



Compatibility of state constraints and dynamics for multiagent control systems

GIULIA CAVAGNARI , ANTONIO MARIGONDA AND MARC QUINCAMPOIX

Abstract. This study concerns the problem of compatibility of state constraints with a multiagent control system. Such a system deals with a number of agents so large that only a statistical description is available. For this reason, the state variable is described by a probability measure on \mathbb{R}^d representing the density of the agents and evolving according to the so-called continuity equation which is an equation stated in the Wasserstein space of probability measures. The aim of the paper is to provide a necessary and sufficient condition for a given constraint (a closed subset of the Wasserstein space) to be compatible with the controlled continuity equation. This new condition is characterized in a viscosity sense as follows: the distance function to the constraint set is a viscosity supersolution of a suitable Hamilton–Jacobi–Bellman equation stated on the Wasserstein space. As a byproduct and key ingredient of our approach, we obtain a new comparison theorem for evolutionary Hamilton–Jacobi equations in the Wasserstein space.

1. Introduction

In classical control theory, a single agent controls a dynamics (here represented by a differential inclusion)

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0, \quad (1.1)$$

where $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is a set valued map, associating with each $x \in \mathbb{R}^d$ the subset $F(x)$ of \mathbb{R}^d of the admissible velocities from x . A *multiagent system* involves a large number of agents having all a dynamics of the form (1.1). In this model, the number of agents is so large that at each time only a *statistical (macroscopic) description* of the state is available. We thus describe the configuration of the system at time t by a Borel measure μ_t on \mathbb{R}^d , where for every Borel set $A \subseteq \mathbb{R}^d$ the quotient $\frac{\mu_t(A)}{\mu_t(\mathbb{R}^d)}$ represents the fraction of the total amount of agents that are present in A at the time t . Since the total amount of agents is supposed to be fixed in time, $\mu_t(\mathbb{R}^d)$ is constant, and thus, we choose to normalize the measure μ_t assuming $\mu_t(\mathbb{R}^d) = 1$, i.e., $\mu_t \in \mathcal{P}(\mathbb{R}^d)$, the set of Borel probability measures on \mathbb{R}^d .

Hence, the evolution of the controlled multi-agent system can be represented by the following two-scale dynamics

Mathematics Subject Classification: 49J15, 49J52, 49L25, 49K27, 49K21, 49Q20, 34A60, 93C15

Keywords: Optimal control, Optimal transport, Hamilton–Jacobi–Bellman equation, Multi-agent.

- *Microscopic dynamics*: each agent's position at time t is given by $x(t)$, which evolves according to the dynamical system

$$\dot{x}(t) \in F(\mu_t, x(t)), \quad \text{for a.e. } t > 0, \quad (1.2)$$

where F is a set-valued map. It is worth pointing out that each agent's dynamics is nonlocal since it depends also on the instantaneous configuration μ_t of the crowd of agents at time t , described by a probability measure on \mathbb{R}^d .

- *Macroscopic dynamics*: the configuration of the crowd of agents at time t is given by a time-depending measure $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ whose evolution satisfies the following *continuity equation* (to be understood in the sense of distributions)

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \quad t > 0, \quad (1.3)$$

coupled with the control constraint

$$v_t(x) \in F(\mu_t, x) \text{ for } \mu_t\text{-a.e. } x \in \mathbb{R}^d \text{ and for a.e. } t \geq 0. \quad (1.4)$$

which represents the possible (Eulerian) velocity $v_t(x)$ chosen by an external planner for an agent at time t and at the position x .

The investigation of (deterministic) optimal control problems in the space of measures is attracting an increasing interest by the mathematical community in the last years, due to the potential applications in the study of complex systems, or multi-agent systems (see, e.g., [16, 18, 19]). Indeed, in the framework of *mean field approximation* of multi agent system, starting from a control problem for a large number of the (discrete) agents, the problem is recasted in the framework of probability measures (see the recent [15] or the preprint [12] for Γ -convergence results for optimal control problems with nonlocal dynamics). This procedure reduces the dimensionality and the number of equations, possibly leading to a simpler and treatable problem from the point of view of numerics. The reader can find a comprehensive overview of the literature about such formulations and applications, together with some insights on research perspective, in the recent survey [1], and references therein. We refer to [7] for further results on mean field control problems.

The problem we address in this paper is the compatibility of the above dynamical system (1.3)–(1.4) with a given closed constraint $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$. Here, $\mathcal{P}_2(\mathbb{R}^d)$ is the set of Borel probability measures with finite second moment; this set is equipped with the 2-Wasserstein distance (see Sect. 2). This compatibility property could be understood in two ways

- \mathcal{K} is *viable* for the dynamics F if and only if for any $\mu \in \mathcal{K}$ there exists a solution $t \mapsto \mu_t$ of the controlled continuity Eqs. (1.3)–(1.4) with $\mu_0 = \mu$ such that $\mu_t \in \mathcal{K}$ for all $t \geq 0$;
- \mathcal{K} is *invariant* for the dynamics F if and only if for any $\mu \in \mathcal{K}$ and for any solution $t \mapsto \mu_t$ of the controlled continuity Eqs. (1.3)–(1.4) with $\mu_0 = \mu$ we have $\mu_t \in \mathcal{K}$ for all $t \geq 0$.

Inspired by a characterization of the viability property via supersolution of Hamilton–Jacobi–Bellman equations, which was first obtained in [9] in the framework of stochastic control, we develop an approach for the present multiagent control problem with deterministic dynamics (1.3)–(1.4).

The main result of our paper (Theorems 6.6 and 6.7) can be roughly summarized as follows

Theorem 1.1. *Let $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ be a closed set and $d_{\mathcal{K}}$ the associated distance function. Assume that the set valued map F is L -Lipschitz.*

- \mathcal{K} is viable iff the function $\mu \mapsto d_{\mathcal{K}}(\mu)$ is a viscosity supersolution of

$$(L + 2)u(\mu) + \mathcal{H}_F^{\text{viab}}(\mu, D_{\mu}u(\mu)) = 0,$$

where, for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $p \in L_{\mu}^2(\mathbb{R}^d; \mathbb{R}^d)$,

$$\mathcal{H}_F^{\text{viab}}(\mu, p) := -d_{\mathcal{K}}(\mu) - \inf_{\substack{v(\cdot) \in L_{\mu}^2(\mathbb{R}^d) \\ v(x) \in F(\mu, x) \mu\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle v(x), p(x) \rangle d\mu(x).$$

- \mathcal{K} is invariant iff the function $\mu \mapsto d_{\mathcal{K}}(\mu)$ is a viscosity supersolution of

$$(L + 2)u(\mu) + \mathcal{H}_F^{\text{inv}}(\mu, D_{\mu}u(\mu)) = 0,$$

where, for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $p \in L_{\mu}^2(\mathbb{R}^d; \mathbb{R}^d)$,

$$\mathcal{H}_F^{\text{inv}}(\mu, p) := -d_{\mathcal{K}}(\mu) - \sup_{\substack{v(\cdot) \in L_{\mu}^2(\mathbb{R}^d) \\ v(x) \in F(\mu, x) \mu\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle v(x), p(x) \rangle d\mu(x).$$

For a completely different approach to the viability problem, we refer to [5], where the author provides a characterization of the viability property for a closed set $\mathcal{K} \subseteq \mathcal{P}_1(\mathbb{T}^d)$ by mean of a condition involving a suitable notion of tangent cone to \mathcal{K} in the Wasserstein space $\mathcal{P}_1(\mathbb{T}^d)$, where \mathbb{T}^d denotes the d -dimensional torus.

The paper is organized as follows: in Sect. 2, we fix the notations and provide some background results; Sect. 3 is devoted to the properties of the set of solutions of the controlled continuity Eqs. (1.3)–(1.4); Sect. 4 establishes the link between the viability/invariance problem with the value function of a suitable control problem in Wasserstein space; Sect. 5 introduces the viscosity solutions of Hamilton–Jacobi–Bellman equations in the Wasserstein space, together with a uniqueness result; in Sect. 6, we apply the results of Sect. 5 to the problem outlined in Sect. 4 deriving our main characterization of viability/invariance. Finally, in Sect. 7 we provide an example illustrating the main results. Some proofs of technical results are postponed to “Appendix.”

2. Notations

Throughout the paper, we will use the following notation and we address to [2] as a relevant resource for preliminaries on measure theory.

$B(x, r)$	the open ball of radius r of a metric space (X, d_X) , i.e., $B(x, r) := \{y \in X : d_X(y, x) < r\}$;
\overline{K}	the closure of a subset K of a topological space X ;
$d_K(\cdot)$	the distance function from a subset K of a metric space (X, d) , i.e., $d_K(x) := \inf\{d(x, y) : y \in K\}$;
$C_b^0(X; Y)$	the set of continuous bounded functions from a Banach space X to Y , endowed with $\ f\ _\infty = \sup_{x \in X} f(x) $ (if $Y = \mathbb{R}$, Y will be omitted);
$C_c^0(X; Y)$	the set of compactly supported functions of $C_b^0(X; Y)$, with the topology induced by $C_b^0(X; Y)$;
$BUC(X; \mathbb{R})$	the space of bounded real-valued uniformly continuous functions defined on X
Γ_I	the set of continuous curves from a real interval I to \mathbb{R}^d ;
Γ_T	the set of continuous curves from $[0, T]$ to \mathbb{R}^d ;
e_t	the evaluation operator $e_t : \mathbb{R}^d \times \Gamma_I \rightarrow \mathbb{R}^d$ defined by $e_t(x, \gamma) = \gamma(t)$ for all $t \in I$;
$\mathcal{P}(X)$	the set of Borel probability measures on a Banach space X , endowed with the weak* topology induced from $C_b^0(X)$;
$\mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$	the set of vector-valued Borel measures on \mathbb{R}^d with values in \mathbb{R}^d , endowed with the weak* topology induced from $C_c^0(\mathbb{R}^d; \mathbb{R}^d)$;
$ v $	the total variation of a measure $v \in \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$;
\ll	the absolutely continuity relation between measures defined on the same σ -algebra;
$m_2(\mu)$	the second moment of a probability measure $\mu \in \mathcal{P}(X)$;
$r\#\mu$	the push-forward of the measure μ by the Borel map r ;
$\mu \otimes \pi_x$	the product measure of $\mu \in \mathcal{P}(X)$ with the Borel family of measures $\{\pi_x\}_{x \in X} \subseteq \mathcal{P}(Y)$ (see Theorem 2.1);
pr_i	the i -th projection map $\text{pr}_i(x_1, \dots, x_N) = x_i$;
$\Pi(\mu, \nu)$	the set of admissible transport plans from μ to ν ;
$\Pi_o(\mu, \nu)$	the set of optimal transport plans from μ to ν ;
$W_2(\mu, \nu)$	the 2-Wasserstein distance between μ and ν ;
$\mathcal{P}_2(X)$	the subset of the elements $\mathcal{P}(X)$ with finite second moment, endowed with the 2-Wasserstein distance;
$\frac{\nu}{\mu}$	the Radon–Nikodym derivative of the measure ν w.r.t. the measure μ ;
$\text{Lip}(f)$	the Lipschitz constant of a function f ;
$(f)^+$	the positive part of a real valued function f , i.e., $(f)^+ = \max\{0, f\}$.

Given Banach spaces X, Y , we denote by $\mathcal{P}(X)$ the set of Borel probability measures on X endowed with the weak* topology induced by the duality with the Banach space $C_b^0(X)$ of the real-valued continuous bounded functions on X with the uniform convergence norm. The second moment of $\mu \in \mathcal{P}(X)$ is defined by $m_2(\mu) = \int_X \|x\|_X^2 d\mu(x)$, and we set $\mathcal{P}_2(X) = \{\mu \in \mathcal{P}(X) : m_2(\mu) < +\infty\}$. For any Borel map $r : X \rightarrow Y$ and $\mu \in \mathcal{P}(X)$, we define the *push forward measure* $r\#\mu \in \mathcal{P}(Y)$ by setting $r\#\mu(B) = \mu(r^{-1}(B))$ for any Borel set B of Y . In other words,

$$\int_Y \varphi(y) d[r\#\mu](y) = \int_X \varphi(r(x)) d\mu(x),$$

for any bounded Borel measurable function $\varphi : Y \rightarrow \mathbb{R}$.

We denote by $\mathcal{M}(X; Y)$ the set of Y -valued Borel measures defined on X . The total variation measure of $\nu \in \mathcal{M}(X; Y)$ is defined for every Borel set $B \subseteq X$ as

$$|\nu|(B) := \sup_{\{B_i\}_{i \in \mathbb{N}}} \left\{ \sum \|\nu(B_i)\|_Y \right\},$$

where the sup ranges on countable Borel partitions of B .

We now recall the definitions of transport plans and Wasserstein distance (cf. for instance Chapter 6 in [2]). Let X be a complete separable Banach space, $\mu_1, \mu_2 \in \mathcal{P}(X)$. The set of *admissible transport plans* between μ_1 and μ_2 is

$$\Pi(\mu_1, \mu_2) = \{\pi \in \mathcal{P}(X \times X) : \text{pr}_i\#\pi = \mu_i, i = 1, 2\},$$

where for $i = 1, 2$, $\text{pr}_i : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a projection $\text{pr}_i(x_1, x_2) = x_i$. The *Wasserstein distance* between μ_1 and μ_2 is

$$W_2^2(\mu_1, \mu_2) = \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{X \times X} |x_1 - x_2|^2 d\pi(x_1, x_2).$$

If $\mu_1, \mu_2 \in \mathcal{P}_2(X)$, then the above infimum is actually a minimum, and the set of minima is denoted by

$$\Pi_o(\mu_1, \mu_2) := \left\{ \pi \in \Pi(\mu_1, \mu_2) : W_2^2(\mu_1, \mu_2) = \int_{X \times X} |x_1 - x_2|^p d\pi(x_1, x_2) \right\}.$$

Recall that $\mathcal{P}_2(X)$ endowed with the W_2 -Wasserstein distance is a complete separable metric space.

The following result is Theorem 5.3.1 in [2].

Theorem 2.1. (Disintegration) *Let \mathbb{X}, X be complete separable metric spaces. Given a measure $\mu \in \mathcal{P}(\mathbb{X})$ and a Borel map $r : \mathbb{X} \rightarrow X$, there exists a Borel family of probability measures $\{\mu_x\}_{x \in X} \subseteq \mathcal{P}(\mathbb{X})$, uniquely defined for $r\#\mu$ -a.e. $x \in X$, such that $\mu_x(\mathbb{X} \setminus r^{-1}(x)) = 0$ for $r\#\mu$ -a.e. $x \in X$, and for any Borel map $\varphi : \mathbb{X} \rightarrow [0, +\infty]$ we have*

$$\int_{\mathbb{X}} \varphi(z) d\mu(z) = \int_X \left[\int_{r^{-1}(x)} \varphi(z) d\mu_x(z) \right] d(r\#\mu)(x).$$

We will write $\mu = (r\sharp\mu) \otimes \mu_x$. If $\mathbb{X} = X \times Y$ and $r^{-1}(x) \subseteq \{x\} \times Y$ for all $x \in X$, we can identify each measure $\mu_x \in \mathcal{P}(X \times Y)$ with a measure on Y .

3. Admissible trajectories

The goal of this section is to give a precise definition of the macroscopic dynamics (1.3, 1.4) and to study its trajectories. To maintain the flow of the paper, the proofs of the results of this section are postponed to ‘‘Appendix A.’’

Definition 3.1. (*Admissible trajectories*) Let $I = [a, b]$ be a closed real interval, $\mu = \{\mu_t\}_{t \in I} \subseteq \mathcal{P}_2(\mathbb{R}^d)$, $\nu = \{\nu_t\}_{t \in I} \subseteq \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$, $F : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued map.

