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+} \left(V \circ s_{x'_\alpha - y'_\alpha}\right)(x'_\alpha) \\
&\quad + \frac{\varepsilon}{1-\varepsilon}\varphi \mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(\Xi_1 \circ s_{x_\alpha - y_\alpha}\right)(x_\alpha) \\
&\quad + \frac{\varepsilon}{1+\varepsilon}\varphi \mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(\Xi_2 \circ s_{x'_\alpha - y'_\alpha}\right)(x'_\alpha).
\end{aligned} \tag{6.6.15}$$

We first consider the terms involving V and Ξ. By Proposition 6.5.1 (j), we have that, along subsequences, the optimizers $(x_\alpha, y_\alpha, y_{\alpha,0}, y'_{\alpha,0}, y'_\alpha, x'_\alpha)$ converge to (z,z,z,z,z,z) with $z \in \widehat{K}$ and $p_\alpha, p'_\alpha \in B_{1/\alpha}(0)$. Then, using the compatibility of $\mathbb{A}_{\theta_1,\theta_2}$, cf. Assumption 6.3.5, we find

$$\begin{aligned}
\liminf_{\alpha\to\infty} &\frac{\varepsilon}{1-\varepsilon}(1-\varphi) A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+} \left(V \circ s_{x_\alpha - y_\alpha}\right)(x_\alpha) \\
&+ \frac{\varepsilon}{1+\varepsilon}(1-\varphi) A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha},+} \left(V \circ s_{x'_\alpha - y'_\alpha}\right)(x'_\alpha) \\
&+ \frac{\varepsilon}{1-\varepsilon}\varphi \mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(\Xi_1 \circ s_{x_\alpha - y_\alpha}\right)(x_\alpha) \\
&+ \frac{\varepsilon}{1+\varepsilon}\varphi \mathbb{A}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(\Xi_2 \circ s_{x'_\alpha - y'_\alpha}\right)(x'_\alpha) \\
&\leq \frac{2\varepsilon}{1-\varepsilon^2}\Big((1-\varphi) A_{\theta^*_1,\theta^*_2,+}(V)(z) + \varphi \mathbb{A}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\Big).
\end{aligned} \tag{6.6.16}$$

Next, we consider the second line in (6.6.15). Using that, for all θ_1, θ_2, $\mathbb{A}_{\theta_1,\theta_2}$ has a controlled growth coupling $\widehat{\mathbb{A}}_{\theta_1,\theta_2}$ with a modulus uniform in θ_1 and θ_2 satisfying the maximum principle and Proposition 6.5.3 (a), we find

$$\begin{aligned}
A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_1(x_\alpha) - A_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} f_2(x'_\alpha) &= \widehat{\mathbb{A}}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(f_1 \ominus f_2\right)(x_\alpha, x'_\alpha) \\
&\leq \widehat{\mathbb{A}}_{\theta^*_{1,\alpha},\theta^*_{2,\alpha}} \left(\frac{\alpha}{2} d^2_{x_\alpha - y_\alpha, x'_\alpha - y'_\alpha}\right)(x_\alpha, x'_\alpha) \\
&\leq \omega_{\widehat{\mathbb{A}},\widehat{K}} \Big(\alpha \left(d(x_\alpha,y_\alpha) + d(y_\alpha,y'_\alpha) + d(y'_\alpha,x'_\alpha)\right)^2 \\
&\qquad + \left(d(x_\alpha,y_\alpha) + d(y_\alpha,y'_\alpha) + d(y'_\alpha,x'_\alpha)\right)\Big),
\end{aligned} \tag{6.6.17}$$

which converges to 0 as $\alpha \to \infty$ by Proposition 6.5.1 (g).

Estimate (2)**:** By using the expansions of $B_+ f_\dagger$ and $B_- f_\ddagger$ obtained in Lemma 6.6.4, we find

$$\begin{aligned}
\frac{\mathbb{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} f_\ddagger(x'_\alpha)}{1+\varepsilon} &= \frac{B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,-} f_\ddagger(x'_\alpha)}{1+\varepsilon} \\
&\leq B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \widehat{f}_1(y_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \widehat{f}_2(y'_\alpha) \\
&\quad + \frac{1}{1-\varepsilon}\Big(B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} f_\dagger(x_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} \widehat{f}_\dagger(y_\alpha)\Big) \\
&\quad + \frac{1}{1+\varepsilon}\Big(B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,-} \widehat{f}_\ddagger(y'_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,-} f_\ddagger(x'_\alpha)\Big) \\
&\quad + \frac{\varepsilon}{1-\varepsilon}(1-\varphi) B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi) B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} V(y'_\alpha) \\
&\quad + \frac{\varepsilon}{1-\varepsilon}\varphi \mathbb{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \Xi_1(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi \mathbb{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \Xi_2(y'_\alpha).
\end{aligned}$$

Again, by sending $\alpha \to \infty$, using Proposition 6.5.1 (j), and the compatibility of $\mathbb{B}_{\theta_1,\theta_2}$, cf. Assumption 6.3.5, we obtain that

$$\begin{aligned}
\liminf_{\alpha\to\infty} & \frac{\varepsilon}{1-\varepsilon}(1-\varphi) B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} V(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}(1-\varphi) B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} V(y'_\alpha) \\
&+ \frac{\varepsilon}{1-\varepsilon}\varphi \mathbb{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \Xi_1(y_\alpha) + \frac{\varepsilon}{1+\varepsilon}\varphi \mathbb{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \Xi_2(y'_\alpha) \\
&\leq \frac{2\varepsilon}{1-\varepsilon^2}\Big((1-\varphi) B_{\theta_1^*,\theta_2^*,+}(V)(z) + \varphi \mathbb{B}_{\theta_1^*,\theta_2^*}(\Xi_{z,0,z})(z)\Big).
\end{aligned} \tag{6.6.18}$$

Using that, for all θ_1, θ_2, $\mathbb{B}_{\theta_1,\theta_2}$ is semi-monotone with $\mathcal{B}_{\theta_1,\theta_2}$ and the expressions for the gradients obtained in Proposition 6.5.3, we find that

$$\begin{aligned}
\frac{1}{1-\varepsilon} & \Big(B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} f_\dagger(x_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,+} \widehat{f}_\dagger(y_\alpha)\Big) \\
&+ B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \widehat{f}_1(y_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*} \widehat{f}_2(y'_\alpha) \\
&+ \frac{1}{1+\varepsilon}\Big(B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,-} \widehat{f}_\ddagger(y'_\alpha) - B_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*,-} f_\ddagger(x'_\alpha)\Big) \\
= \frac{1}{1-\varepsilon} & \Big(\mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(x_\alpha, \alpha(x_\alpha - y_\alpha)) - \mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(y_\alpha, \alpha(x_\alpha - y_\alpha))\Big) \\
&+ \mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(y_\alpha, \alpha(y_\alpha - y'_\alpha)) - \mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(y'_\alpha, \alpha(y_\alpha - y'_\alpha)) \\
&+ \frac{1}{1+\varepsilon}\Big(\mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(y_\alpha, \alpha(y'_\alpha - x'_\alpha)) - \mathcal{B}_{\theta_{1,\alpha}^*,\theta_{2,\alpha}^*}(x'_\alpha, \alpha(y'_\alpha - y'_\alpha))\Big).
\end{aligned} \tag{6.6.19}$$

By the semi-monotonicity property of $\mathbb{B}_{\theta_1,\theta_2}$, (6.6.19) is bounded by

$$\begin{aligned}
\frac{1}{1-\varepsilon} & \omega_{\mathcal{B},\widehat{K}}(d(x_\alpha, y_\alpha) + \alpha d^2(x_\alpha, y_\alpha)) + \omega_{\mathcal{B},\widehat{K}}(d(y_\alpha, y'_\alpha) + \alpha d^2(y_\alpha, y'_\alpha)) \\
&+ \frac{1}{1+\varepsilon}\omega_{\mathcal{B},\widehat{K}}(d(y'_\alpha, x'_\alpha) + \alpha d^2(y'_\alpha, x'_\alpha)).
\end{aligned} \tag{6.6.20}$$

Thus, taking the $\liminf_{\alpha\to\infty}$ gives 0 by Proposition 6.5.1 (g).

