ESAIM: COCV 32 (2026) 5
https://doi.org/10.1051/cocv/2025092

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

DYNAMIC PROGRAMMING PRINCIPLE AND HAMILTON–JACOBI–BELLMAN EQUATION FOR OPTIMAL CONTROL PROBLEMS WITH UNCERTAINTY

Maria Soledad Aronna[1,*], Michele Palladino[2] and
Oscar Sierra[1]

Abstract. We study the properties of the value function associated with an optimal control problem with uncertainties, known as *average* or *Riemann–Stieltjes* problem. Uncertainties are assumed to belong to a compact metric probability space, and appear in the dynamics, in the terminal cost and in the initial condition, which yield an infinite-dimensional formulation. By stating the problem as an evolution equation in a Hilbert space, we show that the value function is the unique lower semi-continuous proximal solution of the Hamilton–Jacobi–Bellman (HJB) equation. Our approach relies on invariance properties and the dynamic programming principle.

Mathematics Subject Classification. 35F21, 49K45, 49K27, 49L25.

Received July 17, 2024. Accepted November 23, 2025.

1. Introduction

This work studies the properties of the value function and the Hamilton–Jacobi–Bellman equation related to a class of optimal control problems that consider uncertainty in the dynamics and initial state. From a practical point of view, the framework covers models with stochasticity in the parameter values and in the initial conditions. The uncertain values are assumed to belong to a probability space, and the cost involves the average of a terminal cost computed among the corresponding probability measure. This framework is known in the literature as *average optimal control* [1, 2] or *Riemman-Stieltjes* problems [3–5].

The main contribution of the work is providing a theoretical framework by stating the problem in a (possibly infinite dimensional) Hilbert space for the time dependent state variable and proving that the associated value function is the unique lower semi-continuous solution of a Hamilton–Jacobi–Bellman (HJB) equation, under mild regularity assumptions on the governing dynamics (it is required to be merely measurable in time). More specifically, we study the following parametrized Riemann–Stieltjes problem with respect to the initial time $s \in [0,T]$ and the initial state, the latter being an L^2-function of the parameter ω, *i.e.,* the initial condition is

---

Keywords and phrases: Optimal control, Hamilton–Jacobi–Bellman equation, dynamic programming principle, invariance principles, Riemann–Stieltjes optimal control problems, uncertain dynamics.

[1] Escola de Matemática Aplicada, FGV EMAp, Praia de Botafogo 190, 22250-900 Rio de Janeiro, RJ, Brazil.
[2] University of L'Aquila – Department of Information Engineering, Computer Science and Mathematics (DISIM), via Vetoio, 67100, L'Aquila, Italy.

[*] Corresponding author: `soledad.aronna@fgv.br`

represented by elements φ in $L^2(\mu,\Omega;\mathbb{R}^n)$:

$$(P)_{s,\varphi} \qquad \begin{aligned} &\min_{u} \int_\Omega g(x(T,\omega),\omega)\,\mathrm{d}\mu(\omega),\\ &\text{s.t.}\\ &\begin{cases} \dot{x}(t,\omega) = f(t,x(t,\omega),u(t),\omega), & \text{a.e. } t\in[s,T],\ \omega\in\Omega,\\ x(s,\omega)=\varphi(\omega), & \omega\in\Omega,\\ u(t)\in U(t), & \text{a.e. } t\in[s,T]. \end{cases} \end{aligned} \tag{1.1}$$

Here $u:[s,T]\to\mathbb{R}^m$ is the control function, which belongs to the *admissible control set*

$$\mathcal{U}_{\mathrm{ad}}[s,T] := \left\{u\in L^\infty([s,T];\mathbb{R}^m) : u(t)\in U(t) \quad \text{a.e. } t\in[s,T]\right\},$$

and the triple (Ω, d_Ω, μ) is a metric measure space.

This type of problem has been addressed using various approaches, including different theoretical frameworks and computational techniques. In [6], the authors propose models from aerospace engineering in the Lebesgue–Stieltjes framework, and in [5], they formally derive a Pontryagin Maximum Principle for that class of average problems. In [7], the authors provide a computational framework for this uncertain dynamic optimization problem, providing consistent approximation techniques. The authors in [8] deal with the linear quadratic case, showing convergence results in model-based reinforcement learning scenarios, while the authors in [9] deal with a related problem with control-affine dynamics and quadratic cost. In [10], this class of problems is referred to as *ensemble optimal control*. The latter work establishes convergence results and numerical algorithms for their solution. It is worth highlighting that our framework involves an average cost, while other authors have studied the worst-case scenario formulation (see, *e.g.,* [11] for *minimax optimal control*).

The problem $(P)_{s,\varphi}$ was also studied in [3] when the initial condition is constant-valued, *i.e.,* $\varphi\equiv x_0\in\mathbb{R}^n$, and where an additional averaged right end-point constraint is considered. The authors in [3] state necessary conditions for optimality by considering, for each control u, the set of pointwise trajectories $\{x(\cdot,\omega)\in W^{1,1}([0,T]:\mathbb{R}^n):\omega\in\Omega\}$ satisfying the constraints. Here we extend the set of constraints to an ODE defined in a Hilbert space, that is, given an initial condition $\varphi\in L^2(\mu,\Omega;\mathbb{R}^n)$ and a control variable $u\in\mathcal{U}_{\mathrm{ad}}[s,T]$, the corresponding feasible trajectory $x(\cdot,\cdot)$ belongs to $C([s,T]:L^2(\mu,\Omega;\mathbb{R}^n))$ and satisfies (1.1).

The *value function* for the Riemann–Stieltjes problem $(P)_{s,\varphi}$ is the extended-valued mapping given by

$$\begin{aligned} V\colon [0,T]\times L^2(\mu,\Omega;\mathbb{R}^n) &\to \mathbb{R}\cup\{\infty\}\\ (s,\varphi) &\mapsto V(s,\varphi) := \inf_{u\in\mathcal{U}_{\mathrm{ad}}[s,T]} (P)_{s,\varphi}. \end{aligned}$$

When V is Fréchet differentiable with respect to (t,φ) in $(0,T)\times L^2(\mu,\Omega;\mathbb{R}^n)$, it satisfies the HJB equation

$$\begin{cases} -\left[V_t(t,\varphi) + H(t,\varphi,V_\varphi(t,\varphi))\right] = 0,\\[2ex] V(T,\varphi) = \displaystyle\int_\Omega g\left(\varphi(\omega),\omega\right)\mathrm{d}\mu(\omega), \end{cases} \tag{1.2}$$

where the Hamiltonian, $H:[0,T]\times L^2(\mu,\Omega;\mathbb{R}^n)\times L^2(\mu,\Omega;\mathbb{R}^n)^*\to\mathbb{R}$ is given by

$$H(t,\varphi,p) := \inf_{u\in U(t)} \left\langle p(\cdot), f(t,\varphi(\cdot),u,\cdot)\right\rangle, \tag{1.3}$$

and $(V_t,V_\varphi)\in(\mathbb{R}\times L^2(\mu,\Omega;\mathbb{R}^n))^*$ denotes the Fréchet derivative of V with respect to $(t,\varphi)\in(0,T)\times L^2(\mu,\Omega;\mathbb{R}^n)$ Here, we deal with non-smooth solutions of the HJB equation, *i.e.,* we consider (1.2) for V merely lower semi-continuous (l.s.c.) and we adopt the notions of *proximal subdifferential* (denoted by $\partial_P V$) and *proximal solutions* [12]. To prove that V is the unique l.s.c. proximal solution of (1.2), we rely on results of *invariance* from [13]. The latter work consists of an extension to the infinite-dimensional case of the work [14] by Clarke *et al.*, where the HJB equation is related to the concepts of weak and strong invariance (often called *viability* and invariance, respectively [15]).

Problems where the state space is finite-dimensional and the cost function is merely lower semi-continuous have been extensively studied. For instance, Frankowska in [16] addresses the case when the dynamics is continuous in time; later in [17], the result is extended to measurable-in-time dynamics. In both cases, the approach is to prove the invariance properties of an associated autonomous system along with the well-known *Dynamic Programming Principle.* In [17], the authors show that the non-smooth solutions of the HJB equation belong to the subdifferential of the value function and are characterized by means of the *contingent cone* [15, 18] to the epigraph of the value function. Such a characterization fails when the space is infinite dimensional (see [18], Prop. 6.4.8). In our framework, we use analogous objects from the proximal analysis known as *proximal subdifferentials* [19] (see also Sect. 4), which are characterized by the elements of the *normal proximal cone* to the epigraph of the value function; these gradients will be involved in the concept of *proximal solution* of the HJB equation. The results of existence and uniqueness of the HJB equation provided in [16, 17] rely on the existence of a minimum of the cost function, which requires the compactness of the set of trajectories. In general, this property does not hold for infinite-dimensional spaces. However, in the semilinear case, when defined on separable Banach spaces, the lack of compactness of the set of trajectories can be overcome by a compactness assumption on the semigroup related to the differential operator, and by applying the Arzelà-Ascoli Theorem [20–22].

In the present work, the cost is given by an integral functional in which the lower semi-continuity is inherited by the function g, and the dynamics are purely nonlinear. The compactness of the set of trajectories is addressed by introducing additional assumptions on the measure μ and ensuring some regularity properties for the mapping $\omega \mapsto \int_s^t f(\sigma, \varphi(\omega), u(\sigma), \omega)\mathrm{d}\sigma$. Additionally, we employ a Rellich–Kondrakov-type theorem [23, 24].

The paper is outlined as follows. Section 2 includes properties on the behavior of the trajectories and the statement of the Dynamic Programming Principle. Section 3 is dedicated to the study of the compactness properties of the set of trajectories, characterizing the lower semi-continuity of the value function, and showing the existence of minimizers. Section 4 is devoted to proving that the value function is the unique proximal solution of the HJB equation. Finally, the paper ends with an appendix, presenting some technical results needed in our proofs.

1.1. Notations, basic assumptions and definitions

Throughout the paper, we impose the following set of hypotheses, which we refer to hereafter with **(H)**:

(i) (Ω, d_Ω, μ) is a compact metric measure space, μ being a finite Radon measure.

(ii) $U : [0,T] \rightsquigarrow \mathbf{U}$ is an upper semi-continuous set-valued function taking nonempty closed values, included in a compact set $\mathbf{U} \subset \mathbb{R}^m$.

(iii) $f\colon [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$ is measurable w.r.t. t and continuous on the other variables, and there exist constants $c, k_f > 0$ such that, for all $x \in \mathbb{R}^n, u \in \mathbf{U}$, $\omega \in \Omega$, a.e. $t \in (0,T)$, f satisfies

$$|f(t,x,u,\omega)| \le c(1+|x|), \tag{1.4}$$

and

$$|f(t,x,u,\omega) - f(t,x',u,\omega)| \le k_f|x-x'|. \tag{1.5}$$

(iv) The extended-valued map $g : \mathbb{R}^n \times \Omega \to \mathbb{R}\cup\{\infty\}$ is such that

$$g(x,\omega) \ge a(\omega) - b|x|^2, \quad \text{for every } \ x \in \mathbb{R}^n \ \text{ and a.e. } \omega \in \Omega,$$

for some $a \in L^1(\mu,\Omega) := \{a : \Omega \to \mathbb{R} : \int_\Omega |a(\omega)|\,\mathrm{d}\mu(\omega) < \infty\}$, and $b \geq 0$. Additionally, g satisfies the following two conditions:

- g is an $\mathcal{L} \otimes \mu$ -measurable function, $\mathcal{L}$ here denoting the Lebesgue measure;
- $g(\cdot,\omega)$ is lower semi-continuous for every fixed $\omega \in \Omega$.

The conditions on g in item (iv) above are equivalent to stating that g is a *normal integrand* [25], p. 195.

The values of our state variables will be taken in the space

$$L^2(\mu,\Omega;\mathbb{R}^n) := \Big\{\varphi : \Omega \to \mathbb{R}^n : \int_\Omega |\varphi(\omega)|^2\,\mathrm{d}\mu(\omega) < \infty\Big\},$$

which is a Hilbert space endowed with the scalar product

$$\langle \varphi,\psi\rangle := \int_\Omega \varphi(\omega)\cdot\psi(\omega)\,\mathrm{d}\mu(\omega) \quad \text{for} \quad \varphi,\psi \in L^2(\mu,\Omega;\mathbb{R}^n).$$

Throughout this paper, we will frequently use the notation $\|\cdot\|_{L^2(\mu,\Omega;\mathbb{R}^n)}$ and $\|\cdot\|_{L^2}$ to refer to the norm associated with the scalar product above. For $x \in L^1([0,T])$ its *weak derivative* is a function $v \in L^1([0,T])$ such that $\int_0^T x\gamma'\,\mathrm{d}t = -\int_0^T v\gamma\,\mathrm{d}t$ for all infinitely differentiable functions γ such that $\gamma(0) = \gamma(T) = 0$. For $1 \leq p < \infty$, we denote by $W^{1,p}([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$ the Sobolev space consisting of all the functions $x \in L^p([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$ such that $\dot{x}$ exists in the weak sense and belongs to $L^p([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$. Furthermore,

$$\|x\|^p_{W^{1,p}([s,T];L^2(\mu,\Omega;\mathbb{R}^n))} := \int_s^T \left(\|x(t)\|^p_{L^2} + \|\dot{x}(t)\|^p_{L^2}\right)\,\mathrm{d}t.$$

We have that (see, *e.g.,* [26]), for $x \in W^{1,p}([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$, there exists a continuous representative $y \in C([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$ such that $x(t) = y(t)$ a.e. $t \in [s,T]$. The mapping $t \mapsto x(t)$ from $[s,T]$ to $L^2(\mu,\Omega;\mathbb{R}^n)$ can be regarded as a curve on the Hilbert space $L^2(\mu,\Omega;\mathbb{R}^n)$.