We say that μ is an *admissible trajectory driven by* ν defined on I with underlying dynamics F if

- $|\nu_t| \ll \mu_t$ for a.e. $t \in I$;
- $\nu_t(x) := \frac{\nu_t}{\mu_t}(x) \in F(\mu_t, x)$ for a.e. $t \in I$ and μ_t -a.e. $x \in \mathbb{R}^d$;
- $\partial_t \mu_t + \operatorname{div} \nu_t = 0$ in the sense of distributions in $[a, b] \times \mathbb{R}^d$.

Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we define the set of *admissible trajectories* as

$$\mathcal{A}_I(\mu) := \left\{ \mu = \{\mu_t\}_{t \in I} : \exists \nu = \{\nu_t\}_{t \in I} \subseteq \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d) \text{ s.t. } \mu \text{ is an admissible traj. driven by } \nu, \text{ defined on } I \text{ with underlying dynamics } F \text{ and } \mu_a = \mu \right\}.$$

We make the following assumptions on the set-valued map F :

(F₁) $F : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ is continuous with convex, compact and nonempty images, where on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ we consider the metric

$$d_{\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d}((\mu_1, x_1), (\mu_2, x_2)) = W_2(\mu_1, \mu_2) + |x_1 - x_2|.$$

(F₂) there exists $L > 0$, a compact metric space U and a continuous map $f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ satisfying

$$|f(\mu_1, x_1, u) - f(\mu_2, x_2, u)| \leq L(W_2(\mu_1, \mu_2) + |x_1 - x_2|),$$

for all $\mu_i \in \mathcal{P}_2(\mathbb{R}^d)$, $x_i \in \mathbb{R}^d$, $i = 1, 2$, $u \in U$, such that the set-valued map F can be represented as

$$F(\mu, x) = \{f(\mu, x, u) : u \in U\}.$$

As pointed out also in Remark 2 of [16], from the Lipschitz continuity of the set-valued map F coming from assumption (F₂), we easily get

$$F(\mu, x) \subseteq F(\nu, y) + L(W_2(\mu, \nu) + |x - y|)\overline{B(0, 1)},$$

for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$. From which, for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, we have

$$F(\mu, x) \subseteq C(1 + m_2^{1/2}(\mu))(1 + |x|)\overline{B(0, 1)}, \quad (3.1)$$

where $C := \max\{1, L \max\{|y| : y \in F(\delta_0, 0)\}\}$.

Definition 3.2. Let $\{\mu^{(n)}\}_{n \in \mathbb{N}} \subseteq \text{AC}([a, b]; \mathcal{P}_2(\mathbb{R}^d))$. We say that $\{\mu^{(n)}\}_{n \in \mathbb{N}}$ uniformly W_2 -converges to μ , $\mu^{(n)} \rightrightarrows \mu$, if we have

$$\sup_{t \in [a, b]} W_2(\mu_t^{(n)}, \mu_t) \rightarrow 0.$$

We recall the following result taken from [16].

Lemma 3.3. (Grönwall-like estimate (Prop. 2 in [16])) *Assume $(F_1) - (F_2)$. Let $\mu_0, \theta_0 \in \mathcal{P}_2(\mathbb{R}^d)$, and $\mu = \{\mu_t\}_{t \in [a, b]} \in \mathcal{A}_{[a, b]}(\mu_0)$ an admissible trajectory. Then, there exists an admissible trajectory $\theta = \{\theta_t\}_{t \in [a, b]} \in \mathcal{A}_{[a, b]}(\theta_0)$, such that for all $t \in [a, b]$ we have*

$$W_2(\mu_t, \theta_t) \leq e^{L(b-a) + (b-a)e^{L(b-a)}} \cdot W_2(\mu_0, \theta_0),$$

where L is as in (F_2) .

Proposition 3.4. *Assume $(F_1) - (F_2)$. Let $\mu = \{\mu_t\}_{t \in [a, b]}$ be an admissible trajectory, with $0 \leq a < b < +\infty$. Then, there exists $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a, b]})$ such that $e_t \# \eta = \mu_t$ for all $t \in [a, b]$, and for η -a.e. (x, γ)*

$$\begin{cases} \dot{\gamma}(t) \in F(\mu_t, \gamma(t)), & \text{for a.e. } t \in [a, b], \\ \gamma(a) = x. \end{cases}$$

Moreover, for any η as above and for all $t, s \in [a, b]$ with $s < t$, we have

(1) for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a, b]}$,

$$\begin{aligned} & \frac{e_t - e_s}{t - s}(x, \gamma) \in F(\mu_s, \gamma(s)) + \\ & + \left[\frac{L}{t - s} \int_s^t [W_2(\mu_\tau, \mu_s) + |(e_\tau - e_s)(x, \gamma)|] d\tau \right] \cdot \overline{B(0, 1)}; \end{aligned}$$

(2) $\|e_t - e_s\|_{L_\eta^2} \leq e^{L(t-s)} \left[(t - s)(K + 2Lm_2^{1/2}(\mu_s)) + L \int_s^t W_2(\mu_\tau, \mu_s) d\tau \right] =: h(t, s);$

(3) $\lim_{t \rightarrow s^+} \left\| \frac{e_t - e_s}{t - s} \right\|_{L_\eta^2} = K + 2Lm_2^{1/2}(\mu_s),$

where $L = \text{Lip}(F)$ and $K = \max\{|y| : y \in F(\delta_0, 0)\}$.

In particular, there exists a Borel map $w : \mathbb{R}^d \times \Gamma_{[a, b]} \rightarrow \mathbb{R}^d$, with $w(x, \gamma) \in F(\mu_s, \gamma(s))$ for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a, b]}$, such that

$$\left\| \frac{e_t - e_s}{t - s} - w \right\|_{L_\eta^2} \leq \frac{L}{t - s} \int_s^t [W_2(\mu_\tau, \mu_s) + h(\tau, s)] d\tau.$$

Proposition 3.5. (Compactness of $\mathcal{A}_{[a,b]}(\mu)$) Assume $(F_1) - (F_2)$ and let $0 \leq a < b < +\infty$ and $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Then, the set of admissible trajectories $\mathcal{A}_{[a,b]}(\mu_0)$ is nonempty and compact w.r.t. uniform W_2 -convergence (see Definition 3.2).

4. Viability problem and the value function

Throughout the paper, let $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ be closed w.r.t. the metric W_2 . We are interested in the definitions of compatibility of our dynamics w.r.t. the state constraint given by \mathcal{K} (cf. introduction of the present paper).

Notice that, since concatenation of admissible trajectories is an admissible trajectory (see the note before Prop. 3 in [16]), if \mathcal{K} is viable (resp. invariant) in $[t_0, T]$ then it is viable (resp. invariant) in $[0, \hat{T}]$ for any $\hat{T} > T$.

As we will investigate in Sect. 5, the viability and invariance properties of a closed set $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ are closely related to the following optimal control problems, with fixed time-horizon $T > 0$.

Definition 4.1. (Value functions) Given $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ closed, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $t_0 \in [0, T]$, we set

(1) $V^{\text{viab}} : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty)$ as follows

$$V^{\text{viab}}(t_0, \mu) := \inf_{\mu \in \mathcal{A}_{[t_0, T]}(\mu)} \int_{t_0}^T d_{\mathcal{K}}(\mu_t) dt, \quad (4.1)$$

where $d_{\mathcal{K}} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty)$, $d_{\mathcal{K}}(\mu) := \inf_{\sigma \in \mathcal{K}} W_2(\mu, \sigma)$.

We say that $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$ is an *optimal trajectory* for V^{viab} starting from μ at time t_0 if it achieves the minimum in (4.1).

(2) $V^{\text{inv}} : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty)$ as follows

$$V^{\text{inv}}(t_0, \mu) := \sup_{\mu \in \mathcal{A}_{[t_0, T]}(\mu)} \int_{t_0}^T d_{\mathcal{K}}(\mu_t) dt. \quad (4.2)$$

We say that $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$ is an *optimal trajectory* for V^{inv} starting from μ at time t_0 if it achieves the maximum in (4.2).

The main interest in the above value functions lies in the fact that they give a characterization of the viability/invariance as explained in Proposition 4.3. We first state a regularity result of the above value functions and the existence of optimal trajectories.

Proposition 4.2. Assume $(F_1) - (F_2)$. Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $t_0 \in [0, T]$, there exist an optimal trajectory $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$ for V^{viab} and an optimal trajectory $\mu' \in \mathcal{A}_{[t_0, T]}(\mu)$ for V^{inv} .

Proof. We prove the existence of an optimal trajectory for V^{viab} . Take any $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$. By passing to the infimum over $\sigma \in \mathcal{K}$ on the triangular inequality

$$W_2(\mu^1, \sigma) \leq W_2(\mu^1, \mu^2) + W_2(\mu^2, \sigma),$$

we have $d_{\mathcal{K}}(\mu^1) \leq W_2(\mu^1, \mu^2) + d_{\mathcal{K}}(\mu^2)$. Reversing the roles of μ^1 and μ^2 , we get the 1-Lipschitz continuity of $d_{\mathcal{K}}$. Hence, by Fatou's Lemma, we get the l.s.c. of the cost functional, i.e.,

$$\int_{t_0}^T d_{\mathcal{K}}(\mu_t) dt \leq \liminf_{n \rightarrow +\infty} \int_{t_0}^T d_{\mathcal{K}}(\mu_t^{(n)}) dt,$$

for any sequence $\{\mu^{(n)}\}_{n \in \mathbb{N}} \subseteq \text{AC}([t_0, T]; \mathcal{P}_2(\mathbb{R}^d))$ uniformly W_2 -converging to μ . Combining this with the W_2 -compactness property of Proposition 3.5, we get the desired result.

We prove the existence of an optimal trajectory for V^{inv} . We fix $\{\mu^{(n)}\}_{n \in \mathbb{N}} \subset \mathcal{A}_{[t_0, T]}(\mu)$ and $\hat{\sigma} \in \mathcal{K}$. For any $t \in [t_0, T]$, by triangular inequality and recalling that by definition we have the equivalence $m_2^{1/2}(\rho) = W_2(\rho, \delta_0)$, we get the following uniform bound

$$\begin{aligned} d_{\mathcal{K}}(\mu_t^{(n)}) &\leq W_2(\mu_t^{(n)}, \hat{\sigma}) \leq W_2(\mu_t^{(n)}, \delta_0) + W_2(\delta_0, \hat{\sigma}) \\ &= m_2^{1/2}(\mu_t^{(n)}) + m_2^{1/2}(\hat{\sigma}) \leq \tilde{C}(1 + m_2^{1/2}(\mu)) + m_2^{1/2}(\hat{\sigma}), \end{aligned}$$

for some constant $\tilde{C} > 0$ coming from estimate (A.2) proved in ‘‘Appendix A’’. Thus, as for the proof of the existence of a minimizer for V^{viab} , we can apply Fatou's Lemma to get the u.s.c. of the cost functional and conclude. \square

We state here a first characterization of viability/invariance in terms of the optimal control problems introduced in Definition 4.1.

Proposition 4.3. *Assume (F₁) – (F₂). Let $\mathcal{K} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ be closed in the W_2 -topology, $t_0 \in [0, T]$. Then,*

- (1) \mathcal{K} is viable for F if and only if $V^{\text{viab}}(t_0, \mu_0) = 0$ for all $\mu_0 \in \mathcal{K}$;
- (2) \mathcal{K} is invariant for F if and only if $V^{\text{inv}}(t_0, \mu_0) = 0$ for all $\mu_0 \in \mathcal{K}$.

Proof. We just prove (1), since the proof of (2) is similar. One implication follows directly by definition, so we prove the other direction assuming $V^{\text{viab}}(t_0, \mu_0) = 0$ for all $\mu_0 \in \mathcal{K}$. By Proposition 4.2, for all $\mu_0 \in \mathcal{K}$, there exists an optimal trajectory $\bar{\mu} \in \mathcal{A}_{[t_0, T]}(\mu_0)$ such that

$$0 = V^{\text{viab}}(t_0, \mu_0) = \int_{t_0}^T d_{\mathcal{K}}(\bar{\mu}_t) dt.$$

This implies that $d_{\mathcal{K}}(\bar{\mu}_t) = 0$ for a.e. $t \in [t_0, T]$. By continuity of $\bar{\mu}$ and by closedness of \mathcal{K} w.r.t. W_2 -topology, we obtain the viability property for \mathcal{K} . \square

As usual, the value function satisfies a Dynamic Programming Principle.

Lemma 4.4. (DPP) *The function $V^{\text{viab}} : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty)$ satisfies*

$$V^{\text{viab}}(t_0, \mu) = \inf \left\{ \int_{t_0}^t d_{\mathcal{X}}(\mu_s) ds + V^{\text{viab}}(t, \mu_t) : t \in [t_0, T], \mu \in \mathcal{A}_{[t_0, T]}(\mu) \right\}. \tag{4.3}$$

Furthermore, for any $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$, the map

$$t \mapsto g_{\mu}(t) := \int_{t_0}^t d_{\mathcal{X}}(\mu_s) ds + V^{\text{viab}}(t, \mu_t)$$

is nondecreasing in $[t_0, T]$, and it is constant if and only if μ is an optimal trajectory.

Proof. We prove one inequality (\geq). By definition of $V^{\text{viab}}(t_0, \mu)$, for any $\varepsilon > 0$ there exists $\mu^\varepsilon \in \mathcal{A}_{[t_0, T]}(\mu)$ s.t.

$$V^{\text{viab}}(t_0, \mu) + \varepsilon \geq \int_{t_0}^t d_{\mathcal{X}}(\mu_s^\varepsilon) ds + \int_t^T d_{\mathcal{X}}(\mu_s^\varepsilon) ds \geq \int_{t_0}^t d_{\mathcal{X}}(\mu_s^\varepsilon) ds + V^{\text{viab}}(t, \mu_t^\varepsilon),$$

for any $t \in [t_0, T]$, since the truncated trajectory $\hat{\mu} := \mu^\varepsilon|_{[t, T]}$ belongs to $\mathcal{A}_{[t, T]}(\mu_t^\varepsilon)$. We conclude by passing to the infimum on $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$ and $t \in [t_0, T]$ on the right-hand side and then letting $\varepsilon \rightarrow 0^+$.

Concerning the other inequality, fix any $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$ and $t \in [t_0, T]$. By definition of $V^{\text{viab}}(t, \mu_t)$, for all $\varepsilon > 0$ there exists $\mu^\varepsilon \in \mathcal{A}_{[t, T]}(\mu_t)$ s.t. $V^{\text{viab}}(t, \mu_t) + \varepsilon \geq \int_t^T d_{\mathcal{X}}(\mu_s^\varepsilon) ds$. Now, defining

$$\hat{\mu}_s := \begin{cases} \mu_s, & \text{if } s \in [t_0, t], \\ \mu_s^\varepsilon, & \text{if } s \in [t, T], \end{cases}$$

we see that $\hat{\mu} \in \mathcal{A}_{[t_0, T]}(\mu)$. Thus,

$$V^{\text{viab}}(t_0, \mu) \leq \int_{t_0}^t d_{\mathcal{X}}(\hat{\mu}_s) ds = \int_{t_0}^t d_{\mathcal{X}}(\mu_s) ds + \int_t^T d_{\mathcal{X}}(\mu_s^\varepsilon) ds \leq \int_{t_0}^t d_{\mathcal{X}}(\mu_s) ds + V^{\text{viab}}(t, \mu_t) + \varepsilon.$$

By passing to the inf on $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$, and then letting $\varepsilon \rightarrow 0^+$, we conclude.

The proof of the second part of the statement is standard and follows straightforwardly from (4.3) (see for instance Prop. 3 in [16]). □

We come now to the formulation of a Dynamic Programming Principle for the value function V^{inv} whose proof is omitted since it is similar to that of Lemma 4.4.