Estimate (3)**:** We have

$$\begin{aligned}
&\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\
&= \left[\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right] \\
&\qquad - \frac{\varepsilon}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}).
\end{aligned}$$

By assumption, $\mathcal{I}$ admits a modulus of continuity $\omega_{\mathcal{I},K}$, uniform in θ_1, θ_2, implying

$$\begin{aligned}
&\frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\
&\leq \omega_{\mathcal{I},\widehat{K}}(d(x_\alpha,x'_\alpha)) - \frac{\varepsilon}{1-\varepsilon}(1-\varphi)\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}).
\end{aligned}$$

Sending $\alpha \to \infty$, using the lower semi-continuity of $\mathcal{I}$, and using Proposition 6.5.1 (j), we find

$$\begin{aligned}
&\liminf_{\alpha\to\infty} \frac{1}{1+\varepsilon}\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{1}{1-\varepsilon}\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) \\
&\quad \leq \liminf_{\alpha\to\infty} \omega_{\mathcal{I},\widehat{K}}(d(x_\alpha,x'_\alpha)) \\
&\quad\quad + \limsup_{\alpha\to\infty}\left[-\frac{\varepsilon}{1-\varepsilon}(1-\varphi)\mathcal{I}(x_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha}) - \frac{\varepsilon}{1+\varepsilon}(1-\varphi)\mathcal{I}(x'_\alpha,\theta^*_{1,\alpha},\theta^*_{2,\alpha})\right] \\
&\quad \leq -\frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\mathcal{I}(z,\theta^*_1,\theta^*_2).
\end{aligned} \tag{6.6.21}$$

Conclusion: Putting together (6.6.16), (6.6.17), (6.6.18), (6.6.20), and (6.6.21), we can conclude that

$$\begin{aligned}
&\liminf_{\alpha\to\infty} \frac{H_+ f_\dagger(x_\alpha)}{1-\varepsilon} - \frac{H_- f_\ddagger(x'_\alpha)}{1+\varepsilon} \\
&\leq \frac{2\varepsilon}{1-\varepsilon^2}\Big((1-\varphi)A_{\theta^*_1,\theta^*_2,+}V(z) + \varphi\mathbb{A}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\Big) \\
&\qquad + \frac{2\varepsilon}{1-\varepsilon^2}\Big((1-\varphi)B_{\theta^*_1,\theta^*_2,+}V(z) + \varphi\mathbb{B}_{\theta^*_1,\theta^*_2}(\Xi_{z,0,z})(z)\Big) \\
&\qquad - \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\mathcal{I}(z,\theta^*_1,\theta^*_2) \\
&\leq \frac{2\varepsilon}{1-\varepsilon^2}(1-\varphi)\left[(A_{\theta^*_1,\theta^*_2,+} + B_{\theta^*_1,\theta^*_2,+})(V) - \mathcal{I}(\cdot,\theta^*_1,\theta^*_2)\right] \\
&\quad + \frac{2\varepsilon}{1-\varepsilon^2}\varphi\left[(\mathbb{A}_{\theta^*_1,\theta^*_2} + \mathbb{B}_{\theta^*_1,\theta^*_2})(\Xi_{\cdot,0,\cdot})\right]_{\widehat{K}_\varepsilon} \\
&\leq \varepsilon\left(\frac{2}{1-\varepsilon^2}c_V + \frac{2}{1-\varepsilon^2}\varphi\left[(\mathbb{A}_{\theta^*_1,\theta^*_2} + \mathbb{B}_{\theta^*_1,\theta^*_2})(\Xi_{\cdot,0,\cdot})\right]_{\widehat{K}_\varepsilon}\right) \\
&\leq \varepsilon\left(C^0_\varepsilon + C_{\varepsilon,\varphi}\right)
\end{aligned}$$

with c_V given by (6.3.1), and C^0_ε and $C_{\varepsilon,\varphi}$ defined via the last two lines. The estimate on the difference of Hamiltonians (6.6.14) and thus (6.6.4) of Proposition 6.6.3 are satisfied. As a consequence our final estimate (6.3.3) and, consequently, the strong comparison principle follow. □

6.7 APPENDICES

6.7.1 *The Jensen perturbation*

The main result of this section is Proposition 6.7.1 that allows us to perturb a semi-convex function with a unique extreme point such that we get a new extreme point close by, in which the function is twice differentiable. The result is a variant of the well-known perturbation result by Jensen, see e.g. [CIL92, Lemma A.3].

Proposition 6.7.1. *Fix $\eta > 0$. Let $\phi : E \times E \to \mathbb{R}$ be bounded above and semi-convex with convexity constant $\kappa \geq 1$. Suppose that (x_0, y_0) is an optimizer of*

$$\phi(x_0, y_0) = \lceil \phi \rceil .$$

Let $R > 0$, $\{\zeta_{z,p}\}_{z\in E, p\in\mathbb{R}^q} \subset C(E)$ and $\{\xi_z\}_{z\in E} \subset C^1(E)$ and semi-concavity constant κ_ξ be as in Definition 6.2.11.

Fix $\varepsilon_1, \varepsilon_2 > 0$ such that $1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$. Furthermore, define for $p = (p_1, p_2) \in \mathbb{R}^q \times \mathbb{R}^q$ the perturbed functions

$$\phi_p(x, y) := \phi(x, y) - \varepsilon_1 \left(\xi_{x_0}(x) + \zeta_{x_0,p_1}(x)\right) - \varepsilon_2 \left(\xi_{y_0}(y) + \zeta_{y_0,p_2}(y)\right) . \tag{6.7.1}$$

Then there exist $p_1, p_2 \in B_\eta(0)$, and a pair $(x_1, y_1) \in B_\eta(x_0) \times B_\eta(y_0)$ globally maximizing ϕ_p at which ϕ_p is twice differentiable.

Corollary 6.7.2. *For $\eta > 0$, p and (x_1, y_1) as in Proposition 6.7.1, we have*

$$0 \leq -\varepsilon_1 \left(\xi_{x_0}(x_1) + \zeta_{x_0,p_1}(x_1)\right) - \varepsilon_2 \left(\xi_{y_0}(y_1) + \zeta_{y_0,p_2}(y_1)\right) \leq \varepsilon_1\eta + \varepsilon\eta_2, \tag{6.7.2}$$

and

$$\lceil \phi \rceil \leq \phi_{p,\varepsilon}(x_1, y_1) \leq \lceil \phi \rceil + \varepsilon_1\eta + \varepsilon\eta_2. \tag{6.7.3}$$

The proof of the perturbation proposition is based partly on results from set-valued analysis. To facilitate the proof, we first introduce the necessary auxiliary definitions and results.

Definition 6.7.3. A set-valued function $\Gamma : A \rightrightarrows B$ is called *upper hemi-continuous at* $a \in A$, if, for all open neighbourhoods $V \subseteq B$ of $\Gamma(a)$ (meaning that $\Gamma(a) \subseteq V$), there exists a neighbourhood U of a such that, for all $x \in U$, we have $\Gamma(x) \subseteq V$.

If A, B are metric, this can equivalently formulated in terms of sequences: A set-valued map $\Gamma : A \rightrightarrows B$, which takes closed values, is upper hemi-continuous at a, if, for any sequence $a_n \to a$ and $b_n \in \Gamma(a_n)$ satisfying $b_n \to b$, we have $b \in \Gamma(a)$.

We say that Γ is upper hemi-continuous, if it is upper hemi-continuous at all points.

Lemma 6.7.4. *Let K be a compact metric space and let Ξ be a metric space.*

For any $\xi \in \Xi$, let $\phi_\xi \in C(K)$ and suppose that the map $\xi \mapsto \phi_\xi$ is continuous from Ξ to $C(K)$, endowed with the supremum norm on K. Then the set-valued map Opt : $\Xi \rightrightarrows K$ defined by

$$\mathit{Opt}(\xi) := \left\{ x \in K \,\middle|\, \phi_\xi \text{ has a maximum at } x \right\}$$

is upper hemi-continuous.

Proof. The result follows immediately from Berge's Maximum Theorem [AB06, Theorem 17.31] with $\xi \mapsto \mathrm{image}_{\phi_\xi}(K)$ being the relevant set-valued map. □

Remark 6.7.5. In the proof below, we will make use of the notion of a $\limsup$ of sets. For a sequence of sets $(A_n)_{n\in\mathbb{N}}$ denote

$$\limsup_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{m\geq n} A_m$$

to be interpreted as $x \in \limsup_{n\to\infty} A_n$ if and only if there are infinitely many $n \in \mathbb{N}$ such that $x \in A_n$.

The following proof is a variant of the proof of [CIL92, Lemma A.3] and [CS04, Theorem 2.3.3].

Proof of Proposition 6.7.1. For notational convenience, we will write $w = (x, y)$ and $w_0 = (x_0, y_0)$. Let $R > 0$ and $\{\zeta_{z,p}\}_{z\in E, p\in\mathbb{R}^q} \subset C(E)$ and $\{\xi_z\}_{z\in E} \subset C(E)$ be two collections of functions as in Definition 6.2.11. Without loss of generality, we can assume that $R \geq \eta$.

We start out by making z_0 the unique optimizer by replacing ϕ by

$$\widehat{\phi}(w) = \phi(w) - \varepsilon_1 \xi_{x_0}(x) - \varepsilon_2 \xi_{y_0}(y).$$

Note that as $1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$ the map $\widehat{\phi}$ is semi-convex and bounded from above with a unique optimizer w_0.

Our next step is to locally, linearly perturb $\widehat{\phi}$ to obtain ϕ_p as in equation (6.7.1). This procedure produces a new optimizer close to w_0 in which the perturbed function ϕ_p is twice differentiable.

To further facilitate the analysis of optimizers, we smoothen out ϕ. To that end, let $C_\delta \colon C_b(E) \to C_b^2(E)$ be a mollifier with $\sup_{\delta>0} ||C_\delta f|| < \infty$ and $C_\delta f \to f$ uniformly on compacts as $\delta \downarrow 0$. Define

$$\phi_{p,\delta}(w) := (C_\delta\phi)(w) - \varepsilon_1 \left(\xi_{x_0}(x) + \zeta_{x_0,p_1}(x)\right) - \varepsilon_2 \left(\xi_{y_0}(y) + \zeta_{y_0,p_2}(y)\right),$$

where we will read $C_0 = \mathbb{1}$ such that $\phi_{p,0} = \phi_p$ and $\phi_{0,0} = \widehat{\phi}$.