For a metric space X, $\mathcal{B}_r(x)$ represents the open ball of radius $r > 0$ centered at $x \in X$. We call *process* a state-control pair (x,u) composed of an absolutely continuous function $x \in W^{1,1}([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$ and a measurable function $u : [s,T] \to \mathbb{R}^m$, which together satisfy the constraints (1.1) of problem $(P)_{s,\varphi}$. Additionally, x is referred as *trajectory* associated with the *control* u. If there exists a process $(\bar{x},\bar{u})$ solving the problem $(P)_{s,\varphi}$, then it is called *optimal pair*, and we will refer to $\bar{x}$ and $\bar{u}$ as an *optimal trajectory* and *control*, respectively. A trajectory x for the problem $(P)_{s,\varphi}$ can be written in the form of the following evolution integral equation

$$x(t,\cdot) = \varphi(\cdot) + \int_s^t f(\sigma,x(\sigma,\cdot),u(\sigma),\cdot)\,\mathrm{d}\sigma, \quad \text{for a.e. } t \in [s,T]. \tag{1.6}$$

It is clear (see, *e.g.,* [22], Ch. 2, Prop. 5.3) that under hypotheses (**H**), for any $(\varphi,u) \in L^2(\mu,\Omega;\mathbb{R}^n) \times \mathcal{U}_{\mathrm{ad}}[s,T]$, there exists a unique solution $x \in C([s,T];L^2(\mu,\Omega;\mathbb{R}^n))$ of (1.6).

The *reduced cost* of $(P)_{s,\varphi}$ will be defined as

$$J_{[s,T]}(\varphi,u) := \int_\Omega g(x_{\varphi,u}(T,\omega),\omega)\mathrm{d}\mu(\omega),$$

where $x_{\varphi,u}$ is the trajectory corresponding to the initial condition $x(s,\cdot) = \varphi(\cdot)$, and the choice of control u. With these considerations, the value function can be written as

$$V(s,\varphi) := \inf_{u\in\mathcal{U}_{\mathrm{ad}}[s,T]} J_{[s,T]}(\varphi,u) = \inf_{u\in\mathcal{U}_{\mathrm{ad}}[s,T]} \int_\Omega g(x_{\varphi,u}(T,\omega),\omega)\mathrm{d}\mu(\omega). \tag{1.7}$$

2. Dynamic programming principle

In this section, the Dynamic Programming Principle for problem family $(P)_{s,\varphi}$ is stated. This property is the key tool needed to obtain the results of Section 4 and establishes that the value function is nondecreasing along trajectories of (1.1) and constant over optimal trajectories. In other words, optimal trajectories are the characteristic curves for the HJB equation.

The following result embraces the properties of the trajectories of (1.1). Here we will write $x_{s,\varphi}$ for the solution of (1.1) with initial time $s \in [0,T]$ and datum $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$. In what follows, $C(T)$ is a positive constant depending on T, which can vary depending on the context.

Lemma 2.1. *Let the assumptions in* **(H)** *hold. Then, for any s and τ with $0 \le s \le \tau \le T$, $\varphi, \bar{\varphi} \in L^2(\mu,\Omega;\mathbb{R}^n)$, and any control $u \in \mathcal{U}_{\rm ad}[s,T]$, the following inequalities hold*

$$\text{i) } \|x_{s,\varphi}(t)\|_{L^2} \le C(T)\mathrm{e}^{c(t-s)}\left(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2}\right), \quad t \in [s,T];$$

$$\text{ii) } \|x_{s,\varphi}(t) - x_{s,\bar{\varphi}}(t)\|_{L^2} \le \mathrm{e}^{k_f(t-s)}\|\varphi - \bar{\varphi}\|_{L^2}, \quad t \in [s,T];$$

$$\text{iii) } \|x_{\tau,\varphi}(t) - x_{s,\varphi}(t)\|_{L^2} \le C(T)\mathrm{e}^{k_f(t-\tau)}\big(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2}\big)(\tau - s), \quad t \in [\tau,T];$$

$$\text{iv) } \|x_{s,\varphi}(t) - x_{s,\varphi}(\tau)\|_{L^2} \le C(T)\mathrm{e}^{c(t-s)}\big(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2}\big)|t-\tau|, \quad t \in [s,T].$$

Proof. It follows directly from conditions (1.4) and (1.5) in **(H)**, and a standard application of Grönwall's Lemma (see for instance [22], Ch. 6, Lem. 2.1). ☐

The following result holds.

Theorem 2.2 (Dynamic Programming Principle). *Under the hypotheses* **(H)**, *the following conditions hold true:*

i) for every $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$, and every $s_1, s_2 \in [0,T]$, with $s_1 \le s_2$, one has

$$V(s_1,\varphi) \le V(s_2, x_{\varphi,u}(s_2;\cdot)), \tag{2.1}$$

for any measurable control $u \in \mathcal{U}_{\rm ad}[s_1,s_2]$, where $x_{\varphi,u}$ is the trajectory with initial condition $x_{\varphi,u}(s_1,\cdot) = \varphi(\cdot)$ and associated with control $u \in \mathcal{U}_{\rm ad}[s_1,s_2]$;

ii) for every $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$ and every $s_1, s_2 \in [0,T]$, with $s_1 \le s_2$, one has

$$V(s_1,\varphi) = \inf_{u \in \mathcal{U}_{\rm ad}[s_1,T]} V(s_2, x_{\varphi,u}(s_2,\cdot)). \tag{2.2}$$

Moreover, if x^*_{φ,u^*} is an optimal trajectory, from (2.2) it follows that $V(t, x^*_{\varphi,u^*}(t,\cdot))$ is constant for all $t \in [s_1,T]$.

Proof. See Appendix A. ☐

3. Existence of optimal controls

In this section, it is proved that the problem $(P)_{s,\varphi}$ admits at least one solution. The approach involves using differential inclusions to show the compactness of the set of trajectories which, together with the lower semi-continuity of the functional $\mathcal{J} : L^2(\mu,\Omega;\mathbb{R}^n) \to \mathbb{R}\cup\{\infty\}$, given by

$$\mathcal{J}(\varphi) := \int_\Omega g(\varphi(\omega),\omega)\mathrm{d}\mu(\omega), \tag{3.1}$$

guarantees the result. At this point, it is important to mention that, in view of Proposition B.7 in the Appendix, assumption **(H)**(ii) on the regularity properties of the set-valued mapping U ensures its pseudo-continuity (*Definition* B.6). We make use of this property of U in the proof of closedness of the set of trajectories.

Additionally, we consider the set-valued map

$$\begin{aligned} F : [0,T] \times L^2(\mu,\Omega;\mathbb{R}^n) &\rightsquigarrow L^2(\mu,\Omega;\mathbb{R}^n) \\ (t,\varphi) &\mapsto F(t,\varphi(\cdot)) := f(t,\varphi(\cdot),U(t),\cdot) \end{aligned} \tag{3.2}$$

and the associated differential inclusion

$$\dot{x}(t,\cdot) \in F(t,x(t,\cdot)) \quad \text{a.e. } t \in [s,T]. \tag{3.3}$$

Fillipov's Lemma (see Cor. B.9 in the Appendix) allows us to reduce the differential equation

$$\dot{x}(t,\cdot) = f(t,x(t,\cdot),u(t),\cdot), \quad u(t) \in U(t) \quad \text{a.e. } t \in [s,T] \tag{3.4}$$

to the differential inclusion (3.3). More precisely, Fillipov's Lemma states that (3.3) and (3.4) have the same set of trajectories. We will use $S_{[s,T]}(\varphi)$ to denote the set of trajectories of (3.3), with initial data $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$ at time $s \in [0,T]$, *i.e.,*

$$S_{[s,T]}(\varphi) := \left\{ x \in C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)) : \ x \text{ solves } (3.3),\ x(s,\cdot) = \varphi(\cdot) \right\}.$$

3.1. Compactness of the set of trajectories

We are interested in the compactness properties of the set trajectories $S_{[s,T]}(\varphi)$ of (3.3), with fixed initial time s and datum φ. For purely nonlinear dynamics defined in a Banach space, this property can be obtained by assuming an inequality that involves a measure of non-compactness [27], a situation not treated here. On the other hand, for semilinear evolutionary distributed parameter systems in a separable Banach space X, the compactness of such a set is achieved through compactness assumptions either on a C_0-semigroup $\mathcal{T}$ or on the nonlinear part. In [22], p. 110, this result is obtained assuming that the semigroup $\mathcal{T}$ is compact, and in [21, 28], it is achieved when at least one of the following two conditions is satisfied: compactness of $\mathcal{T}$ or inclusion of the nonlinear part in a compact set. When $\mathcal{T}$ is compact or the nonlinear part maps into a compact set, the first step to obtain the compactness of the set of trajectories is proving the same property for the operator

$$\mathcal{S} : L^p([s,T] : X) \to C([s,T] : X), \quad p > 1,$$

given by

$$\mathcal{S}(f(\cdot)) := \int_s^{\cdot} \mathcal{T}(\cdot - \sigma) f(\sigma) \mathrm{d}\sigma,$$

which corresponds to the integral component of the mild solution of the associated semilinear system [22]. The compactness of $\mathcal{S}$ yields the relative compactness of the set of trajectories. Then, the application of Mazur's Lemma to the differential inclusion approach establishes the closedness of $S_{[s,T]}(\varphi)$. Our method here consists of tailoring and extending the described approaches to our framework. Thus, for the nonlinear equation (3.4), we will introduce additional conditions on the measure and dynamics to establish the compactness of the analogous operator

$$\mathcal{I} : C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)) \times \mathcal{U}_{\text{ad}}[s,T] \to C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$$

defined by

$$\mathcal{I}(x,u) := \int_s^{\cdot} f(\sigma, x(\sigma,\cdot), u(\sigma), \cdot)\mathrm{d}\sigma, \text{ where } (x,u) \text{ is a process.} \tag{3.5}$$

Thus, if $\{(x_k,u_k)\}_k$ is a bounded sequence of processes, the corresponding sequence $\{\mathcal{I}(x_k,u_k)\}_k$ contains a subsequence converging to some function I^* in $C([s,T]:L^2(\mu,\Omega;\mathbb{R}^n))$, then, the Mazur's Lemma and Fillipov's Lemma will ensure that I^* is the integral part of a trajectory. Following Fattorini [20], that property means that the control system is *trajectory complete.*

Remark 3.1. Note that, when $\operatorname{supp}\mu$ is finite, the resulting optimal control problem is finite-dimensional, so compactness of trajectories can be proved in a standard way (see, *e.g.,* [29], Thm. 23.2).

The following set of hypotheses will play a crucial role in the proofs of the results in this subsection. We consider the property below on the measure μ consisting of these two conditions:

$$\begin{gathered}\text{for all } \ r>0: \quad h(r) \coloneqq \inf\{\mu(\mathcal{B}_r(\omega)) : \omega\in\Omega\} > 0, \text{ and}\\ \limsup_{r\to 0}\left[\sup_{\omega\in\Omega}\frac{\mu(\mathcal{B}_{2r}(\omega))}{\mu(\mathcal{B}_r(\omega))}\right] < \infty.\end{gathered} \tag{$\mathbf{H}_\mu$}$$

There exists a modulus of continuity $\theta_f(\cdot)$ such that, for all $\omega_1,\omega_2\in\Omega$,

$$\int_0^T \sup_{u\in U(t),\ x\in\mathbb{R}^n} |f(t,x,u,\omega_1) - f(t,x,u,\omega_2)|\mathrm{d}t \le \theta_f(d_\Omega(\omega_1,\omega_2)). \tag{3.6}$$

Additionally, the following property is introduced:

$$F(t,\varphi) \text{ takes convex values for all } (t,\varphi)\in[0,T]\times L^2(\mu,\Omega;\mathbb{R}^n). \tag{$\mathbf{C}$}$$

Also, as noted in Remark B.11 in the Appendix, $F(t,\varphi)$ takes compact values for almost all $t\in[s,T]$ and all $\varphi\in L^2(\mu,\Omega;\mathbb{R}^n)$.

Remark 3.2.

- Note that, whenever μ is a *doubling measure*, *i.e.,* a positive Borel measure satisfying the condition

  $$0 < \mu(\mathcal{B}_{2r}(\omega)) \le C'_\mu \mu(\mathcal{B}_r(\omega)) < \infty$$

  for all $\omega\in\Omega$, $r>0$, and some positive constant C'_μ, then both conditions in ($\mathbf{H}_\mu$) hold. On the other hand, if $\Omega = \operatorname{supp}\mu$, the condition $h(r)>0$ holds trivially. In fact, let $\Omega=\operatorname{supp}\mu$, and suppose that $h(r)>0$ does not hold. Let $r>0$ be such that $h(r)=0$. Then, there necessarily exists $(\omega_n)\subseteq\Omega$ with $\mu(\mathcal{B}_r(\omega_n))\to 0$ and $\omega_n\to\omega^*$, for some $\omega^*\in\Omega$, in view of the definition of h and the compactness of Ω. For some sufficiently large N, one has $\mathcal{B}_{r/2}(\omega^*)\subset\mathcal{B}_r(\omega_n)$ if $n\ge N$. This implies that $\mu(\mathcal{B}_{r/2}(\omega^*))=0$, which contradicts the fact that ω^* is in $\operatorname{supp}\mu$.
- The asymptotic property in ($\mathbf{H}_\mu$) implies that, if $\varphi\in L^2(\mu,\Omega;\mathbb{R}^n)$, then almost every point $\omega\in\Omega$ is a Lebesgue point of φ, (see [30], Sect. 3.4).