Lemma 4.5. (DPP) *The function $V^{\text{inv}} : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow [0, +\infty)$ satisfies*

$$V^{\text{inv}}(t_0, \mu) = \sup \left\{ \int_{t_0}^t d_{\mathcal{X}}(\mu_s) ds + V^{\text{inv}}(t, \mu_t) : t \in [t_0, T], \mu \in \mathcal{A}_{[t_0, T]}(\mu) \right\}. \tag{4.4}$$

Furthermore, for any $\mu \in \mathcal{A}_{[t_0, T]}(\mu)$, the map

$$t \mapsto j_\mu(t) := \int_{t_0}^t d_{\mathcal{H}}(\mu_s) ds + V^{\text{inv}}(t, \mu_t)$$

is nonincreasing in $[t_0, T]$, and it is constant if and only if μ is an optimal trajectory.

As in the classical case, the infinitesimal version of the Dynamic Programming Principle gives rise to a Hamilton–Jacobi–Bellman equation. The next section is devoted to such a Hamilton–Jacobi equation.

Proposition 4.6. *Assume (F_1) – (F_2) . The value functions $V^{\text{viab}}(t, \mu)$ and $V^{\text{inv}}(t, \mu)$ are uniformly continuous in $t \in [0, T]$ and Lipschitz continuous in $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ w.r.t. the W_2 -metric.*

Proof. We prove the statement for V^{viab} , since the proof for V^{inv} is analogous. Fix $t_0 \in [0, T]$ and take any $\mu^1, \mu^2 \in \mathcal{P}_2(\mathbb{R}^d)$. By Proposition 4.2, there exists an optimal trajectory $\bar{\mu}^2 \in \mathcal{A}_{[t_0, T]}(\mu^2)$ starting from μ^2 . Thus, for any admissible $\mu^1 \in \mathcal{A}_{[t_0, T]}(\mu^1)$, we have

$$\left| V^{\text{viab}}(t_0, \mu^1) - V^{\text{viab}}(t_0, \mu^2) \right| \leq \int_{t_0}^T \left| d_{\mathcal{H}}(\mu_t^1) - d_{\mathcal{H}}(\bar{\mu}_t^2) \right| dt \leq \int_{t_0}^T W_2(\mu_t^1, \bar{\mu}_t^2) dt.$$

We can now choose $\mu^1 \in \mathcal{A}_{[t_0, T]}(\mu^1)$ such that the Grönwall-like inequality of Lemma 3.3 holds, thus getting

$$\left| V^{\text{viab}}(t_0, \mu^1) - V^{\text{viab}}(t_0, \mu^2) \right| \leq (T - t_0)e^{L(T-t_0) + (T-t_0)e^{L(T-t_0)}} \cdot W_2(\mu^1, \mu^2). \quad (4.5)$$

We now prove the uniform continuity in time of V^{viab} . Let $0 \leq t_1 \leq t_2 \leq T$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\mu \in \mathcal{A}_{[t_1, T]}(\mu)$ an optimal trajectory. Then by the second part of the statement of Lemma 4.4, noticing that in particular $g_\mu(t_1) = V^{\text{viab}}(t_1, \mu)$, we have

$$\begin{aligned} V^{\text{viab}}(t_1, \mu) - V^{\text{viab}}(t_2, \mu) &= \int_{t_1}^{t_2} d_{\mathcal{H}}(\mu_t) dt + V^{\text{viab}}(t_2, \mu|_{t=t_2}) - V(t_2, \mu) \\ &\leq \int_{t_1}^{t_2} d_{\mathcal{H}}(\mu_t) dt + Te^{LT+T}e^{LT} W_2(\mu|_{t=t_2}, \mu). \end{aligned}$$

By continuity of $d_{\mathcal{H}}(\cdot)$ and of $t \mapsto \mu_t$ we have the convergence to zero of the right-hand-side as $t_2 \rightarrow t_1$. Reversing the roles of t_1 and t_2 we conclude. \square

5. Hamilton Jacobi Bellman equation

As reported in p. 352 in [11] and at the beginning of Sec. 6.1 in [10], we recall the following crucial fact. Throughout the paper, let $(\Omega, \mathcal{B}, \mathbb{P})$ be a sufficiently “rich” probability space, i.e., Ω is a complete, separable metric space, \mathcal{B} is the Borel σ -algebra

on Ω , and \mathbb{P} is an atomless Borel probability measure. We use the notation $L^2_{\mathbb{P}}(\Omega) = L^2_{\mathbb{P}}(\Omega; \mathbb{R}^d)$. Then, given any $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$, there exist $X_1, X_2 \in L^2_{\mathbb{P}}(\Omega)$ such that $\mu_i = X_i \# \mathbb{P}$, $i = 1, 2$, and $W_2(\mu_1, \mu_2) = \|X_1 - X_2\|_{L^2_{\mathbb{P}}}$.

Definition 5.1. (1) Given a function $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, we define its *lift* $U : [0, T] \times L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ by setting $U(t, X) = u(t, X \# \mathbb{P})$ for all $X \in L^2_{\mathbb{P}}(\Omega)$.
 (2) Let $\mathcal{H} = \mathcal{H}(\mu, p)$ be a Hamiltonian function mapping $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $p \in L^2_{\mu}(\mathbb{R}^d)$ into \mathbb{R} . We say that the Hamiltonian function $H : L^2_{\mathbb{P}}(\Omega) \times L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ is a *lift* of \mathcal{H} , if $H(X, p \circ X) = \mathcal{H}(X \# \mathbb{P}, p)$, for all $X \in L^2_{\mathbb{P}}(\Omega)$, $p \in L^2_{X \# \mathbb{P}}(\mathbb{R}^d)$.

Definition 5.2. (*Viscosity solution*) Let \mathcal{H} and H be as in Definition 5.1(2). Given $\lambda \geq 0$, we consider a first-order HJB equation of the form

$$-\partial_t u(t, \mu) + \lambda u(t, \mu) + \mathcal{H}(\mu, D_{\mu}u(t, \mu)) = 0, \quad (5.1)$$

and its lifted form

$$-\partial_t U(t, X) + \lambda U(t, X) + H(X, DU(t, X)) = 0. \quad (5.2)$$

We say that $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is a viscosity subsolution (resp. supersolution) of (5.1) in $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ if and only if its lift is a viscosity subsolution (resp. supersolution) of (5.2) in $[0, T] \times L^2_{\mathbb{P}}(\Omega)$. We recall that $U : [0, T] \times L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ is a

- viscosity subsolution of (5.2) if for any test function $\phi \in C^1([0, T] \times L^2_{\mathbb{P}}(\Omega))$ such that $U - \phi$ has a local maximum at $(t_0, X_0) \in [0, T] \times L^2_{\mathbb{P}}(\Omega)$ it holds $-\partial_t \phi(t_0, X_0) + \lambda U(t_0, X_0) + H(X_0, D\phi(t_0, X_0)) \leq 0$;
- viscosity supersolution of (5.2) if for any test function $\phi \in C^1([0, T] \times L^2_{\mathbb{P}}(\Omega))$ such that $U - \phi$ has a local minimum at $(t_0, X_0) \in [0, T] \times L^2_{\mathbb{P}}(\Omega)$ it holds $-\partial_t \phi(t_0, X_0) + \lambda U(t_0, X_0) + H(X_0, D\phi(t_0, X_0)) \geq 0$;
- viscosity solution of (5.2) if it is both a supersolution and a subsolution.

Remark 5.3. Assume $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is constant in time, i.e., with slight abuse of notation we can identify $u(t, \mu) = u(\mu)$ for any $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$, with $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$. Then, (5.1) and (5.2) become, respectively

$$\lambda u(\mu) + \mathcal{H}(\mu, D_{\mu}u(\mu)) = 0, \quad \lambda U(X) + H(X, DU(X)) = 0, \quad (5.3)$$

where $U : L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ is the lift of u . Moreover, the test functions in Definition 5.2 can be taken independent of t , i.e.,

- U is a viscosity subsolution of (5.3) if for any test function $\phi \in C^1(L^2_{\mathbb{P}}(\Omega))$ such that $U - \phi$ has a local maximum at $X_0 \in L^2_{\mathbb{P}}(\Omega)$ it holds $\lambda U(X_0) + H(X_0, D\phi(X_0)) \leq 0$;
- U is a viscosity supersolution of (5.3) if for any test function $\phi \in C^1(L^2_{\mathbb{P}}(\Omega))$ such that $U - \phi$ has a local minimum at $X_0 \in L^2_{\mathbb{P}}(\Omega)$ it holds $\lambda U(X_0) + H(X_0, D\phi(X_0)) \geq 0$.

- U is a viscosity solution of (5.3) if it is both a supersolution and a subsolution.

Theorem 5.4. (Comparison principle) *Assume that there exists $L, C > 0$ such that the Hamiltonian function $H : L_{\mathbb{P}}^2(\Omega) \times L_{\mathbb{P}}^2(\Omega) \rightarrow \mathbb{R}$ satisfies the following assumption:*

(H) *for any $X, Y \in L_{\mathbb{P}}^2(\Omega)$, any $a, b_1, b_2 > 0$ and $C_1, C_2 \in L_{\mathbb{P}}^2(\Omega)$,*

$$\begin{aligned} & H(Y, a(X - Y) - b_1 Y - C_1) - H(X, a(X - Y) + b_2 X + C_2) \\ & \leq \|X - Y\|_{L_{\mathbb{P}}^2} + 2aL \|X - Y\|_{L_{\mathbb{P}}^2}^2 + \\ & \quad + C(1 + m_2^{1/2}(Y \sharp \mathbb{P})) (1 + \|Y\|_{L_{\mathbb{P}}^2}) (\|C_1\|_{L_{\mathbb{P}}^2} + b_1 \|Y\|_{L_{\mathbb{P}}^2}) + \\ & \quad + C(1 + m_2^{1/2}(X \sharp \mathbb{P})) (1 + \|X\|_{L_{\mathbb{P}}^2}) (\|C_2\|_{L_{\mathbb{P}}^2} + b_2 \|X\|_{L_{\mathbb{P}}^2}). \end{aligned}$$

Let $\lambda \geq 0$. Then, if $u_1, u_2 \in UC([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ are a subsolution and a supersolution of (5.1), respectively, we have

$$\sup_{[0, T] \times \mathcal{P}_2(\mathbb{R}^d)} (u_1 - u_2) \leq \sup_{\{T\} \times \mathcal{P}_2(\mathbb{R}^d)} (u_1 - u_2)^+. \quad (5.4)$$

Proof. The proof follows the line of the corresponding classical finite-dimensional argument (see, e.g., Theorem II.2.12 p. 107) in [6]. In the following, we define $\mathbb{G} := \mathbb{R} \times L_{\mathbb{P}}^2(\Omega)$ and, for any $(t, X) \in \mathbb{G}$, we set $\|(t, X)\|_{\mathbb{G}}^2 := |t|^2 + \|X\|_{L_{\mathbb{P}}^2}^2$. We denote $\mathbb{A} := [0, T] \times L_{\mathbb{P}}^2(\Omega) \subset \mathbb{G}$, that is a complete metric space with distance induced by the norm $\|\cdot\|_{\mathbb{G}}$ of \mathbb{G} .

Let $U_1, U_2 : \mathbb{A} \rightarrow \mathbb{R}$ be, respectively, the lift functionals for u_1 and u_2 as in Definition 5.1(1). We define the functional $\Phi : \mathbb{A}^2 \rightarrow \mathbb{R}$ by setting

$$\begin{aligned} \Phi(t, X, s, Y) := & U_1(t, X) - U_2(s, Y) - \frac{\|(t, X) - (s, Y)\|_{\mathbb{G}}^2}{2\varepsilon} + \\ & - \beta \left((1 + \|X\|_{L_{\mathbb{P}}^2}^2)^{m/2} + (1 + \|Y\|_{L_{\mathbb{P}}^2}^2)^{m/2} \right) + \eta(t + s), \end{aligned}$$

where $\varepsilon, \beta, m, \eta > 0$ are positive constants which will be chosen later. Notice that since $u_i \in UC([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$, $i = 1, 2$, we have $U_i \in UC([0, T] \times L_{\mathbb{P}}^2(\Omega))$. Indeed, for all $X, Y \in L_{\mathbb{P}}^2(\Omega)$, $t, s \in [0, T]$,

$$\begin{aligned} |U_i(t, X) - U_i(s, Y)| & = |u_i(t, X \sharp \mathbb{P}) - u_i(s, Y \sharp \mathbb{P})| \leq \omega_{u_i} \left(\sqrt{|t - s|^2 + W_2^2(X \sharp \mathbb{P}, Y \sharp \mathbb{P})} \right) \\ & \leq \omega_{u_i} (\|(t, X) - (s, Y)\|_{\mathbb{G}}), \end{aligned}$$

where $\omega_{u_i}(\cdot)$ is the modulus of continuity of u_i and where we used the fact that $W_2(X \sharp \mathbb{P}, Y \sharp \mathbb{P}) \leq \|X - Y\|_{L_{\mathbb{P}}^2}$. Set

$$A := \sup_{\{T\} \times \mathcal{P}_2(\mathbb{R}^d)} (u_1 - u_2)^+ = \sup_{\{T\} \times L_{\mathbb{P}}^2} (U_1 - U_2)^+.$$

For $R' > 0$, $i = 1, 2$, set

$$\varrho_i(R') := \sup\{|U_i(t, X) - U_i(s, Y)| : \|(t, X) - (s, Y)\|_{\mathbb{G}} \leq R'\};$$

by uniform continuity we have

$$\sup_{R' \geq 0} \frac{\varrho_i(R')}{1 + R'} < +\infty. \tag{5.5}$$

Thus,

$$\begin{aligned} U_1(t, X) - U_2(s, Y) &= U_1(t, X) - U_1(T, X) + U_1(T, X) - U_2(T, X) \\ &\quad + U_2(T, X) - U_2(s, Y) \\ &\leq \varrho_1(T - t) + A + \varrho_2(\|(T, X) - (s, Y)\|_{\mathbb{G}}) \end{aligned}$$

for all $(t, X, s, Y) \in \mathbb{A}^2$. By (5.5), there exists $\mathcal{C} > 0$ such that

$$|U_1(t, X) - U_2(s, Y)| \leq \mathcal{C}(1 + \|X - Y\|_{L^2_{\mathbb{P}}}), \quad \text{for all } (t, X, s, Y) \in \mathbb{A}^2. \tag{5.6}$$

The proof proceeds by contradiction: assume that there exist $(\tilde{t}, \tilde{\mu}) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ and $\delta > 0$ such that $u_1(\tilde{t}, \tilde{\mu}) - u_2(\tilde{t}, \tilde{\mu}) = A + \delta$. In particular, for any $\tilde{X} \in L^2_{\mathbb{P}}(\Omega)$ such that $\tilde{X} \# \mathbb{P} = \tilde{\mu}$, it holds $U_1(\tilde{t}, \tilde{X}) - U_2(\tilde{t}, \tilde{X}) = A + \delta$.