We next study the optimizers for the map $(p, \delta) \mapsto \phi_{p,\delta}$ on $\Xi = (B_1(0) \times B_1(0)) \times [0, 1]$ using Berge's Maximum Theorem with $K = \overline{B_R(w_0)}$. Set

$$\mathrm{Opt}(p, \delta) := \left\{ w \in \overline{B_R(w_0)} \,\middle|\, \phi_{p,\delta} \text{ has a local maximum at } w \in \overline{B_R(w_0)} \right\}.$$

First note that the local nature of the problem can be removed due to the fact that the perturbations all vanish in w_0, whereas they add up to something negative outside the ball $\overline{B_R(w_0)}$ by Definition 6.2.11 (d), implying that

$$\mathrm{Opt}(p,\delta) = \left\{ w \in \overline{B_R(w_0)} \,\middle|\, \phi_{p,\delta} \text{ has a global maximum at } w \right\}. \tag{6.7.4}$$

Applying Lemma 6.7.4 to $(p,\delta) \mapsto \phi_{p,\delta}$ on $\Xi = (B_\eta(0) \times B_\eta(0)) \times [0,1]$ with $K = \overline{B_R(w_0)}$, we find that the set-valued map Opt: $\Xi \rightrightarrows K \subseteq \mathbb{R}^q \times \mathbb{R}^q$, as defined above, is upper hemi-continuous in the variables (p,δ). We can thus find a closed set U with 0 in its interior satisfying

$$U \subseteq B_\eta(0) \times B_\eta(0) \tag{6.7.5}$$

and $\delta_0 > 0$ such that, if $p = (p_1, p_2) \in U$ and $\delta < \delta_0$, then

$$\mathrm{Opt}(p,\delta) \subseteq \mathrm{Opt}(0,0) \oplus B_\eta(0) = B_\eta(w_0), \tag{6.7.6}$$

as the unique optimizer of $\widehat{\phi}$ is w_0.

We next aim to show that the set of such optimizers has positive Lebesgue measure m. Recall that $\kappa := 1 - (\varepsilon_1 + \varepsilon_2)\kappa_\xi > 0$ is the semi-convexity constant of $\widehat{\phi}$. In particular, we will proceed to show the following steps.

Step 1: For any $\delta \in (0, \delta_0)$ we have $m(\mathrm{Opt}(U,\delta)) \geq |\kappa|^{-2d} m(U) > 0$.

Step 2: We take the limit $\delta \downarrow 0$ to obtain $m(\mathrm{Opt}(U,0)) \geq |\kappa|^{-2d} m(U) > 0$.

Step 1. By definition, all perturbations are at least once continuously differentiable on $\overline{B_R(w_0)}$. It follows that for $p \in U$, $\delta \in (0,\delta_0)$ and $w \in \mathrm{Opt}(p,\delta)$ we have that $D(C_\delta\phi)(z) = p$. This, in turn, implies that, for fixed $\delta \in (0,\delta_0)$,

$$U \subseteq (\mathrm{Opt}(\cdot,\delta))^{-1} (\mathrm{Opt}(U,\delta)) \subseteq D(C_\delta\phi)(\mathrm{Opt}(U,\delta)). \tag{6.7.7}$$

We next argue towards a lower bound on the measure of $\mathrm{Opt}(U,\delta)$ for $\delta \in (0,\delta_0)$. We exclude $\delta = 0$ here, due to the possible non-smoothness of ϕ. As the convolution operator is taking averages, the semi-convexity of ϕ carries over to $C_\delta\phi$, which yields

$$-\kappa I_{2d} \leq D^2(C_\delta\phi)(w) \tag{6.7.8}$$

for all $w \in E^2$. On the other hand, if $w \in \mathrm{Opt}(U,\delta)$, we know that there is some $p \in U$ such that w maximizes $\phi_{p,\delta}$, implying that $D^2(C_\delta\phi)(w) \leq 0$. Applying (6.7.7), the chain rule, and (6.7.8), we thus obtain that, for any $\delta \in (0,\delta_0)$,

$$\begin{aligned} m(U) &\leq m(D(C_\delta\phi)(\mathrm{Opt}(U,\delta))) \\ &= \int_{\mathrm{Opt}(U,\delta)} \left|\det D^2(C_\delta\phi)(w)\right| \mathrm{d}w \leq m(\mathrm{Opt}(U,\delta))|\kappa|^{2q} \end{aligned}$$

leading to the lower bound

$$0 < |\kappa|^{-2q} m(U) \leq m(\mathrm{Opt}(U,\delta)), \tag{6.7.9}$$

as U has non-empty interior, establishing the claim of Step 1.

Step 2. Next, we transfer our bound to $m(\mathrm{Opt}(U,0))$. We first establish that

$$\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta) \subseteq \mathrm{Opt}(U,0), \tag{6.7.10}$$

see Remark 6.7.5 for the definition of the $\limsup$ of sets. To that end, we pick an element $w \in \limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)$. By definition we can find a sequence $\delta_n \downarrow 0$ such that $w \in \mathrm{Opt}(U,\delta_n)$ for all $n \in \mathbb{N}$. Then, there are $p_n \in U$ such that w is an optimizer for $(\phi_{p_n,\delta_n})_{n\in\mathbb{N}}$. By the closedness of U and (6.7.5), U is compact, and we can therefore extract a subsequence from $(p_n)_{n\in\mathbb{N}}$ that converges to some $p_0 \in U$. By upper semi-continuity of the map $(p,\delta) \mapsto \phi_{p,\delta}$, see Lemma 6.7.4, we find that w maximizes ϕ_{p_0}, or in other words, $w \in \mathrm{Opt}(U,0)$.

Thus, by (6.7.10), it suffices to lower bound the volume of $\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)$. As a first step, note that (6.7.9) leads to

$$m\left(\bigcup_{\delta'\leq\delta} \mathrm{Opt}(U,\delta')\right) \geq |\kappa|^{-2q} m(U)$$

for any $\delta \in (0,\delta_0)$. Consequently, as

$$\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta) = \bigcap_{\delta\in(0,\delta_0)} \bigcup_{\delta'\leq\delta} \mathrm{Opt}(U,\delta'),$$

by continuity from above of the Lebesgue measure m, we find that

$$m\left(\limsup_{\delta\downarrow 0} \mathrm{Opt}(U,\delta)\right) = \lim_{\delta\downarrow 0} m\left(\bigcup_{\delta'\leq\delta} \mathrm{Opt}(U,\delta')\right) \geq |\kappa|^{-2q} m(U).$$

By (6.7.10), we conclude that

$$m(\mathrm{Opt}(U,0)) \geq |\kappa|^{-2q} m(U) > 0,$$

establishing the claim of Step 2.

We proceed by verifying that we can now find $p \in U$ with an optimizer z_1 in $B_\eta(z_0)$ in which ϕ_p is twice differentiable.

First of all, recall that, by (6.7.6), we have

$$\mathrm{Opt}(U,0) \subseteq B_\eta(z_0).$$

Furthermore, by Alexandrov's theorem [CS04, Theorem 2.3.1], the set of points in $B_\eta(z_0)$ where ϕ_p is twice differentiable has full measure. As the measure of $\mathrm{Opt}(U,0)$ is positive, it follows that there exist $z_1 \in B_\eta(z_0)$ and $p \in U$ such that ϕ_p is twice differentiable in z_1 and has a local maximum at z_1 in $\overline{B_R(z_0)}$. Finally, recall from (6.7.4) that the local optimizer is in fact a global optimizer. This establishes the claim. □

Proof of Corollary 6.7.2. We stay in the context of Proposition 6.7.1 and proceed with the proof of (6.7.3). By construction, we have

$$\lceil \phi \rceil = \phi(x_0, y_0) = \phi_{p,\varepsilon}(x_0, y_0) \leq \lceil \phi_{p,\varepsilon} \rceil = \phi_{p,\varepsilon}(x_1, y_1). \tag{6.7.11}$$

This implies the lower bound of (6.7.2). Note that by the properties of $\xi_{x_0}, \xi_{y_0}, \zeta_{x_0,p_1}$ and ζ_{y_0,p_2} we have

$$\begin{aligned}
&- \varepsilon_1 \left(\xi_{x_0}(x_1) + \zeta_{x_0,p_1}(x_1) \right) - \varepsilon_2 \left(\xi_{y_0}(y_1) + \zeta_{y_0,p_2}(y_1) \right) \\
&\leq \varepsilon_1 |p_1| d(x_0, x_1) + \varepsilon_2 |p_2| d(y_0, y_1) \\
&\leq \varepsilon_1 \eta + \varepsilon_2 \eta,
\end{aligned}$$

leading to the upper bound of (6.7.2). Consequently,

$$\begin{aligned}
\phi_{p,\varepsilon}(x_1, y_1) &\leq \phi(x_1, y_1) + \varepsilon_1 \eta + \varepsilon_2 \eta \\
&\leq \lceil \phi \rceil + \varepsilon_1 \eta + \varepsilon_2 \eta.
\end{aligned} \tag{6.7.12}$$

Combining (6.7.11) and (6.7.12), finally yields (6.7.3). □

6.7.2 *Smooth test function construction*

The main result of this section is Lemma 6.7.6, in which we construct a smooth test function that lies between a function that is twice differentiable in one point and a perturbed version of that function.