Remark 3.3. It is straightforward to note that the convexity of $F(t,\varphi)$ holds if, $U(t)$ is convex for a.e. $t \in [s,T]$, and the dynamics is affine with respect to the control, *i.e.,*

$$f(t,x,u,\omega) = f_0(t,x,\omega) + \sum_{i=1}^{m} f_i(t,x,\omega)\, u_i, \quad \text{a.e. } t \in [s,T].$$

Moreover, there are more general nonlinear dynamics such that conditions **(H)**(iii), (3.6), and **(C)** hold. As an example, consider $U(t) = [0,1]$, a nonlinear function $\Sigma : \mathbb{R} \to \mathbb{R}$ satisfying the inequalities

$$\begin{cases} |\Sigma(r)| < 1 + |r|, & \text{for all } r \in \mathbb{R}, \\ |\Sigma(r_1) - \Sigma(r_2)| \le C_\Sigma |r_1 - r_2|, & \text{for all } r_1, r_2 \in \mathbb{R}. \end{cases} \tag{3.7}$$

e.g., consider the sigmoid function $\Sigma(r) = (1 + \mathrm{e}^{-r})^{-1}$ or the Gaussian $\Sigma(r) = \mathrm{e}^{-r^2}$, $\mathcal{P} : \Omega \to \mathbb{R}$ a Lipschitz continuous function, $\mathcal{M} : [0,T] \to \mathbb{R}$ a measurable and bounded function, and $\mathcal{C} : \mathbb{R} \to \mathbb{R}$ a nonlinear, continuous, and monotone function. For instance, $\mathcal{C}(u) = u^3 + au$, for $a > 0$ satisfies the conditions above. Thus, $\mathcal{C}([0,1])$ is a closed interval. Then, we define the dynamics

$$f(t,x,u,\omega) := \bar{\Sigma}(x)\mathcal{M}(t)\mathcal{C}(u)\mathcal{P}(\omega),$$

where $\bar{\Sigma}(x) := (\Sigma(x_1), ..., \Sigma(x_n))$. Observe that in view of (3.7), f satisfies (**H**)(iii) and from the Lipschitz continuity of $\mathcal{P}$ observe that

$$\int_0^T \sup_{u \in U(t),\ x \in \mathbb{R}^n} |f(t,x,u,\omega_1) - f(t,x,u,\omega_2)| \mathrm{d}t \le T n C_{u,\mathcal{P}} \max_{t \in [0,T]} \mathcal{M}(t) d_\Omega(\omega_1,\omega_2),$$

and (3.6) holds. Now, from (3.7) we deduce that for $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$, $\bar{\Sigma}(\varphi) \in L^2(\mu,\Omega;\mathbb{R}^n)$, and consequently $f(t,\varphi(\cdot),u,\cdot) \in L^2(\mu,\Omega;\mathbb{R}^n)$ and it implies that the multivalued function

$$F(t,\varphi) = \bar{\Sigma}(\varphi)\mathcal{M}(t)\mathcal{C}([0,1])\mathcal{P}$$

takes convex values in $L^2(\mu,\Omega;\mathbb{R}^n)$.

Henceforth, in cases where the dependence on the parameter variable ω can be omitted, we will adopt the notation $\varphi := \varphi(\cdot)$ and $x(t) := x(t,\cdot)$ for the initial datum and trajectories, respectively, and $f(t,x(t),u(t)) := f(t,x(t,\cdot),u(t),\cdot)$ for the dynamics.

Let $\{x_k\}_k \subset C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$ be a sequence of solutions of (3.4) with $x_k(s) = \varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$ and controls $u_k \in \mathcal{U}_{\mathrm{ad}}[s,T]$. Set

$$I_k(t) := \int_s^t \bar{f}_k(\sigma)\mathrm{d}\sigma, \text{ where } \bar{f}_k(\sigma) := f(\sigma, x_k(\sigma), u_k(\sigma)). \tag{3.8}$$

Proposition 3.4. *Let us assume that* (**H**) *and* (**H**$_\mu$) *hold. Then, for all* $t \in [s,T]$, *the sequence* $\{I_k(t)\}_k$ *introduced in* (3.8) *is relatively compact in* $L^2(\mu,\Omega;\mathbb{R}^n)$.

Proof. We will apply Theorem C.1 (see Appendix C) in $L^2(\mu,\Omega;\mathbb{R}^n)$, *i.e.,* the result follows if we show that for all $t \in [s,T]$ the sequence $\{I_k(t)\}_k$ is bounded and that

$$\sup_k \int_\Omega |I_k(t,\omega) - (I_k(t))_{\mathcal{B}_r(\omega)}|^2 \mathrm{d}\mu(\omega) \to 0, \quad \text{as } r \to 0,$$

where $(I_k(t))_{\mathcal{B}_r(\omega)} := \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)} I_k(t,\omega')\,\mathrm{d}\mu(\omega')$. The inequality (1.4) and Lemma 2.1 guarantee that

$$||\bar{f}_k(\sigma)||_{L^2(\mu,\Omega;\mathbb{R}^n)} \le C\left(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2(\mu,\Omega;\mathbb{R}^n)}\right) \tag{3.9}$$

for a.e. $\sigma \in [s,T]$ and all k. It follows that, for any $t \in [s,T]$, the sequence $\{I_k(t)\}_k$ is bounded in $L^2(\mu,\Omega;\mathbb{R}^n)$. On the other hand, denoting $\bar{f}_k(\sigma,\omega) := f(\sigma, x_k(\sigma,\omega), u_k(\sigma), \omega')$, from (3.6) and (**H**), we observe that, for a.e. $\omega, \omega' \in \Omega$,

$$\begin{aligned}
\left|\int_s^t \bar{f}_k(\sigma,\omega) - \bar{f}_k(\sigma,\omega')\,\mathrm{d}\sigma\right| &\le \int_s^t |\bar{f}_k(\sigma,\omega) - \bar{f}_k(\sigma,\omega')|\,\mathrm{d}\sigma \\
&\le \int_s^t |f(\sigma,x_k(\sigma,\omega),u_k(\sigma),\omega) - f(\sigma,x_k(\sigma,\omega),u_k(\sigma),\omega')|\,\mathrm{d}\sigma \\
&\qquad + \int_s^t |f(\sigma,x_k(\sigma,\omega),u_k(\sigma),\omega') - f(\sigma,x_k(\sigma,\omega'),u_k(\sigma),\omega')|\,\mathrm{d}\sigma \\
&\le \theta_f(d_\Omega(\omega,\omega')) + k_f\int_s^t |x_k(\sigma,\omega) - x_k(\sigma,\omega')|\,\mathrm{d}\sigma \\
&\le \theta_f(d_\Omega(\omega,\omega')) + k_f(t-s)|\varphi(\omega)-\varphi(\omega')| + k_f\int_s^t\int_s^\sigma |\bar{f}_k(\tau,\omega) - \bar{f}_k(\tau,\omega')|\mathrm{d}\tau\,\mathrm{d}\sigma.
\end{aligned}$$

The Grönwall inequality leads us to

$$\int_s^t |\bar{f}_k(\sigma,\omega) - \bar{f}_k(\sigma,\omega')|\,\mathrm{d}\sigma \le \mathrm{e}^{k_f(t-s)}\left(\theta_f(d_\Omega(\omega,\omega') + k_f(t-s)|\varphi(\omega)-\varphi(\omega')|\right). \tag{3.10}$$

Considering the latter inequality, it follows that, for some positive constant $C(T)$

$$\begin{aligned}
&\int_\Omega |I_k(t,\omega) - (I_k(t))_{\mathcal{B}_r(\omega)}|^2\mathrm{d}\mu(\omega) \\
&\quad = \int_\Omega\left|\int_s^t \bar{f}_k(\sigma,\omega)\mathrm{d}\sigma - \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)}\left[\int_s^t \bar{f}_k(\sigma,\omega')\mathrm{d}\sigma\right]\mathrm{d}\mu(\omega')\right|^2\mathrm{d}\mu(\omega) \\
&\quad = \int_\Omega\left|\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)}\int_s^t \left[\bar{f}_k(\sigma,\omega) - \bar{f}_k(\sigma,\omega')\right]\mathrm{d}\sigma\,\mathrm{d}\mu(\omega')\right|^2\mathrm{d}\mu(\omega) \\
&\quad \le C(T)\mu(\Omega)\theta_f(r)^2 + C(T)\int_\Omega \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)}|\varphi(\omega)-\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\mathrm{d}\mu(\omega) \\
&\quad = C(T)\mu(\Omega)\theta_f(r)^2 + C(T)\int_\Omega \Phi_r(\omega)\,\mathrm{d}\mu(\omega),
\end{aligned} \tag{3.11}$$

where we have defined the function $\Phi_r : \Omega \to \mathbb{R}$ as

$$\Phi_r(\omega) := \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)}|\varphi(\omega)-\varphi(\omega')|^2\,\mathrm{d}\mu(\omega').$$

We claim that, for all $r > 0$ and $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$, under condition ($\mathbf{H}_\mu$), the function Φ_r satisfies that $\Phi_r \in L^1(\mu,\Omega)$ and

$$\int_\Omega \Phi_r(\omega)\,\mathrm{d}\mu(\omega) \to 0 \quad \text{as} \quad r \to 0.$$

Indeed, the following inequality holds

$$\begin{aligned}\int_\Omega \Phi_r(\omega)\,\mathrm{d}\mu(\omega) &= \int_\Omega \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)} |\varphi(\omega)-\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\,\mathrm{d}\mu(\omega)\\ &\le 2\int_\Omega \left(|\varphi(\omega)|^2 \ + \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)} |\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\right)\mathrm{d}\mu(\omega)\\ &\le 2\int_\Omega \left(|\varphi(\omega)|^2 \ + \frac{\|\varphi\|_{L^2}^2}{\mu(\mathcal{B}_r(\omega))}\right)\mathrm{d}\mu(\omega)\\ &\le 2\left(1+\frac{\mu(\Omega)}{h(r)}\right)\|\varphi\|_{L^2}^2.\end{aligned} \tag{3.12}$$

On the other hand, since $|\varphi|^2 \in L^1(\mu,\Omega)$, the absolute continuity of the integral with respect to μ implies that for every $\varepsilon > 0$, there exists $\delta_1 > 0$ such that for any μ-measurable set $E \subset \Omega$ with $\mu(E) < \delta_1$, we have

$$\int_E |\varphi(\omega)|^2\,\mathrm{d}\mu(\omega) < \frac{\varepsilon}{6\,(C_\mu+2)}, \tag{3.13}$$

where C_μ is the limit superior given in ($\mathbf{H}_\mu$). By Lusin's Theorem, for any positive $\delta < \delta_1$, there exists a μ-measurable compact set $K \subset \Omega$ such that $\mu(\Omega\setminus K) < \delta$ and the restriction $\varphi|_K$ is continuous. Since K is compact, $\varphi|_K$ is uniformly continuous and bounded. Therefore, there exists $\sigma_1 > 0$ such that, for all $\omega,\omega' \in K$ with $d_\Omega(\omega,\omega') < \sigma_1$, we have

$$|\varphi(\omega)-\varphi(\omega')|^2 < \frac{\varepsilon}{3\,\mu(\Omega)},$$

and for all $\omega \in \Omega\setminus K$, we observe that

$$\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap K} |\varphi(\omega')|^2\,\mathrm{d}\mu(\omega') < \sup_{\omega'\in K}|\varphi(\omega')|^2, \quad \text{and} \quad \lim_{r\to 0}\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap K} |\varphi(\omega')|^2\,\mathrm{d}\mu(\omega') = 0.$$

By the Dominated Convergence Theorem, there exists $\sigma_2 > 0$ such that for all $r \in (0,\sigma_2)$,

$$\int_{\Omega\setminus K}\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap K} |\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\mathrm{d}\mu(\omega) < \frac{\varepsilon}{6}. \tag{3.14}$$

From ($\mathbf{H}_\mu$), there exists $\sigma_3 > 0$ such that for all $r \in (0,\sigma_3)$ and all $\omega \in \Omega$,

$$\frac{\mu(\mathcal{B}_{2r}(\omega))}{\mu(\mathcal{B}_r(\omega))} \le C_\mu + 1.$$

We note that if $\omega' \in \mathcal{B}_r(\omega)$, then $\mathcal{B}_r(\omega') \subset \mathcal{B}_{2r}(\omega)$, and

$$\mu(\mathcal{B}_r(\omega')) \le \mu(\mathcal{B}_{2r}(\omega)) \le (C_\mu+1)\mu(\mathcal{B}_r(\omega)). \tag{3.15}$$

Let us take a positive $r \le \min\{\sigma_1,\sigma_2,\sigma_3\}$, then

$$\int_K \Phi_r(\omega)\,\mathrm{d}\mu(\omega) < \int_K \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)} \frac{\varepsilon}{3\,\mu(\Omega)}\,\mathrm{d}\mu(\omega')\,\mathrm{d}\mu(\omega) = \frac{\varepsilon}{3}\frac{\mu(K)}{\mu(\Omega)} < \frac{\varepsilon}{3}.$$

Next, we estimate the integral over $\Omega \setminus K$. First, note that

$$\{\omega \in \Omega \setminus K : \omega' \in \mathcal{B}_r(\omega)\} = \{\omega \in \Omega \setminus K : \omega \in \mathcal{B}_r(\omega')\} = \mathcal{B}_r(\omega') \cap (\Omega \setminus K).$$

Then, applying Fubini's Theorem (recall (3.12)), we have that

$$\begin{aligned}\int_{\Omega\setminus K} \Phi_r(\omega)\,\mathrm{d}\mu(\omega) &\leq 2\int_{\Omega\setminus K}\left(|\varphi(\omega)|^2 + \frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)}|\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\right)\mathrm{d}\mu(\omega)\\ &= 2\int_{\Omega\setminus K}|\varphi(\omega)|^2\,\mathrm{d}\mu(\omega) + 2\int_{\Omega\setminus K}\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap K}|\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\,\mathrm{d}\mu(\omega)\\ &\qquad + 2\int_{\Omega\setminus K}\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap(\Omega\setminus K)}|\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\,\mathrm{d}\mu(\omega)\\ &= 2\int_{\Omega\setminus K}|\varphi(\omega)|^2\,\mathrm{d}\mu(\omega) + 2\int_{\Omega\setminus K}\frac{1}{\mu(\mathcal{B}_r(\omega))}\int_{\mathcal{B}_r(\omega)\cap K}|\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\,\mathrm{d}\mu(\omega)\\ &\qquad + 2\int_{\Omega\setminus K}|\varphi(\omega')|^2\left(\int_{\mathcal{B}_r(\omega')\cap(\Omega\setminus K)}\frac{1}{\mu(\mathcal{B}_r(\omega))}\,\mathrm{d}\mu(\omega)\right)\mathrm{d}\mu(\omega').\end{aligned}$$