Select $\beta, \eta > 0$ such that

$$A + \frac{\delta}{2} \leq A + \delta - 2\beta(1 + \|\tilde{X}\|_{L^2_{\mathbb{P}}}^2)^{m/2} + 2\eta\tilde{t} = \Phi(\tilde{t}, \tilde{X}, \tilde{t}, \tilde{X}) \leq \sup_{\mathbb{A}^2} \Phi.$$

Noting that $\Phi \in C^0(\mathbb{A}^2)$, by taking $\varepsilon < \frac{1}{2\mathcal{C}}$ and recalling (5.6), we have

$$\lim_{\substack{\|X\|_{L^2_{\mathbb{P}}} \rightarrow +\infty \\ \|Y\|_{L^2_{\mathbb{P}}} \rightarrow +\infty}} \Phi(t, X, s, Y) = -\infty,$$

for any $t, s \in [0, T]$. Therefore, there exists $R > 0$ such that

$$\sup_{\mathbb{A}^2} \Phi(t, X, s, Y) = \sup_{([0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)})^2} \Phi(t, X, s, Y).$$

Thus, by Stegall’s Variational Principle (see, e.g., Theorem 6.3.5 in [8]) for any fixed $\xi > 0$, there exists a linear and continuous functional $\Lambda : \mathbb{G}^2 \rightarrow \mathbb{R}$ with $\|\Lambda\|_{\mathbb{G}^{2*}} < \xi$ and such that $\Phi - \Lambda$ attains a strong maximum in $([0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)})^2$. Moreover, on $([0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)})^2$, we have

$$\Phi(t, X, s, Y) - \Lambda(t, X, s, Y) \geq \Phi(t, X, s, Y) - 2\xi\sqrt{T^2 + R^2},$$

and so

$$\sup_{\mathbb{A}^2} \Phi \leq 2\xi\sqrt{T^2 + R^2} + \sup_{([0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)})^2} (\Phi - \Lambda). \tag{5.7}$$

Let $(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) \in ([0, T] \times \overline{B_{L_{\mathbb{P}}^2}(0, R)})^2$ be a maximizer of $\Phi - \Lambda$ on $([0, T] \times \overline{B_{L_{\mathbb{P}}^2}(0, R)})^2$, obtained by choosing $\xi > 0$ s.t. $2\xi\sqrt{T^2 + R^2} \leq \frac{\delta}{8}$. In particular, we get

$$\begin{aligned} A + \frac{\delta}{2} &\leq \sup_{\mathbb{A}^2} \Phi \leq 2\xi\sqrt{T^2 + R^2} + (\Phi - \Lambda)(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) \leq \frac{\delta}{8} + (\Phi - \Lambda)(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) \\ &\leq \frac{\delta}{8} + \Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) + 2\xi\sqrt{T^2 + R^2} \leq \frac{\delta}{4} + \Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}), \end{aligned}$$

and so

$$\Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) \geq A + \frac{\delta}{4}, \quad (5.8)$$

leading to

$$\beta \left((1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{m/2} + (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{m/2} \right) \leq \sup U_1 - \inf U_2 - A - \frac{\delta}{4} + \eta(\bar{t} + \bar{s}).$$

By choosing $0 < \eta < 1$, we get for all $\varepsilon > 0, m \in (0, 1]$

$$\beta((1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{m/2} + (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{m/2}) \leq \sup U_1 - \inf U_2 - A - \frac{\delta}{4} + 2T =: d > 0. \quad (5.9)$$

By Riesz' representation theorem, there exist unique $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{G}^2$ such that

$$\Lambda(t, X, s, Y) = \lambda_1 t + \langle \lambda_2, X \rangle_{L_{\mathbb{P}}^2} + \lambda_3 s + \langle \lambda_4, Y \rangle_{L_{\mathbb{P}}^2}.$$

From (5.7), we have

$$\begin{aligned} \Phi(\bar{t}, \bar{X}, \bar{t}, \bar{X}) + \Phi(\bar{s}, \bar{Y}, \bar{s}, \bar{Y}) &\leq 2(\Phi - \Lambda)(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) + 4\xi\sqrt{T^2 + R^2} \\ &\leq 2\Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) + 8\xi\sqrt{T^2 + R^2}, \end{aligned}$$

and so

$$\begin{aligned} &U_1(\bar{t}, \bar{X}) - U_2(\bar{t}, \bar{X}) + U_1(\bar{s}, \bar{Y}) - U_2(\bar{s}, \bar{Y}) + \\ &\quad - 2\beta \left((1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{m/2} + (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{m/2} \right) + 2\eta(\bar{t} + \bar{s}) \\ &\leq 2U_1(\bar{t}, \bar{X}) - 2U_2(\bar{s}, \bar{Y}) - \frac{\|(\bar{t}, \bar{X}) - (\bar{s}, \bar{Y})\|_{\mathbb{G}}^2}{\varepsilon} - 2\beta \left((1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{m/2} + (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{m/2} \right) + \\ &\quad + 2\eta(\bar{t} + \bar{s}) + 8\xi\sqrt{T^2 + R^2}, \end{aligned}$$

which leads to

$$\begin{aligned} \frac{\|(\bar{t}, \bar{X}) - (\bar{s}, \bar{Y})\|_{\mathbb{G}}^2}{\varepsilon} &\leq U_1(\bar{t}, \bar{X}) - U_1(\bar{s}, \bar{Y}) + U_2(\bar{t}, \bar{X}) - U_2(\bar{s}, \bar{Y}) \\ &\quad + 8\xi\sqrt{T^2 + R^2}. \end{aligned} \quad (5.10)$$

Take $0 < \xi < \varepsilon < 1$. From the previous inequality, the boundedness of U_1, U_2 in $[0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)}$ gives

$$\|(\bar{t}, \bar{X}) - (\bar{s}, \bar{Y})\|_{\mathbb{G}} \leq B' \sqrt{\varepsilon} + 8\varepsilon \sqrt{T^2 + R^2} \leq B \sqrt{\varepsilon}, \quad (5.11)$$

for suitable constants $B', B > 0$ independent on ε .

By uniform continuity of U_i , $i = 1, 2$, and by plugging the previous relation in (5.10), we can build a modulus of continuity $\omega(\cdot)$ such that

$$\frac{\|(\bar{t}, \bar{X}) - (\bar{s}, \bar{Y})\|_{\mathbb{G}}^2}{\varepsilon} \leq \omega(\varepsilon) := \omega_{u_1}(B\sqrt{\varepsilon}) + \omega_{u_2}(B\sqrt{\varepsilon}) + 8\varepsilon \sqrt{T^2 + R^2}. \quad (5.12)$$

We show that neither \bar{t} nor \bar{s} can be equal to T . Indeed, in $\bar{t} = T$,

$$\begin{aligned} \Phi(T, \bar{X}, \bar{s}, \bar{Y}) &\leq U_1(T, \bar{X}) - U_2(T, \bar{X}) + U_2(T, \bar{X}) - U_2(\bar{s}, \bar{Y}) + \eta(T + \bar{s}) \\ &\leq A + \omega_{u_2}(B\sqrt{\varepsilon}) + 2\eta T, \end{aligned}$$

by definition of A . We thus get a contradiction with (5.8) by choosing ε and η small enough s.t. $\omega_{u_2}(B\sqrt{\varepsilon}) + 2\eta T < \frac{\delta}{4}$. The same reasoning applies for proving $\bar{s} < T$.

We define the $C^1(\mathbb{A})$ test functions

$$\begin{aligned} \phi(t, X) &:= U_2(\bar{s}, \bar{Y}) + \frac{\|(t, X) - (\bar{s}, \bar{Y})\|_{\mathbb{G}}^2}{2\varepsilon} + \\ &\quad + \beta \left((1 + \|X\|_{L^2_{\mathbb{P}}}^2)^{m/2} + (1 + \|\bar{Y}\|_{L^2_{\mathbb{P}}}^2)^{m/2} \right) - \eta(t + \bar{s}) + \\ &\quad + \Lambda(t, X, \bar{s}, \bar{Y}), \\ \psi(s, Y) &:= U_1(\bar{t}, \bar{X}) - \frac{\|(\bar{t}, \bar{X}) - (s, Y)\|_{\mathbb{G}}^2}{2\varepsilon} + \\ &\quad - \beta \left((1 + \|\bar{X}\|_{L^2_{\mathbb{P}}}^2)^{m/2} + (1 + \|Y\|_{L^2_{\mathbb{P}}}^2)^{m/2} \right) + \eta(\bar{t} + s) + \\ &\quad - \Lambda(\bar{t}, \bar{X}, s, Y). \end{aligned}$$

Notice that $(U_1 - \phi)(t, X) = (\Phi - \Lambda)(t, X, \bar{s}, \bar{Y})$, hence, $U_1 - \phi$ attains its maximum at $(\bar{t}, \bar{X}) \in [0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)}$ and, similarly, $U_2 - \psi$ attains its minimum at $(\bar{s}, \bar{Y}) \in [0, T] \times \overline{B_{L^2_{\mathbb{P}}}(0, R)}$. We have

$$\begin{aligned} \partial_t \phi(\bar{t}, \bar{X}) &= \frac{\bar{t} - \bar{s}}{\varepsilon} - \eta + \lambda_1, & \partial_t \psi(\bar{s}, \bar{Y}) &= \frac{\bar{t} - \bar{s}}{\varepsilon} + \eta - \lambda_3, \\ D\phi(\bar{t}, \bar{X}) &= \frac{\bar{X} - \bar{Y}}{\varepsilon} + m\beta(1 + \|\bar{X}\|_{L^2_{\mathbb{P}}}^2)^{\frac{m-2}{2}} \bar{X} + \lambda_2, \\ D\psi(\bar{s}, \bar{Y}) &= \frac{\bar{X} - \bar{Y}}{\varepsilon} - m\beta(1 + \|\bar{Y}\|_{L^2_{\mathbb{P}}}^2)^{\frac{m-2}{2}} \bar{Y} - \lambda_4. \end{aligned}$$

Since $\bar{t}, \bar{s} \in [0, T)$, by definition of viscosity sub/supersolution, we have

$$\begin{aligned} -\partial_t \phi(\bar{t}, \bar{X}) + \lambda U_1(\bar{t}, \bar{X}) + H(\bar{X}, D\phi(\bar{t}, \bar{X})) &\leq 0 \leq \\ \leq -\partial_t \psi(\bar{s}, \bar{Y}) + \lambda U_2(\bar{s}, \bar{Y}) + H(\bar{Y}, D\psi(\bar{s}, \bar{Y})). \end{aligned}$$

Now, by (5.8), we have

$$\begin{aligned} U_1(\bar{t}, \bar{X}) - U_2(\bar{s}, \bar{Y}) &\geq \Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) - \eta(\bar{t} + \bar{s}) \geq \Phi(\bar{t}, \bar{X}, \bar{s}, \bar{Y}) - 2T\eta \\ &\geq A + \frac{\delta}{4} - 2T\eta, \end{aligned}$$

and we can choose η sufficiently small so that $A + \frac{\delta}{4} - 2T\eta \geq 0$. Then, we get

$$\begin{aligned} 2\eta &\leq 2\eta + \lambda(U_1(\bar{t}, \bar{X}) - U_2(\bar{s}, \bar{Y})) \\ &\leq \lambda_1 + \lambda_3 + H(\bar{Y}, D\psi(\bar{s}, \bar{Y})) - H(\bar{X}, D\phi(\bar{t}, \bar{X})). \end{aligned}$$

We can now invoke assumption **(H)** with

$$\begin{aligned} a &= \frac{1}{\varepsilon}, \quad b_1 = m\beta(1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}}, \quad C_1 = \lambda_4, \\ b_2 &= m\beta(1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}}, \quad C_2 = \lambda_2, \end{aligned}$$

recalling that $\lambda_1, \lambda_3 \leq \varepsilon$ and $\|\lambda_2\|_{L_{\mathbb{P}}^2}, \|\lambda_4\|_{L_{\mathbb{P}}^2} \leq \varepsilon$ by the bound on the dual norm of the operator Λ . We get

$$\begin{aligned} 2\eta &\leq 2\varepsilon + \|\bar{X} - \bar{Y}\|_{L_{\mathbb{P}}^2} + 2L \frac{\|\bar{X} - \bar{Y}\|_{L_{\mathbb{P}}^2}^2}{\varepsilon} + \|\lambda_4\|_{L_{\mathbb{P}}^2} C(1 + m_2^{1/2}(\bar{Y} \# \mathbb{P})) (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}) + \\ &\quad + C(1 + m_2^{1/2}(\bar{Y} \# \mathbb{P})) (1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}) m\beta(1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}} \|\bar{Y}\|_{L_{\mathbb{P}}^2} + \\ &\quad + \|\lambda_2\|_{L_{\mathbb{P}}^2} C(1 + m_2^{1/2}(\bar{X} \# \mathbb{P})) (1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}) + \\ &\quad + C(1 + m_2^{1/2}(\bar{X} \# \mathbb{P})) (1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}) m\beta(1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}} \|\bar{X}\|_{L_{\mathbb{P}}^2}. \end{aligned}$$

By (5.11), (5.12) and recalling that $\bar{X}, \bar{Y} \in \overline{B_{L_{\mathbb{P}}^2}(0, R)}$, we have

$$\begin{aligned} 2\eta &\leq 2\varepsilon + B\sqrt{\varepsilon} + 2L\omega(\varepsilon) + \\ &\quad + 2\varepsilon D(1 + R) + D_R m\beta \left((1 + \|\bar{Y}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}} + (1 + \|\bar{X}\|_{L_{\mathbb{P}}^2}^2)^{\frac{m-2}{2}} \right), \end{aligned}$$

where we defined $D_R := D(1 + R)R$, where $D := \max\{C(1 + m_2^{1/2}(\bar{Y} \# \mathbb{P})), C(1 + m_2^{1/2}(\bar{X} \# \mathbb{P}))\} > 0$. Finally, by (5.9) we get

$$2\eta \leq 2\varepsilon + B\sqrt{\varepsilon} + 2L\omega(\varepsilon) + 2\varepsilon D(1 + R) + D_R m d \leq K o(1) + \eta,$$

where for the last passage we choose $m \leq \frac{\eta}{D_R d}$, and $o(1)$ is a function of ε going to 0 as $\varepsilon \rightarrow 0^+$. This leads to a contradiction as $\varepsilon \rightarrow 0^+$. \square

Remark 5.5. As highlighted also in Remark 3.8 p. 154 of [6], if $\lambda = 0$ in (5.1), we can drop the symbol of the positive part in (5.4) and conclude that

$$\sup_{[0, T] \times \mathcal{P}_2(\mathbb{R}^d)} (u_1 - u_2) \leq \sup_{\{T\} \times \mathcal{P}_2(\mathbb{R}^d)} (u_1 - u_2).$$

6. Viscosity characterization of viability and invariance

We now provide the main results of the paper: Theorems 6.6 and 6.7. As pointed out also in Remark 4.2 in [18], by Theorem 8.2.11 in [4], the Hamiltonian $\mathcal{H}_F^{\text{viab}}$ defined in Theorem 1.1 satisfies

$$\mathcal{H}_F^{\text{viab}}(\mu, p) = -d_{\mathcal{X}}(\mu) - \int_{\mathbb{R}^d} \inf_{v \in F(\mu, x)} \langle v, p(x) \rangle d\mu(x). \quad (6.1)$$

Definition 6.1. (*Lifted Hamiltonian for viability*) We define the lifted Hamiltonian in $L_{\mathbb{P}}^2(\Omega)$ associated with $\mathcal{H}_F^{\text{viab}}$

$$H_F^{\text{viab}}(X, Q) := -d_{\mathcal{X}}(X \# \mathbb{P}) - \inf_{\substack{v(\cdot) \in L_{X \# \mathbb{P}}^2(\mathbb{R}^d) \\ v(x) \in F(X \# \mathbb{P}, x) \\ \text{for } X \# \mathbb{P}\text{-a.e. } x}} \int_{\Omega} \langle v \circ X(\omega), Q(\omega) \rangle d\mathbb{P}(\omega),$$

for all $X, Q \in L_{\mathbb{P}}^2(\Omega)$. Note that H_F^{viab} is a lift of $\mathcal{H}_F^{\text{viab}}$ according to Definition 5.1.