Lemma 6.7.6. *Let Π_1, Π_1^0, Π_2, and Π_2^0 be as in the proof of Proposition 6.5.3.*
Then, there exist $f_1, f_2 \in C^\infty(E)$ such that, for all $y \in E$,

$$\begin{aligned}
\Pi_1(y) \leq & f_1(y) \leq \Pi_1^0(y), \\
\Pi_2(y) \geq & f_2(y) \geq \Pi_2^0(y)
\end{aligned}$$

with equality only in y_α and y'_α, respectively.

Proof. As in the proof of Proposition 6.5.3, we only consider the case

$$\Pi_1(y) \leq f_1(y) \leq \Pi_1^0(y),$$

for $y \in E$ with equality only in y_α, since the other statement follows analogously.

Our goal is to find f_1, by first constructing a function that is squeezed between Π_1 and Π_1^0, using the Whitney Extension Theorem [Hö3, Theorem 2.3.6], and then modifying it to obtain f_1.

Recall that, by construction, we have that

$$\Pi_1(y) < \Pi_1^0(y) \qquad \text{for } y \in E \setminus \{y_\alpha\}$$

and

$$\Pi_1(y_\alpha) = \Pi_1^0(y_\alpha), \quad D\Pi_1(y_\alpha) = D\Pi_1^0(y_\alpha), \quad D^2\Pi_1(y_\alpha) < D^2\Pi_1^0(y_\alpha).$$

We apply the Whitney Extension Theorem to $\frac{1}{2}(\Pi_1 + \Pi_1^0)$ on the closed set $A = \{y_\alpha\}$, yielding a function $\psi_1 \in C^2(E)$ such that $\Pi_1 \leq \psi_1 \leq \Pi_1^0$ on $B_{2\delta}(y_\alpha)$ for some $\delta > 0$ with equality only in y_α. Inspecting the construction of ψ_1 in the proof of [Whi34, Theorem II], we find that $\psi_1 \in C^\infty(E)$.

Next, we modify ψ_1 such that the resulting function stays between Π_1 and Π_1^0 on all of E. As smooth functions are dense in the set of continuous functions, we can find a function $\psi_2 \in C^\infty(E)$ such that $\Pi_1 < \psi_1 < \Pi_1^0$ on $E \setminus B_\delta(y_\alpha)$.

Then, defining

$$f_1(y) = \ell(y)\psi_1(y) + (1 - \ell(y))\psi_2(y),$$

where ℓ is a smooth function that is 1 on $B_\delta(y_\alpha)$ and 0 outside of $B_{2\delta}(y_\alpha)$, for example ℓ as defined as point (3) on [Spi70, p. 33]. This concludes the proof. □

6.7.3 *Convergence of integrals*

Lemma 6.7.7. *Let $\mathcal{X}$ be a Polish space, $W : \mathcal{X} \to (0, \infty)$ be a continuous function, and ν_n, ν_∞ be non-negative Borel measures with $\int_{\mathcal{X}} W \, d\nu_n < \infty$ for all $n \in \mathbb{N}$ and*

$$\lim_{n\to\infty} \int \phi \, d\nu_n = \int \phi \, d\nu_\infty \in \mathbb{R} \tag{6.7.13}$$

for every function $\phi \in C(\mathcal{X})$ with $|\phi(x)| \leq W(x)$ for all $x \in \mathcal{X}$. Moreover, let $\phi_n, \phi_\infty \in C(\mathcal{X})$ with $\phi_n \to \phi_\infty$ uniformly on compacts and $\sup_{n\in\mathbb{N}} \sup_{x\in\mathcal{X}} \frac{|\phi_n(x)|}{W(x)} < \infty$. Then,

$$\lim_{n\to\infty} \int \phi_n \, d\nu_n = \int \phi_\infty \, d\nu_\infty.$$

Proof. By assumption, the family $\mu_n := W d\nu_n$ satisfies $C_\mu := \sup_{n\in\mathbb{N}} \mu_n(\mathcal{X}) < \infty$ and

$$C_\phi := \sup_{n\in\mathbb{N}} \sup_{x\in\mathcal{X}} \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} < \infty.$$

Using the fact that a function $\phi \in C(\mathcal{X})$ satisfies $|\phi(x)| \leq W(x)$ for all $x \in \mathcal{X}$ if and only if $\phi = W\psi$ for some $\psi \in C_b(\mathcal{X})$, it follows that $\mu_n \to \mu_\infty := W d\nu_\infty$ weakly. In particular, the family $(\mu_n)_{n\in\mathbb{N}}$ is tight. Hence, for all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subseteq \mathcal{X}$ such that

$$C_\phi \mu_n(\mathcal{X} \setminus K_\varepsilon) < \frac{\epsilon}{3} \quad \text{for all } n \in \mathbb{N}.$$

Now, let $\varepsilon > 0$. By (6.7.13) and since $\phi_n \to \phi_\infty$ uniformly on compacts and W is continuous, there exists some $n_0 \in \mathbb{N}$ such that

$$C_\mu \sup_{x\in K_\varepsilon} \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} < \frac{\varepsilon}{3} \quad \text{and} \quad \left| \int \phi_\infty \, d\nu_n - \int \phi_\infty \, d\nu_\infty \right| < \frac{\varepsilon}{3}.$$

We thus obtain that

$$\begin{aligned}\left| \int \phi_n \, \mathrm{d}\nu_n - \int \phi_\infty \, \mathrm{d}\nu_\infty \right| &\le \int \left|\phi_n - \phi_\infty\right| \mathrm{d}\nu_n + \left| \int \phi_\infty \, \mathrm{d}\nu_n - \int \phi_\infty \, \mathrm{d}\nu_\infty \right| \\ &\le \int_{K_\varepsilon} \left|\phi_n - \phi_\infty\right| \mathrm{d}\nu_n + \int_{\mathcal{X}\setminus K_\varepsilon} \left|\phi_n - \phi_\infty\right| \mathrm{d}\nu_n + \frac{\varepsilon}{3} \\ &\le C_\mu \frac{|\phi_n(x) - \phi_\infty(x)|}{W(x)} + C_\phi \mu_n\big(\mathcal{X} \setminus K_\epsilon\big) + \frac{\varepsilon}{3} < \varepsilon\end{aligned}$$

for all $n \in \mathbb{N}$ with $n \ge n_0$. The proof is complete. □

6.7.4 *Proof of Lemma 6.5.2*

Proof. For the proof of (a), note that, for any $x, y \in E$, we have

$$u(x) - \frac{\alpha}{2} d^2(x, y) \le u(x).$$

This implies that

$$\lceil P^\alpha[u] \rceil = \left\lceil u - \frac{\alpha}{2} d^2 \right\rceil \le \lceil u \rceil \,. \tag{6.7.14}$$

On the other hand, we have

$$u(y) \le \left\lceil u - \frac{\alpha}{2} d^2(\cdot, y) \right\rceil = P^\alpha[u](y).$$

It follows that

$$\lfloor u \rfloor \le \lfloor P^\alpha[u] \rfloor \,. \tag{6.7.15}$$

Now, (a) follows by (6.7.14) and (6.7.15). Part (b) is equivalent to

$$P_\alpha[u] \le u \le P^\alpha[u] \quad \text{on } E,$$

which is immediately clear from the definitions of sup- and inf-convolutions. Part (c) follows similarly from the definitions. For the proof of (d), let $y_0 \in E$. Then, since d is the Euclidean metric, we find

$$P^\alpha[u](y) + \frac{\alpha}{2} d^2(y, y_0) = \left\lceil u + \alpha \left\langle y - y_0, \cdot - y_0 \right\rangle - \frac{\alpha}{2} d^2(\cdot, y_0) \right\rceil,$$

where the right-hand side is convex as it is a supremum over affine functions. By Proposition 2.1.5 and Theorem 2.1.7 of [CS04] the claim follows. Lastly, (e) follows from Theorem 3.4.4 of [CS04] by noting that the sets over which can be optimized are compact due to the boundedness of u and v. □

6.8 CONCLUSIONS E FUTURE PERSPECTIVES

In this chapter, we established a comparison principle for viscosity solutions to second-order Hamilton-Jacobi equations, including Hamilton-Jacobi-Bellman and Isaacs equations, in the presence of general partial integro-differential operators. Our approach provided a novel perspective on the classical doubling-of-variables method by formulating the Ishii-Crandall Lemma within a test function framework. This allowed us to systematically handle non-local integral operators, including those associated with Lévy processes. Additionally, we introduced an adaptation of probabilistic coupling techniques to translate key estimates on the difference of Hamiltonians, thereby unifying the treatment of differential, difference, and integral operators.