Given that $\mu(\Omega \setminus K) < \delta$ we apply (3.13) to the first term on the right-hand side of the second equality above, and apply (3.14) and (3.15) to the remaining terms, respectively, to obtain that

$$\begin{aligned}\int_{\Omega\setminus K}\Phi_r(\omega)\,\mathrm{d}\mu(\omega) &< \frac{\varepsilon}{3(C_\mu+2)} + \frac{\varepsilon}{3} + 2\int_{\Omega\setminus K}|\varphi(\omega')|^2\,\mathrm{d}\mu(\omega')\,(C_\mu+1)\\ &< \frac{\varepsilon}{3(C_\mu+2)} + \frac{\varepsilon}{3(C_\mu+2)}\,(C_\mu+1) + \frac{\varepsilon}{3} = \frac{2}{3}\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the claim holds. It follows that the right-hand side of (3.11) tends to zero as $r \to 0$. As already pointed out at the beginning of the proof, by Theorem C.1, the desired result follows. ☐

Remark 3.5. From Remark 3.2, the inequality (3.10) implies that, for all $t \in [s,T]$, a.e. $\omega \in \Omega$ is a Lebesgue point of $I_k(t)$ for any k. Moreover, suppose we modify condition (3.6) and put the identity on the right-hand side of the inequality, instead of θ_f. In that case, we deduce that, for all $t \in [s,T]$ and all k, $I_k(t) \in W^{1,2}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$ whenever $\varphi \in W^{1,2}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$. In fact, considering inequality (3.10) and following Hajłasz [31], the Sobolev space $W^{1,p}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$ with $p \in [1,+\infty)$ is defined as follows: $f \in W^{1,p}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$ if and only if $f \in L^p(\mu,\Omega;\mathbb{R}^n)$ and there exists a positive function $g_f \in L^p(\mu,\Omega;\mathbb{R}^n)$ such that

$$|f(\omega) - f(\omega')| \leq d_\Omega(\omega,\omega')(g_f(\omega) + g_f(\omega')) \tag{3.16}$$

almost everywhere. The space is equipped with the norm

$$\|f\|_{W^{1,p}(\mu,\Omega,d_\Omega;\mathbb{R}^n)} := \|f\|_{L^p(\mu,\Omega;\mathbb{R}^n)} + \inf_{g_f}\|g_f\|_{L^p(\mu,\Omega;\mathbb{R}^n)},$$

the infimum being taken over all positive L^p functions g_f satisfying (3.16).

It is worth mentioning that the Banach space $W^{1,2}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$ is not necessarily a Hilbert space, unless $\Omega \subset \mathbb{R}^d$, in which case, it coincides with the classical Sobolev space $H^1(\Omega;\mathbb{R}^n)$ and the well-known compact embedding results hold. However, for an arbitrary measurable metric space (Ω, d_Ω, μ), this is not the case. A compact embedding result can be obtained under the assumption that the measure μ is doubling. Under this hypothesis, the space $W^{1,p}(\mu,\Omega,d_\Omega;\mathbb{R}^n)$ is compactly embedded in $L^p(\mu,\Omega;\mathbb{R}^n)$, which gives a version of the

Rellich–Kondrachov Theorem for these Sobolev spaces defined in metric spaces [23, 24]. Note that the conditions in ($\mathbf{H}_\mu$) are much weaker than requiring μ to be a doubling measure (recall Rem. 3.2 above).

Remark 3.6. Observe that if the initial datum is such that $\varphi \in C(\Omega;\mathbb{R}^n)$ (or constant, as supposed in [3]) the condition (3.6) of ($\mathbf{H}_\mu$) in Proposition 3.4 can be dropped. In fact, the Lipschitz continuity of f (condition (1.5)), the continuity of the map $\omega \mapsto f(t,x,u,\omega)$ and the Grönwall's inequality ensure the uniform equi-continuity of $\omega \mapsto I_k(t,\omega)$ for all $t \in [0,T]$ and, consequently, the sequence $\{I_k(t)\}_k$ satisfies the hypotheses of Theorem C.1.

Following Cannarsa & Frankowska [28], Lemma 5.4, one may consider the following alternative and stronger hypothesis:

$$\textbf{(K)}\qquad \begin{array}{l}\text{For all } R>0, \text{there exists a compact set } K_R \subset L^2(\mu,\Omega;\mathbb{R}^n) \text{ such that}\\ F(t,\varphi) \subset K_R \quad \text{for every} \quad (t,\varphi) \in [s,T] \times \mathcal{B}_R(0).\end{array}$$

The following is an example of a dynamics satisfying condition (**K**).

Example 3.7. Consider $n = 1$, $d > 2$, $\Omega \subset \mathbb{R}^d$ compact with regular boundary, μ the Lebesgue measure, and $B : \mathbb{R}^m \to \mathbb{R}$ a bounded operator, for all $t \in [s,T]$ and $\varphi \in L^2(\mu,\Omega)$ consider

$$F(t,\varphi) := \{A\varphi + B[u(t)] : u(t) \in U(t)\}$$

where, $A := (I-\Delta)^{-1}$ (Δ denotes the Laplacian operator in $\mathbb{R}^d$) is the elliptic operator with homogeneous Dirichlet boundary conditions defined in the following way:

$$(I-\Delta)^{-1}\varphi = v \Leftrightarrow \begin{cases} v - \Delta v = \varphi & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.17}$$

For any $\varphi \in L^2(\mu,\Omega)$, the elliptic equation (3.17) has a unique weak solution $v \in H^1_0(\Omega) \cap H^2(\Omega)$. Thus, from the elliptic regularity (see, *e.g.,* [32], Thm. 9.25), for a positive constant C the following inequality holds

$$||v||_{H^1} \le ||v||_{H^2} = ||(I-\Delta)^{-1}\varphi||_{H^2} \le C||\varphi||_{L^2}. \tag{3.18}$$

For $\|\varphi\|_{L^2} < R$, define the set $\mathcal{K} := \left\{\psi \in H^1(\Omega) : \|\psi\|_{H^1} \le CR + \mu(\Omega)^{1/2}\|B\|\mathrm{diam}(\mathbf{U})\right\}$. From the Rellich–Kondrachov theorem [32], Theorem 9.16, $K_R = \bar{\mathcal{K}}^{L^2}$ (closure of $\mathcal{K}$ in $L^2(\mu,\Omega)$) is compact in $L^2(\mu,\Omega)$ and $F(t,\varphi) \subset K_R$ for all $(t,\varphi) \in [s,T] \times \mathcal{B}_R(0)$.

Under condition (**K**), for all $t \in [s,T]$, we observe that, if $\|\varphi\|_{L^2} < R$, from inequality (3.9), the sequence $\{I_k(t)\}_k \subset (T-s)K_R$ is evidently relatively compact in $L^2(\mu,\Omega;\mathbb{R}^n)$. Thus, we obtain the result below.

Theorem 3.8 (Compactness of the set of trajectories)**.** *Under conditions* (**H**), ($\mathbf{H}_\mu$) *(or, alternatively,* (**K**)) *and* (**C**), *for all* $s \in [0,T)$ *and* $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$), *the set* $S_{[s,T]}(\varphi)$ *is compact in* $C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$.

Proof. First, we prove that $S_{[s,T]}(\varphi)$ is relatively compact. To achieve this, it is sufficient to verify the hypotheses of the Arzelà-Ascoli Theorem:

(1) $S_{[s,T]}(\varphi)$ is equicontinuous,

(2) for all $t \in [s,T]$,

$$\mathcal{X}(t) := \{x(t)\}_{x \in S_{[s,T]}(\varphi)} \text{ is relatively compact in } L^2(\mu,\Omega;\mathbb{R}^n). \tag{3.19}$$

Proof of (1): Let $t, \tau \in [s,T]$, then for all $x \in S_{[s,T]}(\varphi)$ from Lemma 2.1, it follows that

$$||x(t) - x(\tau)||_{L^2(\mu,\Omega;\mathbb{R}^n)} \leq C\left(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2(\mu,\Omega;\mathbb{R}^n)}\right)|t-\tau|.$$

Consequently, for all $\varepsilon > 0$, if $|t-\tau| < \delta := \varepsilon\left[C\left(\mu(\Omega)^{1/2} + \|\varphi\|_{L^2(\mu,\Omega;\mathbb{R}^n)}\right)\right]^{-1}$, we have that $||x(t)-x(\tau)||_{L^2} < \varepsilon$, for all $x \in S_{[s,T]}(\varphi)$.

Proof of (2): If **(K)** holds, *(2)* follows straightforwardly. Suppose alternatively that $(\mathbf{H}_\mu)$ holds and let $t \in [s,T]$ be fixed and $\{x_k(t)\}_k$ be a sequence in $\mathcal{X}(t)$. For each k, there exists a control $u_k \in \mathcal{U}_{\mathrm{ad}}[s,T]$ such that

$$x_k(t) = \varphi + I_k(t), \tag{3.20}$$

where I_k was defined in (3.8). From Proposition 3.4, $\{I_k(t)\}_k$ is relatively compact in $L^2(\mu,\Omega;\mathbb{R}^n)$, thus $\{x_k(t)\}$ is relatively compact in $L^2(\mu,\Omega;\mathbb{R}^n)$. Then, the Arzelà-Ascoli Theorem ensures the relative compactness of the set $S_{[s,T]}(\varphi)$ in $C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$. Now, suppose that $\{x_k\}_k$ is a strongly convergent sequence in $S_{[s,T]}(\varphi)$ such that

$$x_k \overset{s}{\to} x \quad \text{in } C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)). \tag{3.21}$$

We prove next that $x \in S_{[s,T]}(\varphi)$. From (3.20) we have that

$$I_k \overset{s}{\to} I^* \quad \text{in } C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)),$$

for some $I^* \in C([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$. Then

$$x(t) = \varphi + I^*(t), \quad t \in [s,T]. \tag{3.22}$$

On the other hand, by (3.9) we observe that there exists a subsequence (keeping the same index) of $\{I_k'\} = \{\bar{f}_k\}$ satisfying

$$\bar{f}_k \overset{*}{\rightharpoonup} \bar{f} \quad \text{in } L^\infty([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)), \tag{3.23}$$

for some $\bar{f} \in L^\infty([s,T] : L^2(\mu,\Omega;\mathbb{R}^n))$. Moreover, in view of Theorem C.2, we can extract a subsequence such that

$$\bar{f}_k \overset{w}{\rightharpoonup} \dot{x} \quad \text{in } L^1([s,T] : L^2(\mu,\Omega;\mathbb{R}^n)).$$

Thus, from the uniqueness of the limits in the latter two equations and using (3.22), we deduce that x satisfies

$$x(t) = \varphi + \int_s^t \bar{f}(\sigma)\mathrm{d}\sigma \qquad \text{for } t \in [s,T].$$

By (3.21), for all $\varepsilon > 0$, there exists k_0 such that

$$x_k(t) \in \mathcal{B}_\varepsilon(x(t)) \quad \text{for all } \ t \in [s,T], \quad k \geq k_0,$$

and

$$u_k(t) \in U(t) \subset U(\mathcal{B}_\varepsilon(t)) \qquad \text{for all } \ t \in [s,T], \quad k \geq k_0.$$

Consequently,

$$\bar{f}_k(t) \in f\left(t, \mathcal{B}_\varepsilon(x(t)), U(\mathcal{B}_\varepsilon(t))\right) \qquad \text{a.e. } t \in [s,T], \quad \text{all} \quad k \geq k_0. \tag{3.24}$$

The limit in (3.23) and the Mazur's Lemma provide us with coefficients $\alpha_{ij} \geq 0$, for which $\sum_{i\geq 1}^{N(i)} \alpha_{ij} = 1$ for each j, such that, for some subsequences $\{\bar{f}_{ij}\}_j$ of $\{\bar{f}_k\}_k$ and some $p > 1$, the following limit stands

$$\psi_j := \sum_{i\geq 1}^{N(i)} \alpha_{ij} \bar{f}_{ij} \xrightarrow{s} \bar{f} \quad \text{in } L^p([0,T] : L^2(\mu, \Omega; \mathbb{R}^n)). \tag{3.25}$$

Then, we deduce that

$$\psi_j(t) \xrightarrow{s} \bar{f}(t) \;\; \text{in} \;\; L^2(\mu, \Omega; \mathbb{R}^n) \qquad \text{a.e. } t \in [s,T],$$

and from (3.24) we have that

$$\psi_j(t) \in \operatorname{co} f\left(t, \mathcal{B}_\varepsilon(x(t)), U(\mathcal{B}_\varepsilon(t))\right) \qquad \text{a.e.} \;\; t \in [s,T].$$

Thus, for all $\varepsilon > 0$

$$\bar{f}(t) \in \overline{\operatorname{co}} f\left(t, \mathcal{B}_\varepsilon(x(t)), U(\mathcal{B}_\varepsilon(t))\right) \qquad \text{a.e.} \;\; t \in [s,T].$$

Since the multifunction U is pseudo-continuous and $f(t, x(t), U(t)) = F(t, x(t))$ is convex (condition **(C)**) and closed for a.e. $t \in [s,T]$. Then, the Proposition B.10 implies that

$$\bar{f}(t) \in f\left(t, x(t), U(t)\right) \;\; \text{a.e.} \;\; t \in [s,T].$$

Then, in view of Corollary B.9, there exists a measurable control $u : [s,T] \to \mathbb{R}^m$, such that

$$\begin{cases} u(t) \in U(t) \qquad \text{a.e. } t \in [s,T], \\ \\ \bar{f}(t) = f(t, x(t), u(t)) \qquad \text{a.e. } t \in [s,T]. \end{cases}$$

This leads us to conclude that $x \in S_{[s,T]}(\varphi)$ and thus the proof follows. □

Remark 3.9. Under hypotheses $(\mathbf{H}_\mu)$ or $(\mathbf{K})$, the arguments employed to establish the Arzelà-Ascoli conditions in the proof of Theorem 3.8 above can be adapted to demonstrate the compactness of the operator $\mathcal{I}$ defined in (3.5).