By disintegrating $\mathbb{P} = (X \# \mathbb{P}) \otimes \mathbb{P}_x$ (see Theorem 2.1), we have

$$\begin{aligned} H_F^{\text{viab}}(X, Q) &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \inf_{\substack{v \in L_{X \# \mathbb{P}}^2(\mathbb{R}^d) \\ v(x) \in F(X \# \mathbb{P}, x) \\ \text{for } X \# \mathbb{P}\text{-a.e. } x}} \int_{\mathbb{R}^d} \int_{X^{-1}(x)} \langle v \circ X(\omega), Q(\omega) \rangle d\mathbb{P}_x(\omega) dX \# \mathbb{P}(x) \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \inf_{\substack{v \in L_{X \# \mathbb{P}}^2(\mathbb{R}^d) \\ v(x) \in F(X \# \mathbb{P}, x) \\ \text{for } X \# \mathbb{P}\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle v(x), \int_{X^{-1}(x)} Q(\omega) d\mathbb{P}_x(\omega) \rangle dX \# \mathbb{P}(x) \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \int_{\mathbb{R}^d} \inf_{v \in F(X \# \mathbb{P}, x)} \langle v, \int_{X^{-1}(x)} Q(\omega) d\mathbb{P}_x(\omega) \rangle dX \# \mathbb{P}(x) \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \int_{\mathbb{R}^d} \int_{X^{-1}(x)} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle v, Q(\omega) \rangle d\mathbb{P}_x(\omega) dX \# \mathbb{P}(x) \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \int_{\Omega} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle v, Q(\omega) \rangle d\mathbb{P}(\omega) \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - \inf_{\substack{v \in L_{\mathbb{P}}^2(\Omega) \\ v(\cdot) \in F(X \# \mathbb{P}, X(\cdot))}} \int_{\Omega} \langle v(\omega), Q(\omega) \rangle d\mathbb{P}(\omega), \end{aligned} \quad (6.2)$$

where in the last equality we used Theorem 8.2.11 in [4] (or Theorem 6.31 in [14]).

Definition 6.2. (*Lifted Hamiltonian for invariance*) Related with the invariance problem and associated with $\mathcal{H}_F^{\text{inv}}$, we define the following lifted Hamiltonian in $L_{\mathbb{P}}^2(\Omega)$

$$H_F^{\text{inv}}(X, Q) := -d_{\mathcal{X}}(X \# \mathbb{P}) - \sup_{\substack{v \in L_{X \# \mathbb{P}}^2(\mathbb{R}^d) \\ v(x) \in F(X \# \mathbb{P}, x) \\ \text{for } X \# \mathbb{P}\text{-a.e. } x}} \int_{\Omega} \langle v \circ X(\omega), Q(\omega) \rangle d\mathbb{P}(\omega),$$

for all $X, Q \in L_{\mathbb{P}}^2(\Omega)$. Notice that H_F^{inv} is a lift of $\mathcal{H}_F^{\text{inv}}$ according to Definition 5.1. Moreover, the equivalences (6.1) and (6.2) hold also in this case replacing, respectively, $\mathcal{H}_F^{\text{viab}}$, H_F^{viab} with $\mathcal{H}_F^{\text{inv}}$, H_F^{inv} , and inf with sup.

Lemma 6.3. *Assume (F₁) – (F₂). Then, both the Hamiltonian functions H_F^{viab} and H_F^{inv} satisfy assumption (H) with L and C , respectively, as in (F₂) and (3.1).*

Proof. We prove here the assertion for H_F^{viab} since the assertion for H_F^{inv} can be proved in the same way. Fix any $X, Y \in L^2_{\mathbb{P}}$, $a, b_1, b_2 > 0$ and $C_1, C_2 \in L^2_{\mathbb{P}}$, and denote $\mu_1 := X \sharp \mathbb{P}$, $\mu_2 := Y \sharp \mathbb{P}$. We have

$$\begin{aligned} & H_F^{\text{viab}}(Y, a(X - Y) - b_1 Y - C_1) - H_F^{\text{viab}}(X, a(X - Y) + b_2 X + C_2) = \\ & -d_{\mathcal{X}}(\mu_2) - \int_{\Omega} \inf_{v \in F(\mu_2, Y(\omega))} \{a \langle v, X(\omega) - Y(\omega) \rangle - b_1 \langle v, Y(\omega) \rangle - \langle v, C_1(\omega) \rangle\} d\mathbb{P} \\ & + d_{\mathcal{X}}(\mu_1) + \int_{\Omega} \inf_{w \in F(\mu_1, X(\omega))} \{a \langle w, X(\omega) - Y(\omega) \rangle + b_2 \langle w, X(\omega) \rangle + \langle w, C_2(\omega) \rangle\} d\mathbb{P}. \end{aligned} \quad (6.3)$$

Let $p \in \mathbb{R}^d$. For any $x, y \in \mathbb{R}^d$, define $\delta_{x,y} := L(W_2(\mu_1, \mu_2) + |x - y|)$. Given any $\varepsilon > 0$, there exists $z_{\varepsilon,p} \in F(\mu_1, x) + \delta_{x,y} \overline{B(0, 1)}$ such that

$$\inf_{v \in F(\mu_2, y)} \langle v, p \rangle \geq \inf_{z \in F(\mu_1, x) + \delta_{x,y} \overline{B(0, 1)}} \langle z, p \rangle \geq \langle z_{\varepsilon,p}, p \rangle - \varepsilon,$$

where the first inequality comes from Lipschitz continuity of the set-valued map F . In particular, we can write $z_{\varepsilon,p} = \hat{w}_{\varepsilon,p} + \delta_{x,y} w_{\varepsilon,p}$, with $\hat{w}_{\varepsilon,p} \in F(\mu_1, x)$ and $w_{\varepsilon,p} \in \overline{B(0, 1)}$, thus getting

$$\begin{aligned} \inf_{v \in F(\mu_2, y)} \langle v, p \rangle & \geq \langle \hat{w}_{\varepsilon,p}, p \rangle + \delta_{x,y} \langle w_{\varepsilon,p}, p \rangle - \varepsilon \\ & \geq \inf_{w \in F(\mu_1, x)} \langle w, p \rangle - \delta_{x,y} |p| - \varepsilon. \end{aligned}$$

Hence, we have

$$\inf_{w \in F(\mu_1, x)} \langle w, p \rangle - \inf_{v \in F(\mu_2, y)} \langle v, p \rangle \leq L(W_2(\mu_1, \mu_2) + |x - y|) |p|. \quad (6.4)$$

Thus, for any $x, y, c_1, c_2 \in \mathbb{R}^d$ and by choosing $p = x - y$, it holds

$$\begin{aligned} & \inf_{w \in F(\mu_1, x)} \{a \langle w, x - y \rangle + b_2 \langle w, x \rangle + \langle w, c_2 \rangle\} + \\ & - \inf_{v \in F(\mu_2, y)} \{a \langle v, x - y \rangle - b_1 \langle v, y \rangle - \langle v, c_1 \rangle\} \\ & \leq a \inf_{w \in F(\mu_1, x)} \langle w, x - y \rangle + \sup_{w \in F(\mu_1, x)} \{b_2 \langle w, x \rangle + \langle w, c_2 \rangle\} + \\ & - a \inf_{v \in F(\mu_2, y)} \langle v, x - y \rangle + \sup_{v \in F(\mu_2, y)} \{b_1 \langle v, y \rangle + \langle v, c_1 \rangle\} \\ & \leq a L(W_2(\mu_1, \mu_2) + |x - y|) |x - y| + \\ & + \sup_{w \in F(\mu_1, x)} \{b_2 \langle w, x \rangle + \langle w, c_2 \rangle\} + \sup_{v \in F(\mu_2, y)} \{b_1 \langle v, y \rangle + \langle v, c_1 \rangle\} \\ & \leq a L(W_2(\mu_1, \mu_2) + |x - y|) |x - y| + \\ & + b_2 \sup_{w \in F(\mu_1, x)} \langle w, x \rangle + \sup_{w \in F(\mu_1, x)} \langle w, c_2 \rangle + b_1 \sup_{v \in F(\mu_2, y)} \langle v, y \rangle + \sup_{v \in F(\mu_2, y)} \langle v, c_1 \rangle \\ & \leq a L(W_2(\mu_1, \mu_2) + |x - y|) |x - y| + \\ & + b_2 |x| \sup_{w \in F(\mu_1, x)} |w| + |c_2| \sup_{w \in F(\mu_1, x)} |w| + b_1 |y| \sup_{v \in F(\mu_2, y)} |v| + |c_1| \sup_{v \in F(\mu_2, y)} |v|, \end{aligned}$$

where we used the Cauchy–Schwarz’s inequality. Integrating with respect to the measure $(X, Y, C_1, C_2) \# \mathbb{P}$ on the variables (x, y, c_1, c_2) and by (3.1), we get

$$\begin{aligned} & \int_{\Omega} \inf_{w \in F(\mu_1, X(\omega))} \{a(w, X(\omega) - Y(\omega)) + b_2 \langle w, X(\omega) \rangle + \langle w, C_2(\omega) \rangle\} \, d\mathbb{P} + \\ & - \int_{\Omega} \inf_{v \in F(\mu_2, Y(\omega))} \{a(v, X(\omega) - Y(\omega)) - b_1 \langle v, Y(\omega) \rangle - \langle v, C_1(\omega) \rangle\} \, d\mathbb{P} \\ & \leq 2aL \|X - Y\|_{L^2_{\mathbb{P}}}^2 + b_2 C(1 + m_2^{1/2}(\mu_1))(1 + \|X\|_{L^2_{\mathbb{P}}}) \|X\|_{L^2_{\mathbb{P}}} + \\ & + b_1 C(1 + m_2^{1/2}(\mu_2))(1 + \|Y\|_{L^2_{\mathbb{P}}}) \|Y\|_{L^2_{\mathbb{P}}} + \\ & + \|C_2\|_{L^2_{\mathbb{P}}} C(1 + m_2^{1/2}(\mu_1))(1 + \|X\|_{L^2_{\mathbb{P}}}) + \|C_1\|_{L^2_{\mathbb{P}}} C(1 + m_2^{1/2}(\mu_2))(1 + \|Y\|_{L^2_{\mathbb{P}}}) \end{aligned}$$

recalling that $W_2(X \# \mathbb{P}, Y \# \mathbb{P}) \leq \|X - Y\|_{L^2_{\mathbb{P}}}$.

We conclude from (6.3), thanks to the Lipschitz continuity of $d_{\mathcal{H}}(\cdot)$. □

Remark 6.4. Assume $(F_1) - (F_2)$. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ be fixed. Then, the set of continuous selections of $F(\mu, \cdot)$ is dense in $L^2_{\mu}(\mathbb{R}^d)$ in the set of Borel selections of $F(\mu, \cdot)$. Indeed, let $v(\cdot)$ be a Borel selection of $F(\mu, \cdot)$. By Lusin’s Theorem, for any $\varepsilon > 0$ there exists a compact $K_{\varepsilon} \subseteq \mathbb{R}^d$ and a continuous map $w_{\varepsilon} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $v = w_{\varepsilon}$ on K_{ε} and $\mu(\mathbb{R}^d \setminus K_{\varepsilon}) < \varepsilon$. By Corollary 9.1.3 in [4], we can extend $w_{\varepsilon}|_{K_{\varepsilon}}$ to a continuous selection v_{ε} of $F(\mu, \cdot)$. Moreover, we have

$$\|v_{\varepsilon} - v\|_{L^2_{\mu}} \leq 2\|\chi_{\mathbb{R}^d \setminus K_{\varepsilon}} |F(\mu, \cdot)|\|_{L^2_{\mu}}.$$

Since $|F(\mu, x)| \leq |F(\delta_0, 0)| + Lm_2^{1/2}(\mu) + L|x|$, we have that

$$\|v_{\varepsilon} - v\|_{L^2_{\mu}} \leq 2\left(\left(|F(\delta_0, 0)| + 2Lm_2^{1/2}(\mu)\right) \chi_{\mathbb{R}^d \setminus K_{\varepsilon}}\right)\|_{L^2_{\mu}} \leq \varepsilon\left(|F(\delta_0, 0)| + 2Lm_2^{1/2}(\mu)\right),$$

and the right hand side tends to 0 as $\varepsilon \rightarrow 0^+$.

Now, we deduce that the value functions V^{viab} and V^{inv} satisfy the following Hamilton–Jacobi equations.

Proposition 6.5. *Assume $(F_1) - (F_2)$. Then,*

(1) V^{viab} is a viscosity solution of

$$-\partial_t u(t, \mu) + \mathcal{H}_F^{\text{viab}}(\mu, D_{\mu}u(t, \mu)) = 0; \tag{6.5}$$

(2) V^{inv} is a viscosity solution of

$$-\partial_t u(t, \mu) + \mathcal{H}_F^{\text{inv}}(\mu, D_{\mu}u(t, \mu)) = 0. \tag{6.6}$$

Proof. We prove (1). Let $U : [0, T] \times L^2(\Omega; \mathbb{R}^d) \rightarrow \mathbb{R}$ be the lift of V^{viab} according to Definition 5.1.

Claim 1. U is a viscosity supersolution of $-\partial_t U(t, X) + H_F(X, DU(t, X)) = 0$.

Proof of Claim 1. Let $\phi : [0, T] \times L_{\mathbb{P}}^2(\Omega; \mathbb{R}^d) \rightarrow \mathbb{R}$ be a C^1 map such that $U - \phi$ attains its minimum at (s, X) , and define $\mu = X \sharp \mathbb{P}$. Let $\boldsymbol{\mu} = \{\mu_t\}_{t \in [s, T]}$ be an optimal trajectory defined on $[s, T]$ with $\mu_s = \mu$, its existence being assured by Proposition 4.2, and let $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[s, T]})$ such that $e_t \sharp \boldsymbol{\eta} = \mu_t$ for all $t \in [s, T]$. Fix $\varepsilon > 0$ and choose a family $\{Y_t^\varepsilon\}_{t \in [s, T]} \subseteq L_{\mathbb{P}}^2(\Omega)$ of random variables satisfying the properties of Corollary A.3 related to $\boldsymbol{\mu}$. Then, by the Dynamic Programming Principle in Lemma 4.4 and optimality of $\boldsymbol{\mu}$,

$$\begin{aligned} 0 &= U(t, Y_t^\varepsilon) - U(s, Y_s^\varepsilon) + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau = U(t, Y_t^\varepsilon) - U(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau \\ &\geq \phi(t, Y_t^\varepsilon) - \phi(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau, \end{aligned}$$

where the equality $U(s, Y_s^\varepsilon) = U(s, X)$ holds since $Y_s^\varepsilon \sharp \mathbb{P} = X \sharp \mathbb{P} = \mu$ and since U , as a lift, is law dependent. Therefore, there exists a continuous increasing function $\varrho : [0, +\infty[\rightarrow [0, +\infty[$ with $\varrho(k)/k \rightarrow 0$ as $k \rightarrow 0^+$ such that we have

$$\begin{aligned} 0 &\geq \phi(t, Y_t^\varepsilon) - \phi(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau \\ &\geq \partial_t \phi(s, X)(t - s) + \langle D\phi(s, X), Y_t^\varepsilon - X \rangle_{L_{\mathbb{P}}^2} + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau + \\ &\quad - \varrho \left(|t - s| \left(1 + \left\| \frac{Y_t^\varepsilon - Y_s^\varepsilon}{t - s} \right\|_{L_{\mathbb{P}}^2} \right) + \varepsilon \right) \\ &\geq \partial_t \phi(s, X)(t - s) + \langle D\phi(s, X), Y_t^\varepsilon - Y_s^\varepsilon \rangle_{L_{\mathbb{P}}^2} + \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau - \varepsilon \|D\phi(s, X)\|_{L_{\mathbb{P}}^2} \\ &\quad - \varrho \left(|t - s| \left(1 + \left\| \frac{e_t - e_s}{t - s} \right\|_{L_{\eta}^2} \right) + \varepsilon \right). \end{aligned}$$

Dividing by $t - s > 0$, by Corollary A.3(3), we have

$$\begin{aligned} 0 &\geq \partial_t \phi(s, X) + \int_s^t \inf_{v \in F(X \sharp \mathbb{P}, X(\omega))} \langle D\phi(s, X), v \rangle \, d\mathbb{P}(\omega) + \frac{1}{t - s} \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau + \\ &\quad - \frac{\varepsilon}{t - s} \|D\phi(s, X)\|_{L_{\mathbb{P}}^2} - \frac{1}{t - s} \varrho \left(|t - s| \left(1 + \left\| \frac{e_t - e_s}{t - s} \right\|_{L_{\eta}^2} \right) + \varepsilon \right) + \\ &\quad - (\widehat{\omega}(t) + L\varepsilon) \|D\phi(s, X)\|_{L_{\mathbb{P}}^2}. \end{aligned}$$

By letting $\varepsilon \rightarrow 0^+$, we obtain

$$\begin{aligned} 0 &\geq \partial_t \phi(s, X) + \int_s^t \inf_{v \in F(X \sharp \mathbb{P}, X(\omega))} \langle D\phi(s, X), v \rangle \, d\mathbb{P}(\omega) + \frac{1}{t - s} \int_s^t d_{\mathcal{X}}(\mu_\tau) \, d\tau + \\ &\quad - \frac{1}{t - s} \varrho \left(|t - s| \left(1 + \left\| \frac{e_t - e_s}{t - s} \right\|_{L_{\eta}^2} \right) \right) - \widehat{\omega}(t) \cdot \|D\phi(s, X)\|_{L_{\mathbb{P}}^2}. \end{aligned}$$

Recalling the boundedness of $\left\| \frac{e_t - e_s}{t - s} \right\|_{L^2_{\eta}}$ coming from Proposition 3.4, by letting $t \rightarrow s^+$, we have

$$0 \geq \partial_t \phi(s, X) + \int_{\Omega} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle D\phi(s, X), v \rangle d\mathbb{P}(\omega) + d_{\mathcal{X}}(\mu_s),$$

i.e., $-\partial_t \phi(s, X) + H_F^{\text{viab}}(X, D\phi(s, X)) \geq 0$, where, as already discussed, we have

$$H_F^{\text{viab}}(X, Q) = -d_{\mathcal{X}}(X \# \mathbb{P}) - \int_{\Omega} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle Q(\omega), v \rangle d\mathbb{P}(\omega).$$

Thus, U is a viscosity supersolution of $-\partial_t U(t, X) + H_F^{\text{viab}}(X, DU(t, X)) = 0$.