We conclude with some interesting future directions:

- Our results have relied on bounded semi-continuous viscosity solutions. A natural extension would be to explore unbounded solutions using Lyapunov control methods. This would allow our framework to accommodate a wider range of problems in stochastic control and large deviations.
- We have focused exclusively on the elliptic case. However, it is likely that our results can be extended to the parabolic setting by adapting the test function framework developed for first-order equations (see Chapter 4) to the second-order case. In particular, a time-dependent version of our arguments needs to be developed.
- Our study has considered Hamilton-Jacobi-Bellman and Isaacs equations with compact control sets and cost functions independent of momenta. Extending these results to cases with non-compact control sets and strong coupling regimes, where the cost function explicitly depends on momenta, presents an interesting challenge. However, given the success of similar techniques in the first-order case, we do not anticipate major difficulties in adapting our approach to this more general setting.

BIBLIOGRAPHY

[AB06] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis*. Third. A hitchhiker's guide. Springer, Berlin, 2006, pp. xxii+703.

[AT96] O. Alvarez and A. Tourin. "Viscosity solutions of nonlinear integro-differential equations." In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire* 13.3 (1996), pp. 293–317. DOI: https://doi.org/10.1016/S0294-1449(16)30106-8.

[AF14] Luigi Ambrosio and Jin Feng. "On a class of first order Hamilton-Jacobi equations in metric spaces." In: *Journal of Differential Equations* 256.7 (2014), pp. 2194 –2245. ISSN: 0022-0396. DOI: http://dx.doi.org/10.1016/j.jde.2013.12.018.

[Awa91] Sayah Awatif. "Équations d'Hamilton-Jacobi du premier ordre avec termes intégro-différentiels. I. Unicité des solutions de viscosité." In: *Comm. Partial Differential Equations* 16.6-7 (1991), pp. 1057–1074. ISSN: 0360-5302,1532-4133. DOI: 10.1080/03605309108820789. URL: https://doi.org/10.1080/03605309108820789.

[Bal+06] Karen Ball, Thomas G. Kurtz, Lea Popovic, and Greg Rempala. "Asymptotic analysis of multiscale approximations to reaction networks." In: *The Annals of Applied Probability* 16.4 (2006), pp. 1925–1961.

[BC97] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* English. Boston, MA: Birkhäuser, 1997, pp. xvii + 570. ISBN: 0-8176-3640-4/hbk.

[BS91] Martino Bardi and Pierpaolo Soravia. "Hamilton-Jacobi Equations With Singular Boundary Conditions on a free Boundary and Applications to Differential Games." In: *Transactions of the American Mathematical Society* 325.1 (1991), pp. 205–229. ISSN: 00029947. URL: http://www.jstor.org/stable/2001667 (visited on 04/05/2024).

[BI08] G. Barles and C. Imbert. "Second-order elliptic integro-differential equations: viscosity solutions' theory revisited." In: *Annales de l'IHP Analyse non linéaire*. Vol. 25. 3. 2008, pp. 567–585.

[Bas88] Richard F. Bass. "Uniqueness in law for pure jump Markov processes." In: *Probability Theory and Related Fields* 79.2 (1988), pp. 271–287.

[BFG13] Lorenzo Bertini, Alessandra Faggionato, and Davide Gabrielli. "Large deviation principles for nongradient weakly asymmetric stochastic lattice gases." In: *Ann. Appl. Probab.* 23.1 (2013), pp. 1–65. ISSN: 1050-5164. DOI: 10.1214/11-AAP805. URL: https://doi.org/10.1214/11-AAP805.

[Ber24] Charles Bertucci. *Stochastic optimal transport and Hamilton-Jacobi-Bellman equations on the set of probability measures*. 2024. arXiv: 2306.04283 [math.AP]. URL: https://arxiv.org/abs/2306.04283.

[Ble+23] Jonas Blessing, Robert Denk, Michael Kupper, and Max Nendel. *Convex monotone semigroups and their generators with respect to Γ-convergence*. 2023. arXiv: 2202.08653 [math.AP]. URL: https://arxiv.org/abs/2202.08653.

[Buc58] R. Creighton Buck. "Bounded continuous functions on a locally compact space." In: *Michigan Math. J.* 5.2 (1958), pp. 95–104. DOI: 10.1307/mmj/1028998054.

[BK00] Krzysztof Burdzy and Wilfrid S. Kendall. "Efficient Markovian couplings: examples and counterexamples." In: *Ann. Appl. Probab.* 10.2 (2000), pp. 362–409. ISSN: 1050-5164,2168-8737. DOI: 10.1214/aoap/1019487348.

[CIL87a] Michael G. CRANDALL, Hitoshi ISHII, and Pierre-Louis LIONS. "Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited." In: *Journal of the Mathematical Society of Japan* 39.4 (1987), pp. 581 –596. DOI: 10.2969/jmsj/03940581. URL: https://doi.org/10.2969/jmsj/03940581.

[CL81] Michael G. CRANDALL and Pierre-Louis LIONS. "Conditions d'unicité pour les solutions generalises des equations d'Hamilton-Jacobi du premier ordre." In: *C. R. Acad. Sci. Paris Sér. I Math.* (1981), pp. 487–502.

[CS97] Fabio Camilli and Antonio Siconolfi. "Discontinuous solutions of a hamilton–jacobi equation with infinite speed of propagation." In: *SIAM Journal on Mathematical Analysis* 28.6 (1997), pp. 1420–1426. DOI: https://doi.org/10.1137/S0036141096298047.

[CS05] Fabio Camilli and Antonio Siconolfi. "Time-Dependent Measurable Hamilton–Jacobi Equations." In: *Communications in Partial Differential Equations* 30.5-6 (2005), pp. 813–847. DOI: 10.1081/PDE-200059292. eprint: https://doi.org/10.1081/PDE-200059292. URL: https://doi.org/10.1081/PDE-200059292.

[CS04] Piermarco Cannarsa and Carlo Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Vol. 58. Springer Science & Business Media, 2004.

[CKK03] M. Chipot, D. Kinderlehrer, and M. Kowalczyk. "A Variational Principle for Molecular Motors." In: *Meccanica* 38.5 (2003), pp. 505–518.

[CHK04] Michel Chipot, Stuart Hastings, and David Kinderlehrer. "Transport in a molecular motor system." In: *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique* 38.6 (2004), pp. 1011–1034.

[CK15] Cristina Costantini and Thomas Kurtz. "Viscosity methods giving uniqueness for martingale problems." In: *Electron. J. Probab.* 20 (2015), no. 67, 1–27. ISSN: 1083-6489. DOI: 10.1214/EJP.v20-3624.

[Cra38] Harald Cramér. "Sur un nouveau théorème-limite de la théorie des probabilités." In: (1938).

[CT22] Harald Cramér and Hugo Touchette. *On a new limit theorem in probability theory (Translation of 'Sur un nouveau théorème-limite de la théorie des probabilités')*. 2022. arXiv: 1802.05988 [math.HO]. URL: https://arxiv.org/abs/1802.05988.

[CEL84] M. G. Crandall, L. C. Evans, and P.-L. Lions. "Some properties of viscosity solutions of Hamilton-Jacobi equations." In: *Trans. Amer. Math. Soc.* 282.2 (1984), pp. 487–502. ISSN: 0002-9947. DOI: 10.2307/1999247.

[CL71] M. G. Crandall and T. M. Liggett. "Generation of Semi-Groups of Nonlinear Transformations on General Banach Spaces." In: *American Journal of Mathematics* 93.2 (1971), pp. 265–298.

[CL94] M. G. Crandall and P.-L. Lions. "Hamilton-Jacobi equations in infinite dimensions. VI. Nonlinear A and Tataru's method refined." In: *Evolution equations, control theory, and biomathematics (Han sur Lesse, 1991)*. Vol. 155. Lecture Notes in Pure and Appl. Math. Dekker, New York, 1994, pp. 51–89.

[CI90] Michael G. Crandall and Hitoshi Ishii. "The maximum principle for semicontinuous functions." In: *Differential and Integral Equations* 3.6 (1990), pp. 1001 –1014. DOI: 10.57262/die/1379101975. URL: https://doi.org/10.57262/die/1379101975.

[CIL87b] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. "Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited." In: *J. Math. Soc. Japan* 39.4 (1987), pp. 581–596. ISSN: 0025-5645,1881-1167. DOI: 10.2969/jmsj/03940581. URL: https://doi.org/10.2969/jmsj/03940581.

[CIL92] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. "User's guide to viscosity solutions of second order partial differential equations." In: *Bull. Am. Math. Soc., New Ser.* 27.1 (1992), pp. 1–67. ISSN: 0273-0979; 1088-9485/e. DOI: 10.1090/S0273-0979-1992-00266-5.

[CL83] Michael G. Crandall and Pierre-Louis Lions. "Viscosity solutions of Hamilton-Jacobi equations." In: *Trans. Amer. Math. Soc.* 277.1 (1983), pp. 1–42. ISSN: 0002-9947. DOI: 10.2307/1999343.

[CDL07] Alessandra Cutrì and Francesca Da Lio. "Comparison and existence results for evolutive non-coercive first-order Hamilton-Jacobi equations." In: *ESAIM: Control, Optimisation and Calculus of Variations* 13.3 (2007), 484?502. DOI: 10.1051/cocv:2007021.

[Dei92] Klaus Deimling. *Multivalued differential equations*. Vol. 1. De Gruyter Series in Nonlinear Analysis and Applications. Walter de Gruyter & Co., Berlin, 1992, pp. xii+260. ISBN: 3-11-013212-5. DOI: 10.1515/9783110874228.