3.2. Lower semi-continuity of the value function

With the compactness of the set of trajectories in hand, we can solve problem $(P)_{s,\varphi}$ once we have established the lower semi-continuity of the functional $\mathcal{J}$. The result below characterizes the lower semi-continuity of the value function and gives a result on the existence of optimal trajectories. We observe that the value function V can be redefined as follows: for all $(s, \varphi) \in [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n)$,

$$V\left(s, \varphi\right) := \inf \left\{ \int_\Omega g(x(T, \omega), \omega) \mathrm{d}\mu(\omega) : x \in S_{[s,T]}\left(\varphi\right) \right\}.$$

Furthermore, we define $\operatorname{dom} V := \left\{ (s, \varphi) \in [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n) : \;\; V(s, \varphi) < \infty \right\}$.

Theorem 3.10. *If* $(\mathbf{H})$, $(\mathbf{H}_\mu)$ *(or, alternatively,* $(\mathbf{K})$*), and* $(\mathbf{C})$ *hold true, then* V *is lower semi-continuous and*

$$V(s,\varphi)=\min\left\{\int_\Omega g(x(T,\omega),\omega)\mathrm{d}\mu(\omega) : x\in S_{[s,T]}(\varphi)\right\},\ \textit{for all}\ (s,\varphi)\in[0,T]\times L^2(\mu,\Omega;\mathbb{R}^n). \tag{3.26}$$

Furthermore, for all $\bar\varphi\in L^2(\mu,\Omega;\mathbb{R}^n)$,

$$\begin{gathered}\int_\Omega g(\bar\varphi(\omega),\omega)\mathrm{d}\mu(\omega)=\liminf_{s\to T^-,\varphi\to\bar\varphi} V(s,\varphi),\\ V(0,\bar\varphi)=\liminf_{s\to 0^+,\varphi\to\bar\varphi} V(s,\varphi).\end{gathered} \tag{3.27}$$

Proof. We note that the compactness of $S_{[s,T]}(\varphi)$ implies the compactness of the set $\mathcal{X}(T)$ (defined in (3.19)) in $L^2(\mu,\Omega;\mathbb{R}^n)$ and, from Theorem B.12 (see Appendix), the functional $\mathcal{J}$ is lower semi-continuous. Thus, we obtain (3.26).

In order to prove the semi-continuity of V, we take a sequence (s_k,φ_k) converging strongly to (s,φ) in $[0,T]\times L^2(\mu,\Omega;\mathbb{R}^n)$. Let $\bar x_k\in S_{[s_k,T]}(\varphi_k)$ be the optimal trajectory associated with $V(s_k,\varphi_k)$, and with control $\bar u_k\in\mathcal{U}_{\mathrm{ad}}[s_k,T]$ for all k. We know that the sequence of trajectories $\{\bar x_k\}_k$ can be written as

$$\bar x_k(t)=\varphi_k+\int_{s_k}^t f(\sigma,\bar x_k(\sigma),\bar u_k(\sigma))\mathrm{d}\sigma \quad \text{for all} \quad t\in[s_k,T],$$

or

$$\bar x_k(t)=\bar x_k(s)+\int_{s}^t f(\sigma,\bar x_k(\sigma),\bar u_k(\sigma))\mathrm{d}\sigma \quad \text{for all} \quad t\in[s_k,T]. \tag{3.28}$$

where

$$\bar x_k(s)=\varphi_k+\operatorname{sgn}(s-s_k)\int_{s_k}^s f(\sigma,\bar x_k(\sigma),\bar u_k(\sigma))\,\mathrm{d}\sigma.$$

By employing the convergence assumptions mentioned above, and that the dynamics f satisfies the inequality (3.9), we obtain that

$$\lim_{k\to\infty}\int_{s_k}^s f(\sigma,\bar x_k(\sigma),\bar u_k(\sigma))\,\mathrm{d}\sigma=0 \quad \text{in} \quad L^2(\mu,\Omega;\mathbb{R}^n),$$

which implies that $\bar x_k(s)\to\varphi$ in $L^2(\mu,\Omega;\mathbb{R}^n)$. On the other hand, from the proof of Proposition 3.4 we note that, for all $t\in[s,T]$, the sequence

$$\left\{\int_s^t f(\sigma,\bar x_k(\sigma),\bar u_k(\sigma))\,\mathrm{d}\sigma\right\}_k$$

is bounded in $L^2(\mu,\Omega;\mathbb{R}^n)$ and satisfies the inequality (3.11), then, from Theorem C.1, it is relatively compact in $L^2(\mu,\Omega;\mathbb{R}^n)$, and arguing as in the proof of the Theorem 3.8, from (3.28), we deduce that there exists a subsequence of $\{\bar x_k\}_k$ (using the same index) and a process (x^*,u^*) with $x^*\in S_{[s,T]}(\varphi)$ and $u^*\in\mathcal{U}_{\mathrm{ad}}[s,T]$, such that $\bar x_k\to x^*$ in $C([s,T]:L^2(\mu,\Omega;\mathbb{R}^n))$ as $k\to\infty$, in particular, we observe that $\bar x_k(T)\to x^*(T)$ in $L^2(\mu,\Omega;\mathbb{R}^n)$ as $k\to\infty$, then, from the semi-continuity of the functional $\mathcal{J}$ we obtain

$$\liminf_{k\to\infty} V(s_k,\varphi_k)=\liminf_{k\to\infty}\int_\Omega g(\bar x_k(T,\omega),\omega)\mathrm{d}\mu(\omega)\ge\int_\Omega g(x^*(T,\omega),\omega)\mathrm{d}\mu(\omega)\ge V(s,\varphi).$$

To prove (3.27), we consider φ in $L^2(\mu,\Omega;\mathbb{R}^n)$ and $s_k \to T^-$, $\varphi_k \to \varphi$, then the lower semi-continuity of V implies that

$$V(T,\varphi) \leq \liminf_{k\to\infty} V(s_k,\varphi_k).$$

On the other hand, consider $\{\psi_k\}_k$ in $L^2(\mu,\Omega;\mathbb{R}^n)$ and $y_k \in S_{[s_k,T]}(\psi_k)$ such that $y_k(T)=\varphi$, and since there exists a positive constant C (Lem. 2.1) such that

$$||\varphi-\psi_k||_{L^2(\mu,\Omega;\mathbb{R}^n)} = ||y_k(T)-y_k(s_k)||_{L^2(\mu,\Omega;\mathbb{R}^n)} \leq C|T-s_k|,$$

we conclude that $\psi_k \to \varphi$ in $L^2(\mu,\Omega;\mathbb{R}^n)$ as $k\to\infty$, and

$$V(s_k,\psi_k) \leq \int_\Omega g(y_k(T,\omega),\omega)\mathrm{d}\mu(\omega) = \int_\Omega g(\varphi(\omega),\omega)\mathrm{d}\mu(\omega) = V(T,\varphi),$$

thus,

$$\liminf_{k\to\infty} V(s_k,\psi_k) \leq V(T,\varphi).$$

and the first equality in (3.27) holds. Now, we denote $x \in S_{[0,T]}(\varphi)$ as the optimal trajectory associated with $V(0,\varphi)$, from the Dynamic Programming Principle we know that

$$V(0,\varphi) = V(s,x(s)) \quad \text{for all} \quad s\in[0,T],$$

and the second equality follows thanks to the lower semi-continuity of V. □

4. INVARIANCE PRINCIPLES AND PROXIMAL ANALYSIS

This section aims at proving that the value function is the unique lower semi-continuous solution of the Hamilton–Jacobi–Bellman equation. We will make use of notions from proximal normal analysis, the Dynamic Programming Principle and invariance principles from [13].

We begin by recalling basic definitions of invariance principles. Let X be a Hilbert space, $D:[0,\infty)\rightsquigarrow X$, $\Gamma:[0,\infty)\times X \rightsquigarrow X$ be given multifunctions and $G := \operatorname{Gr} D$.

Definition 4.1. Let $D:[0,\infty)\rightsquigarrow X$ be closed-valued. We say that D is *left absolutely continuous* when, for every T, every bounded subset $B\subset X$ and every $\varepsilon>0$, there exists $\delta>0$ such that, for every sequence $\{(t_i,s_i)\}_{i=1}^\infty$ of open pairwise disjoint subintervals of $[0,T]$, one has that

$$\sum_{i=1}^{\infty}(s_i-t_i)<\delta \implies \sum_{i=1}^{\infty}\operatorname{ex}(D(t_i)\cap B, D(s_i))<\varepsilon, \tag{4.1}$$

where $\operatorname{ex}(U,V) := \sup_{u\in U}\operatorname{dist}(u,V)$ and $\operatorname{dist}(u,V) := \inf_{v\in V}|u-v|$. For right absolute continuity and absolute continuity, one replaces the expression in the sum of the right-hand side of (4.1) by $\operatorname{ex}(D(s_i)\cap B, D(t_i))$ respectively, by $\operatorname{ex}_B(D(t_i),D(s_i))$, where

$$\operatorname{ex}_B(D(t_i),D(s_i)) := \max\{\operatorname{ex}(D(t_i)\cap B, D(s_i)), \operatorname{ex}(D(s_i)\cap B, D(t_i))\}.$$

Definition 4.2 (Weak invariance)**.** The graph G of D is *weakly invariant* w.r.t. the set-valued dynamics $\dot{x} \in \Gamma(t, x)$ (and we write (Γ, G) is weakly invariant) if, for any initial condition $x_0 \in D(t_0)$, there exists $T > t_0$, such that the Cauchy problem

$$\begin{cases} \dot{x}(t) \in \Gamma(t, x(t)) & t \in [t_0, T], \\ x(t_0) = x_0, \end{cases} \tag{4.2}$$

admits a solution $x(t) \in D(t)$ for all $t \in [t_0, T)$.

Definition 4.3 (Strong invariance)**.** The graph G of D is *strongly invariant* w.r.t. the set-valued dynamics $\dot{x} \in \Gamma(t, x)$ (and we write (Γ, G) is strongly invariant) if, for any initial condition $x_0 \in D(t_0)$ and any $T > t_0$, every solution of the Cauchy problem (4.2) satisfies the condition $x(t) \in D(t)$, for all $t \in [t_0, T]$.

Definition 4.4 (Proximal normal [19])**.** Let $K \subset X$ be a closed set. A *proximal normal* to K at a point $x \in K$ is a vector $\xi \in X$ such that there exists $\lambda > 0$ satisfying

$$\langle \xi, x' - x \rangle_X \leq \lambda ||x' - x||_X^2 \quad \text{for all } x' \in K.$$

The set of all such vectors is a cone denoted by $N_K^P(x)$ and called *proximal normal cone* to K at x.

Definition 4.5 (Proximal subgradient and subdifferentials [19])**.** A vector $\xi \in X$ is called a *proximal subgradient* (shortly, P*-subgradient*) of an extended-valued lower semi-continuous function $V : X \to \mathbb{R} \cup \{+\infty\}$ at $x \in \operatorname{dom} V$ whenever

$$(\xi, -1) \in N^P_{\text{epi } V}(x, V(x)).$$

The (possibly empty) set of all such ξ is denoted by $\partial_P V(x)$, and is referred to as *proximal subdifferential* or P*-subdifferential*. If V is an upper semi-continuous function, one sets $\partial^P V(x) := -\partial_P(-V(x))$.

Example 4.6. The proximal subdifferential might be empty. For instance, consider the function $V : \mathbb{R} \to \mathbb{R}$ given by $V(x) = -|x|$, note that $N^P_{\text{epi } V}(0, V(0)) = \{(0, 0)\}$. Therefore $\partial_P V(0) = \emptyset$.

Now, we consider the Hilbert space $X = L^2(\mu, \Omega; \mathbb{R}^n)$ and introduce the notion of *proximal solution* for the HJB equation (1.2) which, in the continuous case, is equivalent to the definition of viscosity solution [12] introduced by Crandall & P.-L. Lions [33].

Remark 4.7. In view of Fillipov's Lemma B.8, the Hamiltonian

$$H : [0, T] \times L^2(\mu, \Omega; \mathbb{R}^n) \times L^2(\mu, \Omega; \mathbb{R}^n)^* \to \mathbb{R}$$

defined in (1.3) can be rewritten as follows

$$H(t, \varphi, p) = \min_{v \in F(t, \varphi)} \langle v, p \rangle.$$

The next definition was provided in [17] in the case in which the control system is merely measurable with respect to time. Such a definition can be regarded as a generalization of the proximal solution defined in [12]. In the remainder of this section, we recall the definition of the functional $\mathcal{J} : L^2(\mu, \Omega; \mathbb{R}^n) \to \mathbb{R} \cup \{\infty\}$, given by

$$\mathcal{J}(\varphi) := \int_\Omega g(\varphi(\omega), \omega) \mathrm{d}\mu(\omega).$$

Definition 4.8 (Proximal solution)**.** Take $W : [0,T] \times L^2(\mu,\Omega;\mathbb{R}^n) \to \mathbb{R} \cup \{\infty\}$ a lower semi-continuous function, and

$$P_W(t) := \left\{(\varphi,\alpha) \in L^2(\mu,\Omega;\mathbb{R}^n) \times \mathbb{R} : W(t,\varphi) \leq \alpha\right\}.$$

We say that W is a *lower semicontinuous proximal solution* of (1.2), if $P_W(t)$ is an absolutely continuous multifunction, $W(T,\varphi) = \mathcal{J}(\varphi)$ for all $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$ such that $(T,\varphi) \in \operatorname{dom} W$, and there exists a full measure set $I \subset (0,T)$ such that for every $(t,\varphi) \in (I \times L^2(\mu,\Omega;\mathbb{R}^n)) \cap \operatorname{dom} W$, one has that

$$\xi_t + \min_{v \in F(t,x)} \langle v, \xi_\varphi \rangle = 0 \quad \text{for all} \quad (\xi_t, \xi_\varphi) \in \partial_P W(t,\varphi).$$

Let us now recall some results of weak and strong invariance, and Hamiltonian characterizations (see, *e.g.*, T. Donchev [13]).