Claim 2. U is a viscosity subsolution of $-\partial_t U(t, X) + H_F^{\text{viab}}(X, DU(t, X)) = 0$.

Proof of Claim 2. Let $\phi : [0, T] \times L^2_{\mathbb{P}}(\Omega; \mathbb{R}^d) \rightarrow \mathbb{R}$ be a C^1 map such that $U - \phi$ attains its maximum at (s, X) and define $\mu = X \# \mathbb{P}$. Fix $\varepsilon > 0$, and let $v_{\varepsilon} \in L^2_{\mu}(\mathbb{R}^d)$ be such that $v_{\varepsilon}(x) \in F(\mu, x)$ for μ -a.e. $x \in \mathbb{R}^d$ and

$$\int_{\Omega} \langle D\phi(s, X)(\omega), v_{\varepsilon} \circ X(\omega) \rangle d\mathbb{P}(\omega) \leq \inf_{\substack{v \in L^2_{X \# \mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X \# \mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(s, X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) + \frac{\varepsilon}{2}.$$

By Remark 6.4, we can suppose that $v_{\varepsilon} \in C^0$, and by Lemma A.4 there exists an admissible trajectory $\mu^{\varepsilon} = \{\mu_t^{\varepsilon}\}_{t \in [s, T]}$ defined on $[s, T]$ with $\mu_s^{\varepsilon} = \mu$, and $\eta^{\varepsilon} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[s, T]})$ such that $e_t \# \eta^{\varepsilon} = \mu_t^{\varepsilon}$ for all $t \in [s, T]$ and

$$\lim_{t \rightarrow s^+} \left\| \frac{e_t - e_s}{t - s} - v_{\varepsilon} \circ e_s \right\|_{L^2_{\eta^{\varepsilon}}} = 0.$$

By density, we can find $\hat{v}_{\varepsilon} \in C_b^0(\mathbb{R}^d)$ such that $\|v_{\varepsilon} - \hat{v}_{\varepsilon}\|_{L^2_{\mu}} \leq \varepsilon$.

Denote by $\mathcal{V}_{\varepsilon} : \Omega \rightarrow \mathbb{R}^d \times \Gamma_{[s, T]}$ a Borel map satisfying $\eta^{\varepsilon} = \mathcal{V}_{\varepsilon} \# \mathbb{P}$. Recalling Lemma A.2, since for all $\varepsilon > 0$ we have $\mu = \mu_s^{\varepsilon} = e_s \# \eta^{\varepsilon} = (e_s \circ \mathcal{V}_{\varepsilon}) \# \mathbb{P} = X \# \mathbb{P}$, we can find a sequence of measure-preserving Borel maps $\{r_n^{\varepsilon}(\cdot)\}_{n \in \mathbb{N}}$ such that

$$\mathbb{P} \left(\left\{ \omega \in \Omega : |X(\omega) - e_s \circ \mathcal{V}_{\varepsilon} \circ r_n^{\varepsilon}(\omega)| \leq \frac{1}{n} \right\} \right) = 1,$$

and we set $Y_t^{\varepsilon, n} = e_t \circ \mathcal{V}_{\varepsilon} \circ r_n^{\varepsilon}$ for all $t \in [s, T]$. In particular, $Y_t^{\varepsilon, n} \# \mathbb{P} = \mu_t^{\varepsilon}$ for all $t \in [s, T]$. We then have

$$\lim_{t \rightarrow s^+} \left\| \frac{Y_t^{\varepsilon, n} - Y_s^{\varepsilon, n}}{t - s} - v_{\varepsilon} \circ Y_s^{\varepsilon, n} \right\|_{L^2_{\mathbb{P}}} = \lim_{t \rightarrow s^+} \left\| \frac{e_t - e_s}{t - s} - v_{\varepsilon} \circ e_s \right\|_{L^2_{\eta^{\varepsilon}}} = 0.$$

Recalling the choice of \hat{v}_{ε} , we have also

$$\|v_{\varepsilon} \circ X - \hat{v}_{\varepsilon} \circ X\|_{L^2_{\mathbb{P}}} = \|v_{\varepsilon} \circ Y_s^{\varepsilon, n} - \hat{v}_{\varepsilon} \circ Y_s^{\varepsilon, n}\|_{L^2_{\mathbb{P}}} = \|v_{\varepsilon} - \hat{v}_{\varepsilon}\|_{L^2_{\mu}} \leq \varepsilon.$$

Since, by Lemma A.2, $\|Y_s^{\varepsilon, n_h} - X\|_{L^2_{\mathbb{P}}} \leq \frac{1}{n}$, we can find a subsequence $\{Y_s^{\varepsilon, n_h}\}_{h \in \mathbb{N}}$ such that for \mathbb{P} -a.e. $\omega \in \Omega$ it holds $\lim_{h \rightarrow +\infty} Y_s^{\varepsilon, n_h}(\omega) = X(\omega)$. Therefore,

$$\lim_{h \rightarrow +\infty} \int |\hat{v}_\varepsilon \circ Y_s^{\varepsilon, n_h}(\omega) - \hat{v}_\varepsilon \circ X(\omega)|^2 d\mathbb{P}(\omega) = 0, \quad (6.7)$$

where we used the Dominated Convergence Theorem to pass to the limit under the integral sign, exploiting the global boundedness of \hat{v}_ε .

From the Dynamic Programming Principle, for all $t \in [s, T]$ we have

$$\begin{aligned} 0 &\leq U(t, Y_t^{\varepsilon, n_h}) - U(s, Y_s^{\varepsilon, n_h}) + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau \\ &= U(t, Y_t^{\varepsilon, n_h}) - U(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau \leq \phi(t, Y_t^{\varepsilon, n_h}) - \phi(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau. \end{aligned}$$

Therefore, there exists a continuous increasing function $\varrho : [0, +\infty[\rightarrow [0, +\infty[$ with $\varrho(k)/k \rightarrow 0$ as $k \rightarrow 0^+$ such that we have

$$\begin{aligned} 0 &\leq \phi(t, Y_t^{\varepsilon, n_h}) - \phi(s, X) + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau \\ &\leq \partial_t \phi(s, X)(t-s) + \langle D\phi(s, X), Y_t^{\varepsilon, n_h} - X \rangle_{L^2_{\mathbb{P}}} + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \\ &\quad + \varrho \left(|t-s| \left(1 + \left\| \frac{Y_t^{\varepsilon, n_h} - Y_s^{\varepsilon, n_h}}{t-s} \right\|_{L^2_{\mathbb{P}}} \right) + \frac{1}{n_h} \right) \\ &\leq \partial_t \phi(s, X)(t-s) + \langle D\phi(s, X), Y_t^{\varepsilon, n_h} - Y_s^{\varepsilon, n_h} \rangle_{L^2_{\mathbb{P}}} + \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \frac{1}{n_h} \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} \\ &\quad + \varrho \left(|t-s| \left(1 + \left\| \frac{e_t - e_s}{t-s} \right\|_{L^2_{\mathcal{V}^\varepsilon}} \right) + \frac{1}{n_h} \right). \end{aligned}$$

Dividing by $t-s > 0$, and recalling the choice of v_ε , we have

$$\begin{aligned} 0 &\leq \partial_t \phi(s, X) + \langle D\phi(s, X), \frac{Y_t^{\varepsilon, n_h} - Y_s^{\varepsilon, n_h}}{t-s} \rangle_{L^2_{\mathbb{P}}} + \frac{1}{t-s} \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \\ &\quad + \frac{1}{n_h} \cdot \frac{1}{t-s} \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} + \frac{1}{t-s} \varrho \left(|t-s| \left(1 + \left\| \frac{e_t - e_s}{t-s} \right\|_{L^2_{\mathcal{V}^\varepsilon}} \right) + \frac{1}{n_h} \right) \\ &\leq \partial_t \phi(s, X) + \langle D\phi(s, X), v_\varepsilon \circ X \rangle_{L^2_{\mathbb{P}}} + \\ &\quad + \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} \left(\|v_\varepsilon \circ X - \hat{v}_\varepsilon \circ X\|_{L^2_{\mathbb{P}}} + \|\hat{v}_\varepsilon \circ X - \hat{v}_\varepsilon \circ Y_s^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \right. \\ &\quad \left. + \|\hat{v}_\varepsilon \circ Y_s^{\varepsilon, n_h} - v_\varepsilon \circ Y_s^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \left\| \frac{Y_t^{\varepsilon, n_h} - Y_s^{\varepsilon, n_h}}{t-s} - v_\varepsilon \circ Y_s^{\varepsilon, n_h} \right\|_{L^2_{\mathbb{P}}} \right) + \\ &\quad + \frac{1}{t-s} \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \frac{1}{n_h} \cdot \frac{1}{t-s} \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} + \\ &\quad + \frac{1}{t-s} \varrho \left(|t-s| \left(1 + \left\| \frac{e_t - e_s}{t-s} \right\|_{L^2_{\mathcal{V}^\varepsilon}} \right) + \frac{1}{n_h} \right) \end{aligned}$$

$$\begin{aligned} &\leq \partial_t \phi(s, X) + \inf_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(s, X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) + \frac{\varepsilon}{2} + \\ &\quad + \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} \left(2\varepsilon + \|\hat{v}_\varepsilon \circ X - \hat{v}_\varepsilon \circ Y_s^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \left\| \frac{e_t - e_s}{t - s} - v_\varepsilon \circ e_s \right\|_{L^2_{\eta^\varepsilon}} \right) + \\ &\quad + \frac{1}{t - s} \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \frac{1}{n_h} \cdot \frac{1}{t - s} \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} + \frac{1}{t - s} \varrho \left(|t - s| \left(1 + \left\| \frac{e_t - e_s}{t - s} \right\|_{L^2_{\eta^\varepsilon}} \right) + \frac{1}{n_h} \right). \end{aligned}$$

By letting $h \rightarrow +\infty$ and thanks to (6.7), we have

$$\begin{aligned} 0 &\leq \partial_t \phi(s, X) + \inf_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(s, X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) + \frac{\varepsilon}{2} + \\ &\quad + \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} \left(2\varepsilon + \left\| \frac{e_t - e_s}{t - s} - v_\varepsilon \circ e_s \right\|_{L^2_{\eta^\varepsilon}} \right) + \\ &\quad + \frac{1}{t - s} \int_s^t d_{\mathcal{X}}(\mu_\tau^\varepsilon) d\tau + \frac{1}{t - s} \varrho \left(|t - s| \left(1 + \left\| \frac{e_t - e_s}{t - s} \right\|_{L^2_{\eta^\varepsilon}} \right) \right). \end{aligned}$$

By letting $t \rightarrow s^+$ and recalling the boundedness of $\left\| \frac{e_t - e_s}{t - s} \right\|_{L^2_{\eta^\varepsilon}}$ coming from Proposition 3.4, we have

$$0 \leq \partial_t \phi(s, X) + \inf_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(s, X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) + \frac{\varepsilon}{2} + 2\varepsilon \|D\phi(s, X)\|_{L^2_{\mathbb{P}}} + d_{\mathcal{X}}(\mu).$$

Finally, letting $\varepsilon \rightarrow 0^+$ yields

$$0 \leq \partial_t \phi(s, X) + \inf_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(s, X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) + d_{\mathcal{X}}(\mu_s),$$

i.e., in view of Definition 6.1, $-\partial_t \phi(s, X) + H_F^{\text{viab}}(X, D\phi(s, X)) \leq 0$.

The proof of item (2) is omitted since it is a straightforward adaption of the previous argument just provided for item (1). We specify that, in this case, the proofs of the assertions regarding subsolutions and supersolutions are reversed, minimum has to be replaced by maximum and vice versa, the inequality signs are reversed and the signs of the terms involving ρ and ε need to be changed accordingly. \square

We finish the section with our main results: a viscosity characterization of viability (Theorem 6.6) and invariance (Theorem 6.7).

Theorem 6.6. (Characterization of viability) *Assume $(F_1) - (F_2)$ and let $L = \text{Lip}(F)$ and $\mathcal{H}_F^{\text{viab}}$ as in Definition 6.1. Consider a W_2 -closed subset $\mathcal{X} \subseteq \mathcal{P}_2(\mathbb{R}^d)$. The following are equivalent:*

(1) the function $z : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, defined by $z(t, \mu) := d_{\mathcal{H}}(\mu)$, is a viscosity supersolution of

$$(L+2)u(t, \mu) + \mathcal{H}_F^{\text{viab}}(\mu, D_\mu u(t, \mu)) = 0, \quad \text{in } [0, T] \times \mathcal{P}_2(\mathbb{R}^d); \quad (6.8)$$

(2) there exists $T > 0$ such that the function $w : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, defined by

$$w(t, \mu) := \frac{e^{-(L+1)(t-T)} - 1}{L+1} d_{\mathcal{H}}(\mu), \quad (6.9)$$

is a viscosity supersolution of

$$-\partial_t u(t, \mu) + \mathcal{H}_F^{\text{viab}}(\mu, D_\mu u(t, \mu)) = 0, \quad \text{in } [0, T] \times \mathcal{P}_2(\mathbb{R}^d); \quad (6.10)$$

(3) \mathcal{H} is viable for the dynamics F .

Proof. For any $T > 0$, consider the decreasing function $\alpha : [0, T] \rightarrow \mathbb{R}$ defined as

$$\alpha(t) = \frac{e^{-(L+1)(t-T)} - 1}{L+1}. \quad (6.11)$$

We denote by $W(t, X) := w(t, X \# \mathbb{P})$ the lift of $w(\cdot)$ according to Definition 5.1(1).