[DCK23] Serena Della Corte and Richard C. Kraaij. *Well-posedness of a Hamilton-Jacobi-Bellman equation in the strong coupling regime*. 2023. arXiv: 2310.05659 [math.AP].

[DK24] Serena Della Corte and Richard C. Kraaij. "Large deviations for Markov processes with switching and homogenisation via Hamilton–Jacobi–Bellman equations." In: *Stochastic Processes and their Applications* (2024), p. 104301. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2024.104301. URL: https://www.sciencedirect.com/science/article/pii/S0304414924000073.

[DFL11] Xiaoxue Deng, Jin Feng, and Yong Liu. "A singular 1-D Hamilton-Jacobi equation, with application to large deviation of diffusions." In: *Communications in Mathematical Sciences* 9.1 (2011).

[DKN21] Robert Denk, Michael Kupper, and Max Nendel. "Convex semigroups on L^p-like spaces." In: *J. Evol. Equ.* 21.2 (2021), pp. 2491–2521. ISSN: 1424-3199,1424-3202. DOI: 10.1007/s00028-021-00693-3. URL: https://doi.org/10.1007/s00028-021-00693-3.

[DV75] M. D. Donsker and S. R. S. Varadhan. "On a Variational Formula for the Principal Eigenvalue for Operators with Maximum Principle." In: *Proceedings of the National Academy of Sciences* 72.3 (1975), pp. 780–783. ISSN: 0027-8424. DOI: 10.1073/pnas.72.3.780. eprint: https://www.pnas.org/content/72/3/780.full.pdf. URL: https://www.pnas.org/content/72/3/780.

[DIS90] Paul Dupuis, Hitoshi Ishii, and H. Mete Soner. "A Viscosity Solution Approach to the Asymptotic Analysis of Queueing Systems." In: *The Annals of Probability* 18.1 (1990), pp. 226–255. ISSN: 00911798.

[Ell85] Richard S. Ellis. *Entropy, large deviations, and statistical mechanics*. New York: Springer-Verlag, 1985.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag, 2000.

[EK09] S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2009. ISBN: 9780470317327. URL: https://books.google.nl/books?id=zvE9RFouKoMC.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and Convergence*. Wiley, 1986, pp. x+534. ISBN: 0-471-08186-8. DOI: 10.1002/9780470316658. URL: https://doi.org/10.1002/9780470316658.

[EI85] Lawrence C Evans and Hitoshi Ishii. "A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities." In: 2.1 (1985), pp. 1–20.

[FGS17] Giorgio Fabbri, Fausto Gozzi, and Andrzej Święch. *Stochastic optimal control in infinite dimension*. Vol. 82. Probability Theory and Stochastic Modelling. Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore. Springer, Cham, 2017, pp. xxiii+916. ISBN: 978-3-319-53066-6; 978-3-319-53067-3. DOI: 10.1007/978-3-319-53067-3. URL: https://doi.org/10.1007/978-3-319-53067-3.

[Fen06] Jin Feng. "Large deviation for diffusions and Hamilton-Jacobi equation in Hilbert spaces." In: *Ann. Probab.* 34.1 (2006), pp. 321–385. ISSN: 0091-1798. DOI: 10.1214/009117905000000567. URL: https://doi.org/10.1214/009117905000000567.

[FK09] Jin Feng and Markos Katsoulakis. "A Comparison Principle for Hamilton-Jacobi Equations Related to Controlled Gradient Flows in Infinite Dimensions." In: *Archive for Rational Mechanics and Analysis* 192.2 (2009), pp. 275–310. ISSN: 1432-0673. DOI: 10.1007/s00205-008-0133-5.

[FK06] Jin Feng and Thomas G. Kurtz. *Large Deviations for Stochastic Processes*. American Mathematical Society, 2006, pp. xii+410. DOI: 10.1090/surv/131. URL: https://doi.org/10.1090/surv/131.

[FMZ21] Jin Feng, Toshio Mikami, and Johannes Zimmer. "A Hamilton-Jacobi PDE associated with hydrodynamic fluctuations from a nonlinear diffusion equation." In: *Comm. Math. Phys.* 385.1 (2021), pp. 1–54. ISSN: 0010-3616. DOI: 10.1007/s00220-021-04110-1. URL: https://doi.org/10.1007/s00220-021-04110-1.

[FG10] Alessio Figalli and Nicola Gigli. "A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions." In: *Journal de Mathématiques Pures et Appliquées* 94.2 (2010), pp. 107–130. ISSN: 0021-7824. DOI: https://doi.org/10.1016/j.matpur.2009.11.005. URL: https://www.sciencedirect.com/science/article/pii/S0021782409001627.

[Fle77] Wendell H. Fleming. "Exit probabilities and optimal stochastic control." In: *Applied Mathematics and Optimization* 4 (1 1977), pp. 329–346. ISSN: 0095-4616. DOI: 10.1007/BF01442148. URL: https://doi.org/10.1007/BF01442148.

[FS06] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions. 2nd ed.* English. 2nd ed. New York, NY: Springer, 2006, pp. xvii + 428. ISBN: 0-387-26045-5/hbk; 978-1-4419-2078-2/pbk; 0-387-31071-1/ebook. DOI: 10.1007/0-387-31071-1.

[FS86] Wendell H. Fleming and Panagiotis E. Souganidis. "PDE-viscosity solution approach to some problems of large deviations." In: *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* Ser. 4, 13.2 (1986), pp. 171–192. ISSN: 0391-173X.

[FW79] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*. second. Springer-Verlag, 1979.

[FW98] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*. second. Springer-Verlag, 1998.

[HUL01] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. ISBN: 3-540-42205-6.

[Hol00] F. Hollander. *Large Deviations*. Fields Institute monographs. American Mathematical Society, 2000. ISBN: 9780821844359. URL: https://books.google.it/books?id=arxAjD_yl4oC.

[Hol16] Julian Hollender. "Lévy-Type Processes under Uncertainty and Related Nonlocal Equations." PhD thesis. Technical University of Dresden, 2016.

[H03] Lars Hörmander. *The analysis of linear partial differential operators. I*. Classics in Mathematics. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)]. Springer-Verlag, Berlin, 2003, pp. x+440. ISBN: 3-540-00662-1. DOI: 10.1007/978-3-642-61497-2. URL: https://doi.org/10.1007/978-3-642-61497-2.

[HP21] M. Hu and S. Peng. "G-Lévy processes under sublinear expectations." In: *Probability, Uncertainty and Quantitative Risk* 6.1 (2021), pp. 1–22.

[IL90] H Ishii and P.L Lions. "Viscosity solutions of fully nonlinear second-order elliptic partial differential equations." In: *Journal of Differential Equations* 83.1 (1990), pp. 26–78. ISSN: 0022-0396. DOI: https://doi.org/10.1016/0022-0396(90)90068-Z. URL: https://www.sciencedirect.com/science/article/pii/002203969090068Z.

[Ish84] Hitoshi Ishii. "Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations." In: *Indiana Univ. Math. J.* 33.5 (1984), pp. 721–748. ISSN: 0022-2518,1943-5258. DOI: 10.1512/iumj.1984.33.33038. URL: https://doi.org/10.1512/iumj.1984.33.33038.

[Ish86] Hitoshi Ishii. "Existence and uniqueness of solutions of Hamilton-Jacobi equations." In: *Funkcial. Ekvac.* 29.2 (1986), pp. 167–188. ISSN: 0532-8721. URL: http://www.math.kobe-u.ac.jp/~fe/xml/mr0877427.xml.

[Ish87] Hitoshi Ishii. "Perron's method for Hamilton-Jacobi equations." In: *Duke math. J* 55.2 (1987), pp. 369–384.

[JLS88] R. Jensen, P-L. Lions, and P. E. Souganidis. “A Uniqueness Result for Viscosity Solutions of Second Order Fully Nonlinear Partial Differential Equations.” In: *Proceedings of the American Mathematical Society* 102.4 (1988), pp. 975–978. URL: http://www.jstor.org/stable/2047344.

[Jen88] Robert Jensen. “The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations.” In: *Archive for Rational Mechanics and Analysis* 101.1 (1988), pp. 1–27.

[JAP97] F. Jülicher, A. Ajdari, and J. Prost. “Modelling molecular motors.” In: *Rev. Mod. Phys.* 69 (4 1997), pp. 1269–1282.

[KK13] Hye-Won Kang and Thomas G. Kurtz. “Separation of time-scales and model reduction for stochastic reaction networks.” In: *The Annals of Applied Probability* 23.2 (2013), pp. 529–583.

[Kle14] Achim Klenke. *Probability theory*. Second. Universitext. A comprehensive course. Springer, London, 2014, pp. xii+638. DOI: 10.1007.

[Kol13] A. B. Kolomeisky. “Motor proteins and molecular motors: How to operate machines at the nanoscale.” In: *Journal of Physics: Condensed Matter* (2013).

[KF07] A. B. Kolomeisky and M. E. Fisher. “Molecular Motors: A Theorist’s Perspective.” In: *Annual Review of Physical Chemistry* 58.1 (2007), pp. 675–695.

[Kra20] Richard C. Kraaij. “The exponential resolvent of a Markov process and large deviations for Markov processes via Hamilton-Jacobi equations.” In: *Electron. J. Probab.* 25 (2020), 39 pp. DOI: 10.1214/20-EJP539. URL: https://doi.org/10.1214/20-EJP539.