Theorem 4.9 (Characterization of weak invariance [13])**.** *Assume that Γ is almost upper semi-continuous (see Def. B.5) with compact and convex values. Suppose that there exists $c \in L^1(0,\infty)$ such that $|\Gamma(t,x)| \leq c(t)$ a.e. $t \in [0,\infty)$. Then the pair (Γ, G) is weakly invariant if and only if the following two conditions hold true:*

i) the set-valued map $t \rightsquigarrow D(t)$ is left absolutely continuous;
ii) there exists a full-measure subset I of $[0,\infty)$ such that, for every $(t,x) \in G$ with $t \in I$, one has

$$\xi_0 + \min_{v \in \Gamma(t,x)} \langle v, \xi \rangle \leq 0 \quad \textit{for all} \;\; (\xi_0, \xi) \in N_G^P(t,x).$$

Theorem 4.10 (Characterization of strong invariance [13])**.** *Assume that Γ is almost lower semi-continuous and $x \rightsquigarrow \Gamma(t,x)$ is Lipschitz continuous, a.e. $t \in [0,\infty)$. Then the pair (Γ, G) is strongly invariant if and only if there exists a full-measure subset I of $[0,\infty)$ such that, for every $(t,x) \in G$ with $t \in I$, one has*

$$\xi_0 + \max_{v \in \Gamma(t,x)} \langle v, \xi \rangle \leq 0 \quad \textit{for all} \;\; (\xi_0, \xi) \in N_G^P(t,x).$$

4.1. Invariance properties and the HJB equation

Now, we will demonstrate the invariance properties of an autonomous system involving the multifunction F (see (3.2)). These results will show that the value function V defined in (1.7) is the unique lower semi-continuous solution of the HJB equation. This is done by applying the Dynamical Programming Principle given in Theorem 2.2.

Other properties of F can be derived from the assumptions (**H**), in fact, one has that

1. for all $t \in [0,T]$ fixed, $F(t,\cdot)$ is continuous; and
2. $F(t,\varphi)$ is $\mathcal{L} \otimes \mathcal{B}$- measurable.

Consequently, from a Scorza–Dragoni type result [34], for all $\varepsilon > 0$, there exists a compact $K_\varepsilon \subset [0,T]$ such that $\operatorname{meas}([0,T] \setminus K_\varepsilon) < \varepsilon$, and $F(t,\varphi)$ is continuous in $K_\varepsilon \times L^2(\mu,\Omega;\mathbb{R}^n)$. Observe that these properties guarantee that the multifunctions

$$\{1\} \times F(t,\varphi) \times \{0\} \quad \text{and} \quad \{1\} \times \{-F(T-t,\varphi)\} \times \{0\}$$

are almost upper semi-continuous and almost lower semi-continuous, respectively. Thus they satisfy the regularity hypotheses of Theorems 4.9 and 4.10, respectively. Recall also that F takes compact values, as noted in Remark B.11. Thus, if F takes convex values, it remains to prove weak and strong invariance to derive the HJB inequalities of Theorems 4.9 and 4.10, respectively. For this, we introduce the following:

Let V be the value function of problem $(P)_{s,\varphi}$, and let us define the set

$$P(t) := \left\{(\varphi,\alpha) \in L^2(\mu,\Omega;\mathbb{R}^n)\times\mathbb{R} : V(t,\varphi)\leq\alpha\right\},$$

and consider the associated set-valued function $t \rightsquigarrow P(t)$. For each $t\in[0,T]$, such a set $P(t)$ can be viewed as the epigraph of the function $L^2(\mu,\Omega;\mathbb{R}^n)\ni\varphi\mapsto V(t,\varphi)$, and $\operatorname{Gr}P$ coincides with

$$\operatorname{epi}V := \left\{(t,\varphi,\alpha) \in L^2(\mu,\Omega;\mathbb{R}^n)\times\mathbb{R} : V(t,\varphi)\leq\alpha\right\},$$

the epigraph of the value function, which is a closed set due to the lower semi-continuity of V.

Next, we prove a regularity result for the multifunction P.

Proposition 4.11. *If hypotheses* **(H)**, $(\mathbf{H}_\mu)$ *(or alternatively* **(K)**), *and* ***(C)*** *hold true, then the set-valued map*

$$t \rightsquigarrow P(t)$$

is absolutely continuous.

Proof. Fix any $0\leq s\leq\tau\leq T$, let $B\times I\subset L^2(\mu,\Omega;\mathbb{R}^n)\times\mathbb{R}$ be a bounded set and $(\varphi,\alpha)\in P(s)\cap B\times I$, consider $\bar{x}\in S_{[s,T]}(\varphi)$ the optimal solution, then $V(t,\bar{x}(t))=\int_\Omega g(\bar{x}(T,\omega),\omega)\mathrm{d}\mu(\omega)$ for all $t\in[s,T]$. This implies that $V(\tau,\bar{x}(\tau))\leq\alpha$, and $(\bar{x}(\tau),\alpha)\in P(\tau)$. Thus, from Lemma 2.1 we have that

$$\begin{aligned}\operatorname{dist}((\varphi,\alpha),P(\tau)) &\leq \|\varphi-\bar{x}(\tau)\|_{L^2(\mu,\Omega;\mathbb{R}^n)}\\ &\leq C\left(\mu(\Omega)^{1/2}+||\varphi||_{L^2(\mu,\Omega;\mathbb{R}^n)}\right)|\tau-s|,\end{aligned}$$

which leads to

$$\operatorname{ex}\left(P\left(s\right)\cap B\times I, P\left(\tau\right)\right)\leq C\left(\mu(\Omega)^{1/2}+\sup_{\varphi\in B}||\varphi||_{L^2(\mu,\Omega;\mathbb{R}^n)}\right)|\tau-s|.$$

On the other hand, let $(\varphi,\alpha)\in P(\tau)\cap B\times I$, and consider a trajectory x with initial time s, and such $x(\tau)=\varphi$. Then $V(s,x(s))\leq V(\tau,\varphi)\leq\alpha$, and $(x(s),\alpha)\in P(s)$. Thus, arguing as above

$$\begin{aligned}\operatorname{dist}((\varphi,\alpha),P(s)) &\leq \|x(\tau)-x(s)\|_{L^2(\mu,\Omega;\mathbb{R}^n)}\\ &\leq C\left(\mu(\Omega)^{1/2}+||x(s)||_{L^2(\mu,\Omega;\mathbb{R}^n)}\right)|\tau-s|,\end{aligned}$$

and

$$\operatorname{ex}\left(P\left(\tau\right)\cap B\times I, P\left(s\right)\right)\leq C\left(\mu(\Omega)^{1/2}+||x||_{C([s,\tau];L^2(\mu,\Omega;\mathbb{R}^n))}\right)|\tau-s|.$$

□

The following result establishes invariance properties related to F and the value function V.

Proposition 4.12 (Invariance results on V)**.** *Let us assume that* **(H)**, $(\mathbf{H}_\mu)$ *(or alternatively* **(K)**), *and the convexity assumption* ***(C)*** *are satisfied. Then:*

i) *the pair* $(\{1\}\times F\times\{0\},\operatorname{Gr}P)$ *is weakly invariant;*

ii) *the pair* $(\{1\}\times\tilde{F}\times\{0\},\operatorname{Gr}\tilde{P})$ *is strongly invariant, where* $\tilde{F}(t,\varphi):=-F(T-t,\varphi)$ *and* $\tilde{P}(t):=P(T-t)$ *for every* $(t,\varphi)\in[0,T]\times L^2(\mu,\Omega;\mathbb{R}^n)$.

Proof. i). Take any $(t_0, \varphi_0, \alpha_0) \in [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R}$ such that $(\varphi_0, \alpha_0) \in P(t_0)$. Take the optimal pair $(\bar{x}, \bar{u})$ of $(P)_{(t_0,\varphi_0)}$. It follows from the Dynamic Programming Principle that $V(t, \bar{x}(t)) = V(t_0, \varphi_0) \leq \alpha_0$ for every $t \in [t_0, T]$, implying that $(\bar{x}(t), \alpha_0) \in P(t)$ for all $t \in [t_0, T]$.

ii). Take any $(t_0, \varphi_0, \alpha_0) \in [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R}$ such that $(\varphi_0, \alpha_0) \in \tilde{P}(t_0)$. Consider any solution $\tilde{x}(\cdot)$ of the problem

$$\begin{cases} \dot{x}(t) \in \tilde{F}(t, x(t)), & \text{a.e. } t \in [t_0, T], \\ x(t_0) = \varphi_0. \end{cases}$$

One needs to show that

$$(\tilde{x}(t), \alpha_0) \in \tilde{P}(t) = \left\{ (\varphi, \alpha) \in L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R} : \quad \tilde{V}(t, \varphi) \leq \alpha \right\},$$

where $\tilde{V}(t,x) := V(T-t, x)$. Let us define the curve $y(s) := \tilde{x}(T-s)$, $s \in [0, T-t_0]$ (namely, $s = T-t$), and observe that

$$\frac{\mathrm{d}}{\mathrm{d}s} y(s) = \frac{\mathrm{d}}{\mathrm{d}s} \tilde{x}(T-s) \in -\tilde{F}(T-s, \tilde{x}(T-s)) = F(s, y(s)).$$

In view of the Dynamic Programming Principle, one has that

$$V(s, y(s)) \leq V(T-t_0, y(T-t_0)) = \tilde{V}(t_0, \tilde{x}(t_0)) \leq \alpha_0, \qquad \forall\, s \in [0, T-t_0].$$

Since $s = T-t$, the left hand side of the previous inequality reads as

$$V(s, y(s)) = V(s, \tilde{x}(T-s)) = V(T-t, \tilde{x}(t)) = \tilde{V}(t, \tilde{x}(t)),$$

providing the inequality

$$\tilde{V}(t, \tilde{x}(t)) \leq \tilde{V}(t_0, \tilde{x}(t_0)) \leq \alpha_0, \qquad \forall\, t \in [t_0, T],$$

which gives the desired condition. ☐

The following result is crucial for proving the uniqueness of the solution to the HJB equation.

Theorem 4.13 (Comparison results for V)**.** *Take a lower semi-continuous function* $\theta : [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n) \to (-\infty, \infty]$ *such that*

$$\theta(T, \varphi) = \mathcal{J}(\varphi) \qquad \text{for all } \varphi \in L^2(\mu, \Omega; \mathbb{R}^n).$$

Define

$$P_\theta(t) := \left\{ (\varphi, \alpha) \in L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R} : \quad \theta(t, \varphi) \leq \alpha \right\}$$

and assume that P_θ is an absolutely continuous multifunction. Then

i) if the pair $(\{1\} \times F \times \{0\}, \operatorname{Gr} P_\theta)$ is weakly invariant, then $V(t,\varphi) \leq \theta(t,\varphi)$ for all $(t,\varphi) \in \operatorname{dom} V \cap \big([0,T) \times L^2(\mu, \Omega; \mathbb{R}^n)\big)$.

ii) if the pair $(\{1\} \times \tilde{F} \times \{0\}, \operatorname{Gr} \tilde{P}_\theta)$ is strongly invariant, then $V(t,\varphi) \geq \theta(t,\varphi)$ for all $(t,\varphi) \in \operatorname{dom} V \cap \big([0,T) \times L^2(\mu, \Omega; \mathbb{R}^n)\big)$, where $\tilde{F}(t,\varphi) := -F(T-t, \varphi)$ and $\tilde{P}_\theta(t) := P_\theta(T-t)$.

Proof. i) Since the pair $(\{1\} \times F \times \{0\}, \mathrm{Gr}P_\theta)$ is weakly invariant then, for every $(t_0, \varphi_0, \alpha) \in [0,T) \times L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R}$ such that $(\varphi_0, \alpha) \in P_\theta(t_0)$, there exists a trajectory x satisfying the Cauchy problem

$$\begin{cases} \dot{x}(t) \in F(t, x(t)) & \text{a.e. } t \in [t_0, T), \\ x(t_0) = \varphi_0, \end{cases} \tag{4.3}$$

and $(x(t), \alpha) \in P_\theta(t)$ for all $t \in [t_0, T)$. Choosing $\alpha = \theta(t_0, \varphi_0)$, the weak invariance property and the absolutely continuity of the multifunction P_θ imply that

$$\theta(t, x(t)) \leq \theta(t_0, \varphi_0) \qquad \text{for all } \ t \in [t_0, T).$$

In view of the lower semi-continuity of θ, letting $t \to T^-$, one obtains the inequality

$$\mathcal{J}(x(T)) \leq \liminf_{t \to T^-} \theta(t, x(t)) \leq \theta(t_0, \varphi_0)$$

and, when $(t_0, \varphi_0) \in \mathrm{dom}V$, taking the infimum over all the trajectories, one gets

$$V(t_0, \varphi_0) \leq \theta(t_0, \varphi_0).$$

ii) If $(t_0, \varphi_0) \in [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n)$ is such that $V(t_0, \varphi_0) = \infty$, then there is nothing to prove. On the other hand, for every $(t_0, \varphi_0) \in \mathrm{dom}\, V$, consider the Cauchy problem (4.3). Let us define $y(s) := x(T - s)$ and observe that

$$\dot{y}(s) = \dot{x}(T-s) \in -F(T-s, x(T-s)) = \tilde{F}(s, y(s)) \quad \text{a.e.} \quad s \in [0, T - t_0],$$

and that $y(0) = x(T)$. Since the pair $(\{1\} \times \tilde{F} \times \{0\}, \mathrm{Gr}\tilde{P}_\theta)$ is strongly invariant, this implies that, for every $\alpha \in \mathbb{R}$, such that $(x(T), \alpha) \in \tilde{P}_\theta(0)$, every solution of the Cauchy problem

$$\begin{cases} \dot{y}(s) \in \tilde{F}(s, y(s)) & \text{a.e. } s \in [0, T - t_0], \\ y(0) = x(T), \end{cases} \tag{4.4}$$

satisfies that $(y(s), \alpha) \in \tilde{P}_\theta(s)$ for all $s \in [0, T - t_0]$. Choosing $\alpha = \theta(T, x(T))$, the strong invariance property implies that for any solution of the Cauchy problem (4.4), it follows that

$$\theta(T-s, y(s)) = \theta(T-s, x(T-s)) \leq \theta(T, x(T)) \quad \text{for all } s \in [0, T - t_0]. \tag{4.5}$$

In particular, (4.5) holds for $\bar{y}(s) := \bar{x}(T-s)$ where $\bar{x}$ is the optimal trajectory of $(P)_{(t_0, \varphi_0)}$. Thus, in view of the Dynamic Programming Principle, one has that

$$\theta(T-s, \bar{x}(T-s)) \leq \theta(T, \bar{x}(T)) = \mathcal{J}(\bar{x}(T)) = V(T, \bar{x}(T)) = V(T-s, \bar{x}(T-s)) \quad \text{for all } s \in [0, T-t_0].$$

Finally, when $s = T - t_0$, we have that $\theta(t_0, \varphi_0) \leq V(t_0, \varphi_0)$. This concludes the proof. □

Finally, we are in a position to prove the main result of this paper.