Proof of (1) \Rightarrow 2). Let $d_{\mathcal{H}}$ be a supersolution to (6.8) (cf. Remark 5.3). Fix $t \in [0, T]$, μ and $X \in L_{\mathbb{P}}^2(\Omega)$. Let $\Psi : [0, T] \times L_{\mathbb{P}}^2(\Omega) \rightarrow \mathbb{R}$ be a C^1 test function such that $W - \Psi$ has a local minimum at (t, X) . We want to prove that

$$-\partial_t \Psi(t, X) + H_F^{\text{viab}}(X, D\Psi(t, X)) \geq 0.$$

Since $s \mapsto \alpha(s)d_{\mathcal{H}}(Y \# \mathbb{P}) = W(s, Y)$ is regular for any $Y \in L_{\mathbb{P}}^2(\Omega)$, then by the minimality we should have

$$\partial_s \Psi(t, X) = \partial_s W(t, X), \quad \text{i.e. } \partial_s \Psi(t, X) = \dot{\alpha}(t)d_{\mathcal{H}}(X \# \mathbb{P}).$$

Hence, for all $(s, Y) \in [0, T] \times L_{\mathbb{P}}^2(\Omega)$ in a small enough neighborhood $I_{t, X}$ of (t, X) ,

$$\begin{aligned} \Psi(s, Y) &= \alpha(s)\varphi(Y) + g(s, Y), \\ \text{with } \varphi &\in C^1(L_{\mathbb{P}}^2(\Omega)) \text{ s.t. } \varphi(X) = d_{\mathcal{H}}(X \# \mathbb{P}), \\ g &\in C^1([0, T] \times L_{\mathbb{P}}^2(\Omega)) \text{ s.t. } \partial_t g(t, X) = 0, \end{aligned} \quad (6.12)$$

and φ, g such that

$$W(s, Y) - \Psi(s, Y) \geq W(t, X) - \Psi(t, X), \quad (6.13)$$

by local minimality of (t, X) . By definition of W and (6.13), we get

$$\alpha(s)[d_{\mathcal{H}}(Y \# \mathbb{P}) - \varphi(Y)] \geq g(s, Y) - g(t, X),$$

for any $(s, Y) \in I_{t,X}$. In particular, by choosing $s = t$, we obtain

$$d_{\mathcal{X}}(Y \sharp \mathbb{P}) \geq \varphi(Y) + \frac{1}{\alpha(t)}[g(t, Y) - g(t, X)],$$

with equality holding when $Y = X$. Thus, denoting with $\Phi_t : L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ the function given by

$$\Phi_t(Y) := \varphi(Y) + \frac{1}{\alpha(t)}[g(t, Y) - g(t, X)],$$

we notice that $\Phi_t \in C^1(L^2_{\mathbb{P}}(\Omega))$ and that the map $Y \mapsto d_{\mathcal{X}}(Y \sharp \mathbb{P}) - \Phi_t(Y)$ attains a local minimum in X . Thus, recalling also Remark 5.3, we can employ Φ_t as a test function for $d_{\mathcal{X}}$ to get

$$(L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{viab}}(X, D\Phi_t(X)) \geq 0. \tag{6.14}$$

Notice that by (6.12),

$$\begin{aligned} \partial_t \Psi(t, X) &= \dot{\alpha}(t)d_{\mathcal{X}}(X \sharp \mathbb{P}) = -[(L + 1)\alpha(t) + 1]d_{\mathcal{X}}(X \sharp \mathbb{P}), \\ D\Psi(t, X) &= \alpha(t)D\varphi(X) + Dg(t, X) = \alpha(t)D\Phi_t(X). \end{aligned} \tag{6.15}$$

Recalling the definition of the lifted Hamiltonian H_F^{viab} , by (6.14) we obtain

$$\begin{aligned} 0 &\leq (L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{viab}}\left(X, \frac{1}{\alpha(t)}D\Psi(t, X)\right) \\ &= (L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) - d_{\mathcal{X}}(X \sharp \mathbb{P}) - \frac{1}{\alpha(t)} \inf_{\substack{v \in L^2_{\mathbb{P}}(\Omega) \\ v(\cdot) \in F(X \sharp \mathbb{P}, X(\cdot))}} \int_{\Omega} \langle v(\omega), D\Psi(t, X)(\omega) \rangle d\mathbb{P}(\omega). \end{aligned}$$

Multiplying by $\alpha(t)$, we finally get

$$[(L + 1)\alpha(t) + 1]d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{viab}}(X, D\Psi(t, X)) \geq 0,$$

thus

$$-\partial_t \Psi(t, X) + H_F^{\text{viab}}(X, D\Psi(t, X)) \geq 0,$$

which concludes that w is a supersolution of (6.10).

Proof of (2 \Rightarrow 3). Let $T > 0$ and assume that $w(t, \mu) = \alpha(t)d_{\mathcal{X}}(\mu)$ is a viscosity supersolution of (6.10). We recall that H_F^{viab} , given in Definition 6.1, satisfies the assumptions of Theorem 5.4 as proved in Lemma 6.3. In particular, if we denote by $U(t, X) := V^{\text{viab}}(t, X \sharp \mathbb{P})$ the lift of the value function of Definition 4.1, we have

$$W(T, X) = U(T, X) = 0, \text{ for every } X \in L^2_{\mathbb{P}}(\Omega).$$

Therefore, since both w and V^{viab} are uniformly continuous (see Proposition 4.6), by Theorem 5.4 and Proposition 6.5, we have $U(t, X) \leq W(t, X)$ for all $(t, X) \in$

$[0, T] \times L^2_{\mathbb{P}}(\Omega)$. Thus, for all $\mu \in \mathcal{X}$ and all $X \in L^2_{\mathbb{P}}(\Omega)$ with $X \# \mathbb{P} = \mu$ we obtain $V^{\text{viab}}(t, \mu) = U(t, X) = W(t, X) = 0$ for all $t \in [0, T]$. By Proposition 4.3, we conclude that there exists an admissible trajectory starting from μ and defined on $[0, T]$, which is entirely contained in \mathcal{X} . So \mathcal{X} is viable.

Proof of (3 \Rightarrow 1). Assume that \mathcal{X} is viable. Set $\hat{d}_{\mathcal{X}}(Y) := d_{\mathcal{X}}(Y \# \mathbb{P})$ for all $Y \in L^2_{\mathbb{P}}(\Omega)$, i.e., $\hat{d}_{\mathcal{X}}$ is the lift of $d_{\mathcal{X}}$. Let $\phi \in C^1(L^2_{\mathbb{P}}(\Omega))$ and $X \in L^2_{\mathbb{P}}(\Omega)$ be such that $\hat{d}_{\mathcal{X}} - \phi$ has a local minimum at X , and set $\mu = X \# \mathbb{P} \in \mathcal{P}_2(\mathbb{R}^d)$.

For any $\varepsilon > 0$ and $T > 0$, there exist $\bar{\mu}^\varepsilon \in \mathcal{X}$, and $\bar{\mu}^\varepsilon \in \mathcal{A}_{[0, T]}(\bar{\mu}^\varepsilon)$ satisfying $W_2(\mu, \bar{\mu}^\varepsilon) \leq d_{\mathcal{X}}(\mu) + \varepsilon$ and $\bar{\mu}^\varepsilon \subseteq \mathcal{X}$. By Grönwall's inequality (Lemma 3.3), there exists $\mu^\varepsilon \in \mathcal{A}_{[0, T]}(\mu)$, $\eta^\varepsilon \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[0, T]})$ such that $\mu^\varepsilon = e_t \# \eta^\varepsilon$, and

$$d_{\mathcal{X}}(\mu^\varepsilon) \leq W_2(\mu^\varepsilon, \bar{\mu}_t^\varepsilon) \leq e^{Lt+te^{Lt}} \cdot W_2(\mu, \bar{\mu}^\varepsilon) \leq e^{Lt+te^{Lt}} \cdot (d_{\mathcal{X}}(\mu) + \varepsilon),$$

for all $t \in [0, T]$.

According to Corollary A.3 applied to μ^ε , set

$$\widehat{\omega}(t) := \frac{L}{t} \int_0^t \left[W_2(\mu_\tau^\varepsilon, \mu_0^\varepsilon) + \|e_\tau - e_0\|_{L^2_{\eta^\varepsilon}} \right] d\tau,$$

there exists a family $\{Y_t^\varepsilon\}_{t \in [0, T]} \subseteq L^2_{\mathbb{P}}(\Omega)$ satisfying $Y_t^\varepsilon \# \mathbb{P} = \mu_t^\varepsilon$ for all $t \in [0, T]$ and

$$\begin{aligned} \left\langle p, \frac{Y_t^\varepsilon - Y_0^\varepsilon}{t} \right\rangle_{L^2_{\mathbb{P}}} &\geq \int_{\Omega} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle p(\omega), v \rangle d\mathbb{P}(\omega) - (\widehat{\omega}(t) + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}} \\ &= -d_{\mathcal{X}}(X \# \mathbb{P}) - H_F^{\text{viab}}(X, p) - (\widehat{\omega}(t) + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}} \end{aligned}$$

for any $p \in L^2_{\mathbb{P}}(\Omega)$ (recall that $\mu = \mu_0 = X \# \mathbb{P} = Y_0^\varepsilon \# \mathbb{P}$). According to the choice of X , we have

$$\frac{d_{\mathcal{X}}(\mu_t^\varepsilon) - d_{\mathcal{X}}(\mu)}{t} = \frac{\hat{d}_{\mathcal{X}}(Y_t^\varepsilon) - \hat{d}_{\mathcal{X}}(X)}{t} \geq \frac{\phi(Y_t^\varepsilon) - \phi(X)}{t}. \quad (6.16)$$

We estimate the first term as follows

$$\begin{aligned} \frac{d_{\mathcal{X}}(\mu_t^\varepsilon) - d_{\mathcal{X}}(\mu)}{t} &\leq \frac{W_2(\mu_t^\varepsilon, \bar{\mu}_t^\varepsilon) - W_2(\mu, \bar{\mu}^\varepsilon)}{t} + \frac{\varepsilon}{t} \leq \frac{e^{Lt+te^{Lt}} - 1}{t} \cdot W_2(\mu, \bar{\mu}^\varepsilon) + \frac{\varepsilon}{t} \\ &\leq \frac{e^{Lt+te^{Lt}} - 1}{t} \cdot (d_{\mathcal{X}}(X \# \mathbb{P}) + \varepsilon) + \frac{\varepsilon}{t}. \end{aligned}$$

Concerning the right hand side of (6.16), we have that there exists a continuous increasing map $\varrho : [0, +\infty) \rightarrow [0, +\infty)$ with $\varrho(r)/r \rightarrow 0$ as $r \rightarrow 0^+$ such that

$$\begin{aligned} \frac{\phi(Y_t^\varepsilon) - \phi(X)}{t} &\geq \langle D\phi(X), \frac{Y_t^\varepsilon - X}{t} \rangle_{L^2_{\mathbb{P}}} - \frac{\varrho(\|Y_t^\varepsilon - X\|_{L^2_{\mathbb{P}}} + t)}{t} \\ &\geq \langle D\phi(X), \frac{Y_t^\varepsilon - Y_0^\varepsilon}{t} \rangle_{L^2_{\mathbb{P}}} - \|D\phi(X)\|_{L^2_{\mathbb{P}}} \cdot \frac{\|X - Y_0^\varepsilon\|_{L^2_{\mathbb{P}}}}{t} + \\ &\quad - \frac{1}{t} \varrho(\|Y_t^\varepsilon - Y_0^\varepsilon\|_{L^2_{\mathbb{P}}} + \|Y_0^\varepsilon - X\|_{L^2_{\mathbb{P}}} + t) \\ &\geq \langle D\phi(X), \frac{Y_t^\varepsilon - Y_0^\varepsilon}{t} \rangle_{L^2_{\mathbb{P}}} - \frac{\varepsilon}{t} \|D\phi(X)\|_{L^2_{\mathbb{P}}} - \frac{\varrho(\|e_t - e_0\|_{L^2_{\mathbb{P}}} + t + \varepsilon)}{t} \\ &\geq -d_{\mathcal{X}}(X \sharp \mathbb{P}) - H_F^{\text{viab}}(X, D\phi(X)) - \left(\widehat{\omega}(t) + L\varepsilon + \frac{\varepsilon}{t} \right) \|D\phi(X)\|_{L^2_{\mathbb{P}}} + \\ &\quad - \frac{\varrho\left(t\left(\| \frac{e_t - e_0}{t} \|_{L^2_{\mathbb{P}}} + 1\right) + \varepsilon\right)}{t}, \end{aligned}$$

where in the third inequality we employed the definition of Y_t^ε provided in the proof of Corollary A.3, i.e., $Y_t^\varepsilon = e_t \circ \mathcal{W}_\varepsilon$ for any $t \in [0, T]$, for some $\mathcal{W}_\varepsilon : \Omega \rightarrow \mathbb{R}^d \times \Gamma_{[0, T]}$ s.t. $\mathcal{W}_\varepsilon \sharp \mathbb{P} = \eta^\varepsilon$. Recalling now the uniform boundedness in ε of $\| \frac{e_t - e_0}{t} \|_{L^2_{\mathbb{P}}}$ coming from Proposition 3.4(3), by letting $\varepsilon \rightarrow 0^+$ and $t \rightarrow 0^+$, and by setting

$$\ell := \liminf_{t \rightarrow 0^+} \liminf_{\varepsilon \rightarrow 0^+} \frac{\phi(Y_t^\varepsilon) - \phi(X)}{t},$$

we have

$$-d_{\mathcal{X}}(X \sharp \mathbb{P}) - H_F^{\text{viab}}(X, D\phi(X)) \leq \ell \leq (L+1) \cdot d_{\mathcal{X}}(X \sharp \mathbb{P}).$$

This leads to $(L+2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{viab}}(X, D\phi(X)) \geq 0$, i.e., $d_{\mathcal{X}}(\mu)$ is a supersolution of (6.8). \square

Theorem 6.7. (Characterization of invariance) *Assume $(F_1) - (F_2)$ and let $L = \text{Lip}(F)$ and $\mathcal{H}_F^{\text{inv}}$ as in Definition 6.2. Consider a W_2 -closed subset $\mathcal{X} \subseteq \mathcal{P}_2(\mathbb{R}^d)$. The following is equivalent:*

- (1) *the function $z : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, defined by $z(t, \mu) := d_{\mathcal{X}}(\mu)$, is a viscosity supersolution of*

$$(L+2)u(t, \mu) + \mathcal{H}_F^{\text{inv}}(\mu, D_\mu u(t, \mu)) = 0 \quad \text{in } [0, T] \times \mathcal{P}_2(\mathbb{R}^d); \quad (6.17)$$

- (2) *there exists $T > 0$ such that the function $w : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, defined by (6.9), is a viscosity supersolution of*

$$-\partial_t u(t, \mu) + \mathcal{H}_F^{\text{inv}}(\mu, D_\mu u(t, \mu)) = 0 \quad \text{in } [0, T] \times \mathcal{P}_2(\mathbb{R}^d); \quad (6.18)$$

- (3) *\mathcal{X} is invariant for the dynamics F .*

Proof. For any $T > 0$, consider the decreasing function $\alpha : [0, T] \rightarrow \mathbb{R}$ defined as in (6.11).

We denote by $W(t, X) := w(t, X \sharp \mathbb{P})$ the lift of $w(\cdot)$ defined in (6.9) according to Definition 5.1(1).

Proof of (1 \Rightarrow 2). This part of the proof is the same as the one developed in Theorem 6.6 with H_F^{inv} in place of H_F^{viab} .

Proof of (2 \Rightarrow 3). Same as in Theorem 6.6, with V^{viab} replaced by V^{inv} .

Proof of (3 \Rightarrow 1). Assume that \mathcal{X} is invariant. Set $\hat{d}_{\mathcal{X}}(Y) = d_{\mathcal{X}}(Y \sharp \mathbb{P})$ for all $Y \in L_{\mathbb{P}}^2(\Omega)$, i.e., $\hat{d}_{\mathcal{X}}$ is the lift of $d_{\mathcal{X}}$. Let $\phi \in C^1(L_{\mathbb{P}}^2(\Omega))$ and $X \in L_{\mathbb{P}}^2(\Omega)$ be such that $\hat{d}_{\mathcal{X}} - \phi$ has a local minimum at X , and set $\mu = X \sharp \mathbb{P} \in \mathcal{P}_2(\mathbb{R}^d)$.