[Kra22] Richard C. Kraaij. “A general convergence result for viscosity solutions of Hamilton-Jacobi equations and non-linear semigroups.” In: *Journal of Functional Analysis* 282.5 (2022), p. 109346. ISSN: 0022-1236. DOI: https://doi.org/10.1016/j.jfa.2021.109346. URL: https://www.sciencedirect.com/science/article/pii/S0022123621004286.

[KM20] Richard C. Kraaij and Louis Mahé. “Well-posedness of Hamilton-Jacobi equations in population dynamics and applications to large deviations.” In: *Stochastic Processes and their Applications* (2020). ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2020.03.013. URL: http://www.sciencedirect.com/science/article/pii/S0304414919300821.

[KRV19] Richard C. Kraaij, Frank Redig, and Rik Versendaal. “Classical large deviation theorems on complete Riemannian manifolds.” In: *Stochastic Processes and their Applications* 129.11 (2019), pp. 4294 –4334. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2018.11.019. URL: http://www.sciencedirect.com/science/article/pii/S0304414918306781.

[KS20] Richard C. Kraaij and Mikola C. Schlottke. “A large deviation principle for Markovian slow-fast systems.” In: *preprint; ArXiv:2011.05686* (2020). arXiv: 2011.05686.

[KS21] Richard C. Kraaij and Mikola C. Schlottke. "Comparison Principle for Hamilton - Jacobi - Bellman Equations via a Bootstrapping Procedure." In: *Nonlinear Differential Equations and Applications NoDEA* 28.2 (2021), p. 22. DOI: 10.1007/s00030-021-00680-0. URL: https://doi.org/10.1007/s00030-021-00680-0.

[KP17] Rohini Kumar and Lea Popovic. "Large deviations for multi-scale jump-diffusion processes." In: *Stochastic Processes and their Applications* 127.4 (2017), pp. 1297 –1320. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2016.07.016. URL: https://www.sciencedirect.com/science/article/pii/S0304414916301284.

[Lin92] Torgny Lindvall. *Lectures on the coupling method*. Dover Publications, 1992.

[Lio84] P.-L. Lions. "Solutions de viscosité des équations de Hamilton-Jacobi du premier ordre et applications." In: *Goulaouic-Meyer-Schwartz seminar, 1983–1984*. École Polytech., Palaiseau, 1984, Exp. No. 6, 13. ISBN: 2-7302-0077-0.

[Lip96] Robert Liptser. "Large deviations for two scaled diffusions." In: *Probability Theory and Related Fields* 106.1 (Sept. 1996), pp. 71–104. ISSN: 1432-2064. URL: https://doi.org/10.1007/s004400050058.

[Nen24] Max Nendel. "Lower semicontinuity of monotone functionals in the mixed topology on C_b." In: (2024). arXiv: 2210.09133 [q-fin.MF]. URL: https://arxiv.org/abs/2210.09133.

[PS19] Mark A. Peletier and Mikola C. Schlottke. *Large-deviation principles of switching Markov processes via Hamilton-Jacobi equations*. 2019. eprint: 1901.08478.

[PS07] Benoit Perthame and Panagiotis Souganidis. "Asymmetric Potentials and Motor Effect: A Large Deviation Approach." In: *Archive for Rational Mechanics and Analysis* 193 (June 2007).

[Pin07] Ross G. Pinsky. "Regularity properties of the Donsker-Varadhan rate functional for non-reversible diffusions and random evolutions." In: *Stoch. Dyn.* 7.2 (2007), pp. 123–140. ISSN: 0219-4937. DOI: 10.1142/S0219493707001998.

[Pop18] Lea Popovic. "Large deviations of Markov chains with multiple time-scales." In: *Stochastic Processes and their Applications* (2018). ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2018.09.009. URL: http://www.sciencedirect.com/science/article/pii/S0304414918300942.

[Roc70] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970, pp. xviii–451.

[Sen72] F. Dennis Sentilles. "Bounded continuous functions on a completely regular space." In: *Trans. Amer. Math. Soc.* 168 (1972), pp. 311–336. ISSN: 0002-9947. DOI: 10.2307/1996178. URL: https://doi.org/10.2307/1996178.

[SW05] Adam Shwartz and Alan Weiss. "Large Deviations with Diminishing Rates." In: *Mathematics of Operations Research* 30.2 (2005), pp. 281–310. ISSN: 0364765X, 15265471. URL: http://www.jstor.org/stable/25151653.

[Son86] Halil Mete Soner. "Optimal control with state-space constraint. II." In: *SIAM J. Control Optim.* 24.6 (1986), pp. 1110–1122. ISSN: 0363-0129. DOI: 10.1137/0324067. URL: https://doi.org/10.1137/0324067.

[Spi70] Michael Spivak. *A comprehensive introduction to differential geometry. Vol. One*. Published by M. Spivak. 1970, iii+656 pp. (not consecutively paged) paperbound.

[SV79] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Vol. 233. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin-New York, 1979, pp. xii+338. ISBN: 3-540-90353-4.

[Swe92] Guido Sweers. "Strong positivity in $C(\overline{\Omega})$ for elliptic systems." In: *Math. Z.* 209.2 (1992), pp. 251–271. ISSN: 0025-5874. DOI: 10.1007/BF02570833. URL: https://doi.org/10.1007/BF02570833.

[Tat92] Daniel Tataru. "Viscosity solutions of Hamilton-Jacobi equations with unbounded nonlinear terms." In: *J. Math. Anal. Appl.* 163.2 (1992), pp. 345–392. ISSN: 0022-247X. DOI: 10.1016/0022-247X(92)90256-D. URL: https://doi.org/10.1016/0022-247X(92)90256-D.

[Tat94] Daniel Tataru. "Viscosity solutions for Hamilton-Jacobi equations with unbounded nonlinear term: a simplified approach." In: *J. Differential Equations* 111.1 (1994), pp. 123–146. ISSN: 0022-0396. DOI: 10.1006/jdeq.1994.1078. URL: https://doi.org/10.1006/jdeq.1994.1078.

[Tho00] Hermann Thorisson. *Coupling, stationarity, and regeneration*. Springer-Verlag, 2000.

[Tu10] Loring W. Tu. *An Introduction to Manifolds*. Springer New York, NY, 2010.

[Var66] S.R.S Varadhan. "Asymptotic Probabilities and Differential Equations." In: *Communications on Pure and Applied Mathematics* 19.3 (1966), pp. 261–286.

[Whi34] Hassler Whitney. "Analytic extensions of differentiable functions defined in closed sets." In: *Trans. Amer. Math. Soc.* 36.1 (1934), pp. 63–89. ISSN: 0002-9947,1088-6850. DOI: 10.2307/1989708. URL: https://doi.org/10.2307/1989708.

[YFRH02] (Author) YC. Fung and (Reviewer) RL. Huston. "Introduction to Bioengineering." In: *Applied Mechanics Reviews* 55.2 (Apr. 2002), B39–B39. ISSN: 0003-6900. DOI: 10.1115/1.1451233. URL: https://doi.org/10.1115/1.1451233.

[YZ10] G. George Yin and Chao Zhu. *Hybrid Switching Diffusions*. Vol. 63. Springer New York, 2010. ISBN: 978-1-4419-1106-3. DOI: 10.1007/978-1-4419-1105-6.

[YZ99] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls*. Vol. 43. Applications of Mathematics (New York). Hamiltonian systems and HJB equations. Springer-Verlag, New York, 1999, pp. xxii+438. ISBN: 0-387-98723-1. DOI: 10.1007/978-1-4612-1466-3. URL: https://doi.org/10.1007/978-1-4612-1466-3.

SUMMARY

This thesis focuses on two main topics: large deviations for Markov processes and the well-posedness of Hamilton–Jacobi equations.

The first two chapters provide an introduction to both areas. Chapter 1 explores the mathematical foundations of Hamilton–Jacobi equations, highlighting their applications in control theory and emphasizing the role of viscosity solutions in handling situations where classical solutions fail. Chapter 2 introduces large deviations theory, starting from basic examples and leading to rigorous definitions. A key theme is the connection between large deviations and Hamilton–Jacobi equations, introduced through the Feng–Kurtz method.

The subsequent chapters present the main research contributions of this thesis. Chapter 3 studies two examples of two-scale Markov processes and applies the Feng–Kurtz method to establish a large deviations principle.

The first example originates from molecular biology and models the movement of a motor protein along filaments. We consider a two-component stochastic process (X_t, I_t), where X_t represents the protein's spatial position and I_t encodes its molecular configuration. The dynamics follow a stochastic differential equation:

$$\mathrm{d}X_t = -\nabla\psi(X_t, I_t)\mathrm{d}t + \mathrm{d}B_t,$$
$$\mathbb{P}\Big(I(t+\Delta t) = j \,|\, I(t) = i, X(t) = x\Big) = r_{ij}(x)\,\Delta t + \mathcal{O}(\Delta t^2) \quad \text{as } \Delta t \to 0.$$

After rescaling the process in space and time by a small parameter $\varepsilon > 0$, we analyze its behavior as $\varepsilon \to 0$. Using the Feng–Kurtz method, we prove a large deviations principle, which requires establishing a comparison principle for viscosity solutions of a spatially inhomogeneous Hamilton–Jacobi equation.