Theorem 4.14. *Let us assume that* $(\mathbf{H})$, $(\mathbf{H}_\mu)$ *(or alternatively* $(\mathbf{K})$*), and* ***(C)*** *hold true. Suppose that*

$$\mathrm{meas}\{t \in [0,T] \,|\, \exists\, \varphi \in L^2(\mu, \Omega; \mathbb{R}^n), \ \{0\} \neq N^P_{\mathrm{epi}\, V}(t, \varphi, V(t, \varphi)) \subset \mathbb{R} \times L^2(\mu, \Omega; \mathbb{R}^n) \times \{0\}\} = 0. \tag{4.6}$$

Then, the value function V of problem $(P)_{s,\varphi}$ is the unique lower semi-continuous, bounded from below function such that there exists a set $I \subseteq [0,T)$ of full measure for which, for every $(t,\varphi) \in (I \times L^2(\mu,\Omega;\mathbb{R}^n)) \cap \operatorname{dom} V$, one has

$$\begin{aligned} &\xi_t + \min_{v\in F(t,\varphi)} \langle v, \xi_\varphi\rangle = 0 \quad \textit{for all } \ (\xi_t,\xi_\varphi) \in \partial_P V(t,\varphi), \\ &V(T,\varphi) = \mathcal{J}(\varphi). \end{aligned} \tag{4.7}$$

Proof. In view of Theorem 3.10, and Theorem 4.9, the value function is a lower semi-continuous, bounded below function and there exists a full measure interval $I_1 \subseteq [0,T)$ such that, for every $(t,\varphi,\alpha) \in \text{epi}V \cap \big(I_1 \times L^2(\mu,\Omega;\mathbb{R}^n) \times \mathbb{R}\big)$, one has

$$\xi_0 + \min_{v\in F(t,\varphi)} \langle v,\xi\rangle \le 0 \quad \text{for all } (\xi_0,\xi,-q) \in N^P_{\text{Gr}P}(t,\varphi,\alpha). \tag{4.8}$$

Set $s = T-t$ and observe that, in view of the definition of GrP, the condition $(T-s,\varphi,\alpha) \in \text{Gr}P$ implies $(s,\varphi,\alpha) \in \text{Gr}\tilde{P}$. Furthermore, an easy computation of the proximal normal cone $N^P_{\text{Gr}P}$ shows that, for every $(T-s,\varphi,\alpha) \in \text{Gr}P$,

$$\text{if } \ (\xi_0,\xi,-q) \in N^P_{\text{Gr}P}(T-s,\varphi,\alpha), \ \text{ then } \ (-\xi_0,\xi,-q) \in N^P_{\text{Gr}\tilde{P}}(s,\varphi,\alpha).$$

On the other hand, in view of Proposition 4.12 and Theorem 4.10, there exists a full measure interval $I_2 \subseteq [0,T]$ such that the following conditions hold

$$\begin{aligned} &-\xi_0 + \max_{v\in \tilde{F}(s,\varphi)} \langle v,\xi\rangle \le 0 \qquad \text{for all } \ (-\xi_0,\xi,-q) \in N^P_{\text{Gr}\tilde{P}}(s,\varphi,\alpha), \\ &V(T,\varphi) = \mathcal{J}(\varphi), \end{aligned}$$

one can rewrite the conditions above as

$$\begin{aligned} &-\xi_0 + \max_{v\in -F(t,\varphi)} \langle v,\xi\rangle \le 0 \qquad \text{for all } \ (-\xi_0,\xi,-q) \in N^P_{\text{Gr}\tilde{P}}(T-t,\varphi,\alpha), \\ &V(T,\varphi) = \mathcal{J}(\varphi). \end{aligned} \tag{4.9}$$

In particular the relations (4.8) and (4.9) imply that there exists a full measure interval $I_3 := I_1 \bigcap I_2 \subseteq [0,T]$ such that, for every $(t,\varphi,\alpha) \in \text{epi}V \cap \big(I_3 \times L^2(\mu,\Omega;\mathbb{R}^n) \times \mathbb{R}\big)$, the following conditions hold:

$$\begin{aligned} &\xi_0 + \min_{v\in F(t,\varphi)} \langle v,\xi\rangle = 0 \qquad \text{for all } \ (\xi_0,\xi,-q) \in N^P_{\text{epi}V}(t,\varphi,\alpha), \\ &V(T,\varphi) = \mathcal{J}(\varphi). \end{aligned} \tag{4.10}$$

From the properties of the proximal cone $N^P_{\text{epi}V}$ (see, *e.g.*, [19]), we have that $q \ge 0 : \alpha = V(t,\varphi)$ if $q > 0$, and $q = 0$ if $\alpha > V(t,\varphi)$. In this last case, it follows that

$$(\xi_0,\xi,0) \in N^P_{\text{epi}V}(t,\varphi,V(t,\varphi)), \tag{4.11}$$

and from (4.6), we deduce that there exists a full-measure interval $I_4 \subset [0,T]$ such that (4.11) does not hold. Thus, we set $I := I_3 \cap I_4$ and consider the remaining case $q > 0$. Define $\xi_t := \xi_0/q$ and $\xi_\varphi := \xi/q$. The convexity of the proximal cone ensures that

$$(\xi_t,\xi_\varphi,-1) \in N^P_{\text{epi}V}(t,\varphi,V(t,\varphi)) \tag{4.12}$$

and since the Hamiltonian $\min_{v \in F(t,\varphi)} \langle v, \xi \rangle$ is positively homogeneous with respect ξ, the relations (4.10) and (4.12) imply the condition (4.7), for every $(t, \varphi) \in \operatorname{dom} V \cap (I \times L^2(\mu, \Omega; \mathbb{R}^n))$.

Now, suppose that $\theta : [0, T] \times L^2(\mu, \Omega; \mathbb{R}^n) \to (-\infty, \infty]$ is any other lower semi-continuous, bounded from below function satisfying the relation (4.7) and such that the set-valued map

$$t \rightsquigarrow P_\theta(t) := \big\{(\varphi, \alpha) \in L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R} : \; \theta(t, \varphi) \leq \alpha \big\}$$

is absolutely continuous. Then, in view of Theorems 4.9, 4.10 and 4.13, $\theta \equiv V$. This concludes the proof. □

Acknowledgements

This work was carried out while Oscar Sierra Fonseca was a postdoctoral fellow at FGV EMAp, holding a FAPERJ *Nota 10* Scholarship. We are grateful to both FAPERJ and FGV EMAp for their generous support. We would also like to express our sincere gratitude to the reviewer for their thorough reading and insightful comments, which greatly helped us improve the manuscript. The first and third authors acknowledge the support of FAPERJ (Rio de Janeiro, Brazil) via processes SEI-260003/004623/2021 and SEI-260003/000175/2024; and CNPq (Brazil) through process code 312407/2023-8. The second author gratefully acknowledge the partial support by PRIN2022-PNRR "Some mathematical approaches to climate change and its impacts", by the PNRR project "National Centre for HPC, Big Data and Quantum Computing" (CUP: E13C22001000006) and by the PRIN2022 project Progr-2022238YY5-PE1, "Optimal control problems: analysis, approximations and applications". Furthermore, the second author would like to acknowledge the financial support of GNAMPA-INdAM and, in particular, the INdAM-GNAMPA project "Optimal Control and Machine Learning" (CUP_E5324001950001)

Data availability statement

This statement is not applicable, as the study involved neither the creation nor the analysis of new data.

References

[1] J. Lohéac and E. Zuazua, From averaged to simultaneous controllability. *Ann. Fac. Sci. Toulouse Math.* **25** (2016) 785–828.

[2] E. Zuazua, Averaged control. *Automatica* **50** (2014) 3077–3087.

[3] P. Bettiol and N. Khalil, Necessary optimality conditions for average cost minimization problems. *Discrete Continuous Dyn. Syst. B* **24** (2019) 2093–2124.

[4] N. Khalil, *Optimality conditions for optimal control problems and applications.* PhD thesis, Université de Bretagne occidentale-Brest (2017).

[5] I. Ross, R.J. Proulx, M. Karpenko and Q. Gong, Riemann–Stieltjes optimal control problems for uncertain dynamic systems. *J. Guidance Control Dyn.* **38** (2015) 1251–1263.

[6] I. Ross, M. Karpenko and R. Proulx, A Lebesgue–Stieltjes framework for optimal control and allocation. *Proc. Am. Control Conf.* **2015** (2015) 5599–5604.

[7] C. Phelps, J.O. Royset and Q. Gong, Sample average approximations in optimal control of uncertain systems. *52nd IEEE Conference on Decision and Control, Firenze, Italy* (2013) 1958–1965.

[8] A. Pesare, M. Palladino and M. Falcone, Convergence results for an averaged LQR problem with applications to reinforcement learning. *Math. Control Signals Syst.* **33** (2021) 379–411.

[9] M. Palladino, A. Pesare and T. Scarinci, Convergence results for control problems with unknown dynamic and applications to reinforcement learning. *Math. Control Related Fields* (2005) 1–21. DOI: https://doi.org/10.3934/mcrf.2025030

[10] A. Scagliotti, Optimal control of ensembles of dynamical systems. *ESAIM: Control Optim. Calc. Var.* **29** (2023) 22.

[11] R. Vinter, Minimax optimal control. *SIAM J. Control Optim.* **44** (2005) 939–968.

[12] F. Clarke and Y. Ledyaev, Mean value inequalities in Hilbert Space. *Trans. Am. Math. Soc.* **344** (1994) 307–324.

[13] T. Donchev, Properties of the reachable set of control systems. *Syst. Control Lett.* **46** (2002) 379–386.

[14] F. Clarke, Yu.S. Ledyaev, R.J. Stern and P.R. Wolenski, Qualitative properties of trajectories of control systems: a survey. *J. Dyn. Control Syst.* **1** (1995) 1–48.

[15] J.-P. Aubin and A. Cellina, *Differential inclusions*, vol. 264 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (1984). Set-valued maps and viability theory.

[16] H. Frankowska, Lower semicontinuous solutions of Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* **31** (1993) 257–272.

[17] H. Frankowska, S. Plaskacz and T. Rzezuchowski, Measurable viability theorems and the Hamilton–Jacobi–Bellman equation. *J. Differ. Equ.* **116** (1995) 265–305.

[18] J.-P. Aubin and H. Frankowska, Set-valued analysis. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA (2009). Reprint of the 1990 edition [MR1048347].

[19] F. Clarke, Yu.S. Ledyaev, R.J. Stern and P.R. Wolenski, Nonsmooth Analysis and Control Theory. Springer-Verlag, Berlin, Heidelberg (1998).

[20] H.O. Fattorini. Infinite Dimensional Optimization and Control Theory. Encyclopedia of Mathematics and its Applications. Cambridge University Press (1999).

[21] H. Frankowska, A priori estimates for operational differential inclusions. *J. Differ. Equ.* **84** (1990) 100–128.

[22] X. Li and J. Yong, Optimal Control Theory for Infinite Dimensional Systems. Systems & Control: Foundations & Applications. Birkhäuser Boston (2012).

[23] P. Hajłasz and P. Koskela. Sobolev Met Poincaré. *Mem. Amer. Math. Soc.* **145** (2000) ix–101.

[24] A. Kałamajska, On compactness of embedding for Sobolev spaces defined on metric spaces. *Ann. Acad. Sci. Fennicae Math.* **24** (1999) 123–132.

[25] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions.* U.E.R. de Mathématiques (1975).

[26] L.C. Evans, *Partial Differential Equations.* Graduate studies in mathematics. American Mathematical Society (2010).

[27] A. Tolstonogov, Differential Inclusions in a Banach Space. Springer Netherlands (2000).

[28] P. Cannarsa and H. Frankowska, Value function and optimality conditions for semilinear control problems. *Appl. Math. Optim.* **26** (1992) 139–169.

[29] F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264. Springer (2013).

[30] J. Heinonen, P. Koskela, N. Shanmugalingam and J.T. Tyson, *Sobolev Spaces on Metric Measure Spaces: An Approach Based on Upper Gradients.* New Mathematical Monographs. Cambridge University Press (2015).

[31] P. Hajłasz, Sobolev spaces on an arbitrary metric space. *Potential Anal.* **5** (1996) 403–415.

[32] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer Science & Business Media (2010).

[33] M.G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton–Jacobi equations. *Bull. American Mathematical Society* **277** (1983) 1–42.

[34] Z. Artstein and K. Prikry, Carathéodory selections and the Scorza Dragoni property. *J. Math. Anal. Appl.* **127** (1987) 540–547.

[35] R. Lucchetti and F. Patrone, On Nemytskii's operator and its application to the lower semicontinuity of integral functionals. *Indiana Univ. Math. J.* **29** (1980) 703–713.

APPENDIX A. PROOF OF THEOREM 2.2: THE DYNAMIC PROGRAMMING PRINCIPLE

Proof. i). Suppose by contradiction that the assertion in i) of Theorem 2.2 does not hold. Then there exist $s_1, s_2 \in [0,T]$, $s_1 < s_2$, a control $u_1 \in \mathcal{U}_{\text{ad}}[s_1, s_2]$ and $\varepsilon > 0$ such that

$$V(s_1, \varphi) > V(s_2, x_{\varphi,u_1}(s_2;\cdot)) + \varepsilon. \tag{A.1}$$

On the other hand, by characterization of the infimum, there exists a control $u_2 \in \mathcal{U}_{\text{ad}}[s_2, T]$ such that

$$V(s_2, x_{\varphi,u_1}(s_2;\cdot)) > J_{[s_2,T]}(x_{\varphi,u_1}(s_2,\cdot), u_2) - \frac{\varepsilon}{2}.$$

Consider the control $u \in \mathcal{U}_{\text{ad}}[s_1, T]$ obtained from concatenating u_1 and u_2 :

$$u(s) := \begin{cases} u_1(s) & s \in [s_1, s_2), \\ u_2(s) & s \in [s_2, T]. \end{cases}$$

It easily follows from the definition of value function that

$$\begin{aligned} V(s_1, \varphi) \le J_{[s_1,T]}(\varphi, u) &= J_{[s_2,T]}(x_{\varphi,u_1}(s_2,\cdot), u_2) \\ &< V(s_2, x_{\varphi,u_1}(s_2;\cdot)) + \frac{\varepsilon}{2} < V(s_1,\varphi) - \frac{\varepsilon}{2}, \end{aligned} \tag{A.2}$$

which leads a contradiction. This completes the proof of i).

ii) It only remains to prove that

$$V(s_1, \varphi) \ge \inf_{u \in \mathcal{U}_{\text{ad}}[s_1,T]} V(s_2, x_{\varphi,u}(s_2,\cdot)).$$

Let $\varepsilon > 0$ and $u_\varepsilon \in \mathcal{U}_{\text{ad}}[s_1, T]$ be a control such that

$$J_{[s_1,T]}(\varphi, u_\varepsilon) \le V(s_1, \varphi) + \varepsilon.$$

Observe that

$$V(s_2, x_{\varphi,u_\varepsilon}(s_2,\cdot)) \le J_{[s_2,T]}(x_{\varphi,u_\varepsilon}(s_2,\cdot), u_\varepsilon).$$

Then

$$\begin{aligned} \inf_{u \in \mathcal{U}_{\text{ad}}[s_1,T]} V(s_2, x_{\varphi,u}(s_2,\cdot)) &\le V(s_2, x_{\varphi,u_\varepsilon}(s_2,\cdot)) \\ &\le J_{[s_2,T]}(x_{\varphi,u_\varepsilon}(s_2,\cdot), u_\varepsilon) \\ &= J_{[s_1,T]}(\varphi, u_\varepsilon) \le V(s_1,\varphi) + \varepsilon. \end{aligned}$$

This yields the desired inequality and completes the proof. □

APPENDIX B. MEASURABLE SELECTIONS AND INTEGRAL FUNCTIONALS

Let $\mathbf{T}$ and X be two topological spaces. Let 2^X be the set of all nonempty subsets of X. We call any map $\Gamma : \mathbf{T} \to 2^X$ a multifunction, we will denote it as $\Gamma : \mathbf{T} \rightsquigarrow X$.

Definition B.1 (Continuity of set-valued maps)**.** A multifunction $\Gamma : \mathbf{T} \rightsquigarrow X$ is said to be *upper semi-continuous* if, for any closed subset $F \subset X$, the set

$$\Gamma^{-1}(F) \equiv \{t \in \mathbf{T} \mid \Gamma(t) \bigcap F \neq \emptyset\}$$

is closed in $\mathbf{T}$. Γ is said to be *lower semi-continuous* if for any open subset $U \subset X$, $\Gamma^{-1}(U)$ is open in $\mathbf{T}$. Finally, Γ is said to be *continuous* if it is both upper and lower semi-continuous.

When Γ is single-valued, the above three kinds of continuities are equivalent.

Definition B.2 (Polish space)**.** A topological space X is called a *Polish space* if it has a countable base, is metrizable, and the space is complete under a metric compatible with the topology of the space.

Definition B.3 (Souslin space)**.** A topological space X is called a *Souslin space* if it is Hausdorff and there exists a Polish space P and a continuous map $g : P \to X$, such that

$$g(P) = X.$$

A subset $A \subset X$ is said to be *Souslinian* if as a subspace, A is a Souslin space. Sometimes, we also call such an A a *Souslin set.*

Definition B.4. A multifunction $\Gamma : \mathbf{T} \rightsquigarrow X$ is said to be *Lebesgue (resp. Borel, Souslin) measurable* if for any closed set $F \subset X$, the set $\Gamma^{-1}(F)$ is Lebesgue (resp. Borel, Souslin) set in $\mathbf{T}$.

Definition B.5. A multifunction $\Gamma : \mathbf{T} \rightsquigarrow X$ is said to be *almost upper semicontinuous* (resp. *almost lower semicontinuous*) if, for every $\varepsilon > 0$, there exists a compact subset $K_\varepsilon \subset \mathbf{T}$ with $\operatorname{meas}(\mathbf{T} \setminus K_\varepsilon) < \varepsilon$ such that the restriction $\Gamma|_{K_\varepsilon}$ is upper semicontinuous (resp. lower semicontinuous).

Definition B.6 (Pseudocontinuity of set-valued functions)**.** A multifunction $U : \mathbf{T} \rightsquigarrow X$ is said to be *pseudo-continuous* at $t \in \mathbf{T}$ if

$$\bigcap_{\varepsilon > 0} \overline{U\left(\mathcal{B}_\varepsilon(t)\right)} = U(t).$$

We say that U is pseudo-continuous on $\mathbf{T}$ if it is pseudo-continuous at each point $t \in \mathbf{T}$.

Proposition B.7. *Let $\mathbf{U} \subset X$ be a compact metric space and $U : \mathbf{T} \rightsquigarrow \mathbf{U}$ be a multifunction taking closed set values. Then U is upper semi-continuous if and only if it is pseudo-continuous.*

Theorem B.8. *[Filippov's Lemma [22], p. 102] Let $U : \mathbf{T} \rightsquigarrow X$ be measurable taking closed set values and Y a Polish space. Let $f : \mathbf{T} \times X \to Y$ be Souslin measurable and for each $\bar{u} \in X$, $f(\cdot, \bar{u})$ is measurable; for almost all $t \in \mathbf{T}$, $f(t, \cdot)$ is continuous. Let $y : \mathbf{T} \to Y$ be Lebesgue measurable satisfying*

$$y(t) \in f(t, U(t)), \qquad a.e.\ t \in \mathbf{T}.$$

Then there exists a measurable function $u : \mathbf{T} \to X$, such that

$$\begin{cases} u(t) \in U(t) & a.e.\ t \in \mathbf{T}, \\ y(t) = f(t, u(t)) & a.e.\ t \in \mathbf{T}. \end{cases}$$

Corollary B.9. *Let $\mathbf{U} \subset \mathbb{R}^m$ be a compact subset and consider $f : [0,T] \times L^2(\mu, \Omega; \mathbb{R}^n) \times \mathbb{R}^m \to L^2(\mu, \Omega; \mathbb{R}^n)$ and let $U : [0,T] \rightsquigarrow \mathbf{U}$ be a multifunction taking closed set values. If for each $\bar{u} \in \mathbf{U}$, $f(\cdot, \cdot, \bar{u})$ is measurable; for almost all $t \in [s,T]$, $f(t, \cdot, \cdot)$ is continuous and*

$$\dot{x}(t) \in f(t, x(t), U(t)) \quad a.e.\ t \in [s,T],$$

then, there exists a measurable function $u : [s,T] \to \mathbb{R}^m$, such that

$$\begin{cases} u(t) \in U(t), & a.e.\ t \in [s,T], \\ \dot{x}(t) = f(t, x(t), u(t)), & a.e.\ t \in [s,T]. \end{cases}$$

Proof. It follows from Theorem B.8. □

Proposition B.10. *[22] Let $U : [0,T] \rightsquigarrow \mathbb{R}^m$ be a pseudo-continuous multifunction and*

$$f : [0,T] \times L^2(\mu,\Omega;\mathbb{R}^n) \times \mathbb{R}^m \longrightarrow L^2(\mu,\Omega;\mathbb{R}^n)$$

uniformly continuous on $(\varphi,u) \in L^2(\mu,\Omega;\mathbb{R}^n) \times \mathbb{R}^m$ for any $t \in [0,T]$. Then $f(t,\varphi,U(t))$ is closed and convex if and only if, for almost all $t \in [0,\infty)$, the set $f(t,\varphi,U(t))$ satisfies the following:

$$\bigcap_{\delta>0} \overline{\mathrm{co}} f\left(t, \mathcal{B}_\delta(\varphi), U\left(\mathcal{B}_\delta(t)\right)\right) = f(t,\varphi,U(t)), \tag{B.1}$$

where $\overline{\mathrm{co}}\, S$ is the smallest convex and closed set containing S, i.e., $\overline{\mathrm{co}}\, S$ denotes the convex hull of S.

Remark B.11. Let f be as in Proposition B.10. If $U : [0,T] \rightsquigarrow \mathbf{U}$ is a multifunction taking nonempty closed values, included in a compact set $\mathbf{U} \subset \mathbb{R}^m$. Then from the compactness of $U(t)$ for all $t \in [0,T]$ and the continuity of f with respect to u, it follows that $f(t,\varphi,U(t))$ is compact in $L^2(\mu,\Omega;\mathbb{R}^n)$, for almost all $t \in [0,T]$ and all $\varphi \in L^2(\mu,\Omega;\mathbb{R}^n)$.

Let $g : \mathbb{R}^n \times \Omega \to (-\infty,+\infty]$ be a normal integrand (see $(\mathbf{H})$(iv) in Sect. 1.1) and, for $p < \infty$, consider the functional $\mathcal{J}$ (see (3.1)) over the space $L^p(\mu,\Omega;\mathbb{R}^n)$. The following result is a characterization of the lower semi-continuity of $\mathcal{J}$.

Theorem B.12. *[35] Let $\mathcal{J}$ be not identically equal to $+\infty$ and let $p < \infty$. If g is a normal integrand, then the following properties are equivalent:*

(1) $\mathcal{J}$ is strongly lower semi-continuous on $L^p(\mu,\Omega;\mathbb{R}^n)$ and

$$\mathcal{J}(\varphi) > -\infty \quad \forall\, \varphi \in L^p(\mu,\Omega;\mathbb{R}^n).$$

(2) $\mathcal{J} : L^p(\mu,\Omega;\mathbb{R}^n) \to (-\infty,+\infty]$.

(3) $\mathcal{J} : L^p(\mu,\Omega;\mathbb{R}^n) \to [-\infty,+\infty]$ and there exist $k \in \mathbb{R}, h \geq 0$ such that

$$\mathcal{J}(\varphi) \geq k - h\|\varphi(\cdot)\|^p_{L^p(\mu,\Omega;\mathbb{R}^n)} \qquad \textit{for every } \varphi \in L^p(\mu,\Omega;\mathbb{R}^n).$$

(4) There exist $a \in L^1(\mu,\Omega)$, $b \geq 0$ such that

$$g(x,\omega) \geq a(\omega) - b|x|^p. \;\; \textit{for every } x \in \mathbb{R}^n \;\; \textit{and a.e. } \omega \in \Omega.$$

If g is not a normal integrand, then the following hold: (2) ⇔ (3) ⇔ (4) and (1) ⇒ (2).

Appendix C. Compactness results

The following is a criterion for the relative compactness in the Banach space

$$L^p(\mu,\Omega) := \left\{ f : \Omega \to \mathbb{R} : \int_\Omega |f(\omega)|^p \,\mathrm{d}\mu(\omega) < \infty \right\}.$$

Theorem C.1. *[24] Let Ω be a metric space equipped with a finite Borel measure μ such that, for any $r > 0$, $h(r) > 0$ (see $(\mathbf{H}_\mu)$). Then, every bounded sequence $\{f_k\} \subset L^p(\mu,\Omega)$, with $1 \leq p < \infty$ such that*

$$\sup_k \int_\Omega \left| f_k(\omega) - (f_k)_{\mathcal{B}_r(\omega)} \right|^p \mathrm{d}\mu(\omega) \xrightarrow{r\to 0} 0$$

where $(f)_{\mathcal{B}_r(\omega)} := \frac{1}{\mu(\mathcal{B}_r(\omega))} \int_{\mathcal{B}_r(\omega)} f(\omega')\,\mathrm{d}\mu(\omega')$, is relatively compact in $L^p(\mu,\Omega)$.

The following result is a consequence of the Arzelà-Ascoli and Alaoglu Theorems.

Theorem C.2. *[15], p. 13 Let us consider a sequence of absolutely continuous functions $x_k(\cdot)$ from an interval I of $\mathbb{R}$ to a Banach space X satisfying*

$$\begin{cases} \textit{i) for all } t \in I,\ \{x_k(t)\}_k \ \textit{ is a relatively compact subset of } X, \\[2ex] \textit{ii) there exists a positive function } \ c(\cdot) \in L^1(I) \ \textit{ such that} \\ \qquad \textit{for almost all } t \in I,\ \|x_k'(t)\| \le c(t). \end{cases}$$

Then there exists a subsequence (again denoted by $x_k(\cdot)$) converging to an absolutely continuous function $x(\cdot)$ from I to X in the sense that

$$\begin{cases} \textit{i)} \ \ x_k(\cdot) \ \textit{converges uniformly to } x(\cdot) \textit{ over compact subsets of } I, \\[2ex] \textit{ii)} \ x_k'(\cdot) \ \textit{converges weakly to } x'(\cdot) \ \textit{in } L^1(I, X). \end{cases}$$