The second example, inspired by biochemical reaction networks, involves a two-component process $Z = (X, Y)$, where X and Y represent two molecular species. Assuming that the number of X-molecules is significantly larger than that of Y-molecules, we introduce a rescaled process $Z_N = (X_N, Y_N)$ with a time-scale separation of order N. Again, applying the Feng–Kurtz method, we derive a large deviations principle for X_N and prove a comparison principle for the associated Hamilton–Jacobi equation.

Chapter 4 transitions to the second theme of this thesis: the well-posedness of Hamilton–Jacobi equations. Motivated by the previous examples, we analyze a general class of Hamilton–Jacobi equations of the form

$$\mathcal{H}(x,p) = \sup_{\theta\in\Theta} \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta).$$

This Hamiltonian structure encompasses those arising in the large deviations problems studied earlier. We establish a comparison principle for viscosity solutions, demonstrating its applicability in a broad setting.

Chapter 5 extends these results by proving the existence of viscosity solutions for a general class of Hamilton–Jacobi equations using Lyapunov control techniques. Unlike traditional approaches, our method relaxes common assumptions such as convexity and coercivity. This allows us to address Hamilton–Jacobi–Isaacs equations, where the Hamiltonian takes a "sup-inf" or "inf-sup" form.

In the final chapter, Chapter 6, we investigate second-order Hamilton–Jacobi equations, presenting a novel proof of the comparison principle for viscosity solutions. We also examine specific cases, including Hamiltonians with a partial integro-differential structure:

$$\begin{aligned} Hf(x) &= \langle b(x), \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\Sigma\Sigma^T(x) D^2 f(x)\right) \\ &\qquad + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z})\,\langle \mathbf{z}, \nabla f(x)\rangle\right] \mu_x(\mathrm{d}\mathbf{z}) + \mathcal{H}(\nabla f(x)). \end{aligned}$$

This type of operators appear in applications such as financial mathematics (e.g., option pricing and stock price modeling). Additionally, we discuss Hamilton–Jacobi–Bellman and Isaacs equations in this context.

Each chapter concludes with remarks and future research directions.

SAMENVATTING (DUTCH SUMMARY)

Dit proefschrift richt zich op twee hoofdonderwerpen: grote afwijkingen voor Markov-processen en de goed-gesteldheid van Hamilton–Jacobi-vergelijkingen.

De eerste twee hoofdstukken geven een inleiding tot beide onderwerpen. Hoofdstuk 1 behandelt de wiskundige basis van Hamilton–Jacobi-vergelijkingen, met de nadruk op hun toepassingen in de regeltheorie en de rol van viscositeitsoplossingen bij problemen waar klassieke oplossingen tekortschieten. Hoofdstuk 2 introduceert de theorie van grote afwijkingen, beginnend met eenvoudige voorbeelden en geleidelijk overgaand naar rigoureuze definities. Een centrale rol is weggelegd voor de verbinding tussen grote afwijkingen en Hamilton–Jacobi-vergelijkingen, geïntroduceerd via de methode van Feng en Kurtz.

De daaropvolgende hoofdstukken presenteren de belangrijkste onderzoeksbijdragen van dit proefschrift. Hoofdstuk 3 onderzoekt twee voorbeelden van Markov-processen met twee tijdschalen en past de methode van Feng en Kurtz toe om een principe van grote afwijkingen af te leiden.

Het eerste voorbeeld komt voort uit de moleculaire biologie en modelleert de beweging van een motorproteïne langs filamenten. We beschouwen een stochastisch proces met twee componenten (X_t, I_t), waarbij X_t de ruimtelijke positie van het eiwit vertegenwoordigt en I_t de moleculaire configuratie aangeeft. De dynamica worden beschreven door de volgende stochastische differentiaalvergelijking:

$$\begin{aligned} &\mathrm{d}X_t = -\nabla\psi(X_t, I_t)\mathrm{d}t + \mathrm{d}B_t, \\ &\mathbb{P}\Big(I(t+\Delta t) = j \,|\, I(t) = i, X(t) = x\Big) = r_{ij}\,(x)\,\Delta t + \mathcal{O}(\Delta t^2) \quad \text{als } \Delta t \to 0. \end{aligned} \tag{.0.1}$$

Door het proces te herschalen in ruimte en tijd met een kleine parameter $\varepsilon > 0$, bestuderen we het limietgedrag als $\varepsilon \to 0$. Met behulp van de methode van Feng en Kurtz bewijzen we een principe van grote afwijkingen, wat vereist dat we een vergelijkingsprincipe voor viscositeitsoplossingen van een ruimtelijk inhomogene Hamilton–Jacobi-vergelijking vaststellen.

Het tweede voorbeeld, geïnspireerd door biochemische reactienetwerken, beschouwt een tweedelig proces $Z = (X, Y)$, waarbij X en Y twee moleculaire soorten vertegenwoordigen. Aangezien het aantal X-moleculen veel groter is dan dat van Y-moleculen, introduceren we een herschaald proces $Z_N = (X_N, Y_N)$, waarbij de tijdschaalscheiding van orde N is. Opnieuw passen we de methode van Feng en Kurtz toe en leiden we een principe van grote afwijkingen af voor X_N, waarbij we opnieuw een vergelijkingsprincipe voor de bijbehorende Hamilton–Jacobi-vergelijking bewijzen.

Hoofdstuk 4 richt zich op het tweede hoofdthema van dit proefschrift: de goed-gesteldheid van Hamilton–Jacobi-vergelijkingen. Gemotiveerd door de eerdere voorbeelden analyseren we een algemene klasse Hamilton–Jacobi-vergelijkingen van de vorm

$$\mathcal{H}(x,p) = \sup_{\theta \in \Theta} \Lambda(x,p,\theta) - \mathcal{I}(x,p,\theta).$$

Deze Hamiltoniaan omvat ook de structuren die in de eerder bestudeerde problemen van grote afwijkingen voorkomen. We bewijzen een vergelijkingsprincipe voor viscositeitsoplossingen en tonen de toepasbaarheid ervan in een brede context aan.

Hoofdstuk 5 breidt deze resultaten uit door het bestaan van viscositeitsoplossingen voor een algemene klasse van Hamilton–Jacobi-vergelijkingen te bewijzen met behulp van Lyapunov-controletechnieken. In tegenstelling tot traditionele benaderingen versoepelt onze methode veelgebruikte aannames zoals convexiteit en coerciviteit. Dit stelt ons in staat om ook Hamilton–Jacobi–Isaacs-vergelijkingen te behandelen, waarbij de Hamiltoniaan een "sup-inf" of "inf-sup" vorm heeft.

In het laatste hoofdstuk, Hoofdstuk 6, onderzoeken we tweede-orde Hamilton–Jacobi-vergelijkingen en presenteren we een nieuwe bewijsstrategie voor het vergelijkingsprincipe voor viscositeitsoplossingen. We analyseren ook specifieke gevallen, waaronder Hamiltonianen met een partieel integro-differentiële structuur:

$$\begin{aligned} Hf(x) &= \langle b(x), \nabla f(x) \rangle + \frac{1}{2} \operatorname{Tr}\left(\Sigma\Sigma^T(x) D^2 f(x) \right) \\ &\quad + \int \left[f(x+\mathbf{z}) - f(x) - \chi_{B_1(0)}(\mathbf{z}) \langle \mathbf{z}, \nabla f(x) \rangle \right] \mu_x(\mathrm{d}\mathbf{z}) + \mathcal{H}(\nabla f(x)). \end{aligned}$$

Dit type operatoren komt voor in toepassingen zoals financiële wiskunde (bijvoorbeeld optieprijsbepaling en aandelenkoersmodellen). Daarnaast bespreken we in deze context ook Hamilton–Jacobi–Bellman- en Isaacs-vergelijkingen.

Elk hoofdstuk sluit af met opmerkingen en toekomstige onderzoeksperspectieven.

PUBLICATIONS

Included in this thesis:

[1] S. Della Corte and R. C. Kraaij, *Large deviations for Markov processes with switching and homogenisation via Hamilton–Jacobi–Bellman equations.* In: Stochastic Processes and their Applications. (2024)

[2] S. Della Corte and R. C. Kraaij, *Well–posedness of Hamilton–Jacobi–Bellman equations in the strong coupling regime. Preprint. (2023)*

[3] S. Della Corte and R. C. Kraaij, *Existence of viscosity solutions for Hamilton-Jacobi equations on Riemannian Manifolds via Lyapunov control. Preprint. (2024)*

[4] S. Della Corte, F. Fuchs, R. C. Kraaij and M. Nendel, *A comparison principle for Hamiton-Jacobi-Bellman and Isaacs equations based on couplings of integro–differential operators. Preprint. (2024)*

Not included in this thesis:

[1] S. Della Corte, A. Diana and C. Mantegazza, *Global Existence and Stability for the Modified Mullins-Sekerka and Surface Diffusion Flow.* In: Mathematics in Engineering. (2022).

[2] S. Della Corte, A. Diana and C. Mantegazza, *Uniform Sobolev, interpolation and geometric Calderón-Zygmund inequalities for graph hypersurfaces.* In: Note di matematica. (2023)