Fix $\varepsilon > 0$, and let $v_\varepsilon \in L_{\mu}^2(\mathbb{R}^d)$ be such that $v_\varepsilon(x) \in F(\mu, x)$ for μ -a.e. $x \in \mathbb{R}^d$ and

$$\int_{\Omega} \langle D\phi(X)(\omega), v_\varepsilon \circ X(\omega) \rangle d\mathbb{P}(\omega) \geq \sup_{\substack{v \in L_{X \sharp \mathbb{P}}^2(\mathbb{R}^d) \\ v(\cdot) \in F(X \sharp \mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) - \frac{\varepsilon}{2}.$$

By Remark 6.4, we can suppose that $v_\varepsilon \in C^0$, and by Lemma A.4 there exists an admissible trajectory $\mu^\varepsilon = \{\mu_t^\varepsilon\}_{t \in [0, T]}$ defined on $[0, T]$ with $\mu_s^\varepsilon = \mu$, and $\eta^\varepsilon \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[0, T]})$ such that $e_t \sharp \eta^\varepsilon = \mu_t^\varepsilon$ for all $t \in [0, T]$ and

$$\lim_{t \rightarrow 0^+} \left\| \frac{e_t - e_0}{t} - v_\varepsilon \circ e_0 \right\|_{L_{\eta^\varepsilon}^2} = 0.$$

By density, we can find $\hat{v}_\varepsilon \in C_b^0(\mathbb{R}^d)$ such that $\|v_\varepsilon - \hat{v}_\varepsilon\|_{L_{\mu}^2} \leq \varepsilon$.

Denote by $\mathcal{Y}_\varepsilon : \Omega \rightarrow \mathbb{R}^d \times \Gamma_{[0, T]}$ a Borel map satisfying $\eta^\varepsilon = \mathcal{Y}_\varepsilon \sharp \mathbb{P}$. Recalling Lemma A.2, since for all $\varepsilon > 0$ we have $\mu = \mu_0^\varepsilon = e_0 \sharp \eta^\varepsilon = (e_0 \circ \mathcal{Y}_\varepsilon) \sharp \mathbb{P} = X \sharp \mathbb{P}$, we can find a sequence of measure-preserving Borel maps $\{r_n^\varepsilon(\cdot)\}_{n \in \mathbb{N}}$ such that

$$\mathbb{P} \left(\left\{ \omega \in \Omega : |X(\omega) - e_0 \circ \mathcal{Y}_\varepsilon \circ r_n^\varepsilon(\omega)| \leq \frac{1}{n} \right\} \right) = 1,$$

and we set $Y_t^{\varepsilon, n} = e_t \circ \mathcal{Y}_\varepsilon \circ r_n^\varepsilon$ for all $t \in [0, T]$. In particular, $Y_t^{\varepsilon, n} \sharp \mathbb{P} = \mu_t^\varepsilon$ for all $t \in [0, T]$. We then have

$$\lim_{t \rightarrow 0^+} \left\| \frac{Y_t^{\varepsilon, n} - Y_0^{\varepsilon, n}}{t} - v_\varepsilon \circ Y_0^{\varepsilon, n} \right\|_{L_{\mathbb{P}}^2} = \lim_{t \rightarrow 0^+} \left\| \frac{e_t - e_0}{t} - v_\varepsilon \circ e_0 \right\|_{L_{\eta^\varepsilon}^2} = 0.$$

Recalling the choice of \hat{v}_ε , we have also

$$\|v_\varepsilon \circ X - \hat{v}_\varepsilon \circ X\|_{L_{\mathbb{P}}^2} = \|v_\varepsilon \circ Y_0^{\varepsilon, n} - \hat{v}_\varepsilon \circ Y_0^{\varepsilon, n}\|_{L_{\mathbb{P}}^2} = \|v_\varepsilon - \hat{v}_\varepsilon\|_{L_{\mu}^2} \leq \varepsilon.$$

Since, by Lemma A.2, $\|Y_0^{\varepsilon, n} - X\|_{L_{\mathbb{P}}^2} \leq \frac{1}{n}$, we can find a subsequence $\{Y_0^{\varepsilon, n_h}\}_{h \in \mathbb{N}}$ such that for \mathbb{P} -a.e. $\omega \in \Omega$ it holds $\lim_{h \rightarrow +\infty} Y_0^{\varepsilon, n_h}(\omega) = X(\omega)$. Therefore,

$$\lim_{h \rightarrow +\infty} \int |\hat{v}_\varepsilon \circ Y_0^{\varepsilon, n_h}(\omega) - \hat{v}_\varepsilon \circ X(\omega)|^2 d\mathbb{P}(\omega) = 0, \quad (6.19)$$

where we used the Dominated Convergence Theorem to pass to the limit under the integral sign, exploiting the global boundedness of \hat{v}_ε .

Now, let $\bar{\mu}^{nh} \in \mathcal{X}$ such that $W_2(\mu, \bar{\mu}^{nh}) \leq d_{\mathcal{X}}(\mu) + \frac{1}{n_h}$. By Grönwall's inequality (Lemma 3.3), given μ^ε as before there exist $\bar{\mu}^{\varepsilon, nh} \in \mathcal{A}_{[0, T]}(\bar{\mu}^{nh})$ such that

$$d_{\mathcal{X}}(\mu_t^\varepsilon) \leq W_2(\mu_t^\varepsilon, \bar{\mu}_t^{\varepsilon, nh}) \leq e^{Lt+te^{Lt}} \cdot W_2(\mu, \bar{\mu}^{nh}) \leq e^{Lt+te^{Lt}} \cdot \left(d_{\mathcal{X}}(\mu) + \frac{1}{n_h} \right),$$

for all $t \in [0, T]$, where we used the fact that $\bar{\mu}^{\varepsilon, nh} \subseteq \mathcal{X}$ by invariance of the set \mathcal{X} and since $\bar{\mu}_0^{\varepsilon, nh} = \bar{\mu}^{nh} \in \mathcal{X}$. According to the choice of X , we have

$$\frac{d_{\mathcal{X}}(\mu_t^\varepsilon) - d_{\mathcal{X}}(\mu)}{t} = \frac{\hat{d}_{\mathcal{X}}(Y_t^{\varepsilon, nh}) - \hat{d}_{\mathcal{X}}(X)}{t} \geq \frac{\phi(Y_t^{\varepsilon, nh}) - \phi(X)}{t}. \quad (6.20)$$

We estimate the first term as follows

$$\begin{aligned} \frac{d_{\mathcal{X}}(\mu_t^\varepsilon) - d_{\mathcal{X}}(\mu)}{t} &\leq \frac{W_2(\mu_t^\varepsilon, \bar{\mu}_t^{\varepsilon, nh}) - W_2(\mu, \bar{\mu}^{nh})}{t} + \frac{1}{n_h} \cdot \frac{1}{t} \\ &\leq \frac{e^{Lt+te^{Lt}} - 1}{t} \cdot W_2(\mu, \bar{\mu}^{nh}) + \frac{1}{n_h} \cdot \frac{1}{t} \leq \frac{e^{Lt+te^{Lt}} - 1}{t} \cdot \left(d_{\mathcal{X}}(X \# \mu) + \frac{1}{n_h} \right) + \frac{1}{n_h} \cdot \frac{1}{t}. \end{aligned} \quad (6.21)$$

Concerning the right hand side of (6.20), we have that there exists a continuous increasing map $\varrho : [0, +\infty) \rightarrow [0, +\infty)$ with $\varrho(r)/r \rightarrow 0$ as $r \rightarrow 0^+$ such that

$$\begin{aligned} \frac{\phi(Y_t^{\varepsilon, nh}) - \phi(X)}{t} &\geq \langle D\phi(X), \frac{Y_t^{\varepsilon, nh} - X}{t} \rangle_{L_{\mathbb{P}}^2} - \frac{\varrho(\|Y_t^{\varepsilon, nh} - X\|_{L_{\mathbb{P}}^2} + t)}{t} \\ &\geq \langle D\phi(X), \frac{Y_t^{\varepsilon, nh} - Y_0^{\varepsilon, nh}}{t} \rangle_{L_{\mathbb{P}}^2} - \|D\phi(X)\|_{L_{\mathbb{P}}^2} \cdot \frac{\|X - Y_0^{\varepsilon, nh}\|_{L_{\mathbb{P}}^2}}{t} + \\ &\quad - \frac{1}{t} \varrho(\|Y_t^{\varepsilon, nh} - Y_0^{\varepsilon, nh}\|_{L_{\mathbb{P}}^2} + \|Y_0^{\varepsilon, nh} - X\|_{L_{\mathbb{P}}^2} + t) \\ &\geq \langle D\phi(X), \frac{Y_t^{\varepsilon, nh} - Y_0^{\varepsilon, nh}}{t} \rangle_{L_{\mathbb{P}}^2} - \frac{1}{n_h} \cdot \frac{1}{t} \|D\phi(X)\|_{L_{\mathbb{P}}^2} + \\ &\quad - \frac{\varrho\left(\|e_t - e_0\|_{L_{\mathbb{P}^\varepsilon}^2} + t + \frac{1}{n_h}\right)}{t}. \end{aligned}$$

Recalling the choice of v_ε , we have

$$\begin{aligned}
 \frac{\phi(Y_t^{\varepsilon, n_h}) - \phi(X)}{t} &\geq \langle D\phi(X), v_\varepsilon \circ X \rangle_{L^2_{\mathbb{P}}} + \\
 &\quad - \|D\phi(X)\|_{L^2_{\mathbb{P}}} \left(\|v_\varepsilon \circ X - \hat{v}_\varepsilon \circ X\|_{L^2_{\mathbb{P}}} + \|\hat{v}_\varepsilon \circ X - \hat{v}_\varepsilon \circ Y_0^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \right. \\
 &\quad \left. + \|\hat{v}_\varepsilon \circ Y_0^{\varepsilon, n_h} - v_\varepsilon \circ Y_0^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \left\| \frac{Y_t^{\varepsilon, n_h} - Y_0^{\varepsilon, n_h}}{t} - v_\varepsilon \circ Y_0^{\varepsilon, n_h} \right\|_{L^2_{\mathbb{P}}} \right) + \\
 &\quad - \frac{1}{n_h} \cdot \frac{1}{t} \|D\phi(X)\|_{L^2_{\mathbb{P}}} - \frac{1}{t} \varrho \left(t \left(1 + \left\| \frac{e_t - e_0}{t} \right\|_{L^2_{\mathbb{P}^\varepsilon}} \right) + \frac{1}{n_h} \right) \\
 &\geq \sup_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) - \frac{\varepsilon}{2} + \\
 &\quad - \|D\phi(X)\|_{L^2_{\mathbb{P}}} \left(2\varepsilon + \|\hat{v}_\varepsilon \circ X - \hat{v}_\varepsilon \circ Y_0^{\varepsilon, n_h}\|_{L^2_{\mathbb{P}}} + \left\| \frac{e_t - e_0}{t} - v_\varepsilon \circ e_0 \right\|_{L^2_{\mathbb{P}^\varepsilon}} \right) + \\
 &\quad - \frac{1}{n_h} \cdot \frac{1}{t} \|D\phi(X)\|_{L^2_{\mathbb{P}}} - \frac{1}{t} \varrho \left(t \left(1 + \left\| \frac{e_t - e_0}{t} \right\|_{L^2_{\mathbb{P}^\varepsilon}} \right) + \frac{1}{n_h} \right).
 \end{aligned}$$

Recalling now the uniform boundedness in ε of $\left\| \frac{e_t - e_0}{t} \right\|_{L^2_{\mathbb{P}^\varepsilon}}$ coming from Proposition 3.4(3), by letting $h \rightarrow +\infty$, $t \rightarrow 0^+$ and $\varepsilon \rightarrow 0^+$, and by setting

$$\ell := \liminf_{\varepsilon \rightarrow 0^+} \liminf_{t \rightarrow 0^+} \liminf_{h \rightarrow +\infty} \frac{\phi(Y_t^{\varepsilon, n_h}) - \phi(X)}{t},$$

we have, thanks also to (6.19),

$$\ell \geq \sup_{\substack{v \in L^2_{X\sharp\mathbb{P}}(\mathbb{R}^d) \\ v(\cdot) \in F(X\sharp\mathbb{P}, \cdot)}} \int_{\Omega} \langle D\phi(X)(\omega), v \circ X(\omega) \rangle d\mathbb{P}(\omega) = -d_{\mathcal{X}}(X\sharp\mathbb{P}) - H_F^{\text{inv}}(X, D\phi(X)). \tag{6.22}$$

Thus, by passing to the limit also in (6.21) and combining that estimate with (6.22), we get

$$-d_{\mathcal{X}}(X\sharp\mathbb{P}) - H_F^{\text{inv}}(X, D\phi(X)) \leq \ell \leq (L+1) \cdot d_{\mathcal{X}}(X\sharp\mathbb{P}).$$

This leads to $(L+2)d_{\mathcal{X}}(X\sharp\mathbb{P}) + H_F^{\text{inv}}(X, D\phi(X)) \geq 0$, i.e., $d_{\mathcal{X}}(\mu)$ is a supersolution of (6.17) (cf. Remark 5.3). \square

7. An example

Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, $u \in \mathbb{R}$, let $U = [1/2, 3/2]$, $U' = [-3/2, 3/2]$ and define the functions $f, g : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$ as

$$f(\mu, x, u) := u \arctan(1 - \mu_2^{1/2}(\mu)) e^{-|x|^2} x, \quad g(\mu, x, u) := \pi u x.$$

Define the set-valued maps $F, G : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ as

$$F(\mu, x) := \{f(\mu, x, u) : u \in U\}, \quad G(\mu, x) := \{g(\mu, x, u) : u \in U'\},$$

and the closed set

$$\mathcal{K} := \{\mu \in \mathcal{P}_2(\mathbb{R}^d) : m_2(\mu) \leq 1\} = \{X \# \mathbb{P} : \|X\|_{L^2_{\mathbb{P}}(\Omega)} \leq 1\}.$$

Notice that F, G satisfy the assumptions $(F_1) - (F_2)$ and $G(\mu, x) \supseteq F(\mu, x)$. In particular,

$$F(X \# \mathbb{P}, X(\omega)) = \left\{ \lambda \arctan(1 - \|X\|_{L^2_{\mathbb{P}}}) e^{-|X(\omega)|^2} X(\omega) : \lambda \in [1/2, 3/2] \right\}.$$

We have

$$d_{\mathcal{K}}(\mu) = \begin{cases} m_2^{1/2}(\mu) - 1, & \text{if } \mu \notin \mathcal{K}, \\ 0, & \text{if } \mu \in \mathcal{K}. \end{cases}$$

Indeed, to prove that $d_{\mathcal{K}}(\mu) \leq m_2^{1/2}(\mu) - 1$ for all $\mu \notin \mathcal{K}$, take a W_2 -geodesic $\{\xi_t\}_{t \in [0, m_2^{1/2}(\mu)]}$ with constant speed joining δ_0 to $\mu \notin \mathcal{K}$. We have $m_2^{1/2}(\xi_1) = W_2(\delta_0, \xi_1) = 1$, and $W_2(\mu, \delta_0) = W_2(\mu, \xi_1) + 1$. So $\xi_1 \in \mathcal{K}$ and $d_{\mathcal{K}}(\mu) \leq m_2^{1/2}(\mu) - 1$. Conversely, fix $\varepsilon > 0$ and let $\mu_\varepsilon \in \mathcal{K}$ be such that $d_{\mathcal{K}}(\mu) \geq W_2(\mu, \mu_\varepsilon) - \varepsilon$. Then, recalling that $W_2(\mu_\varepsilon, \delta_0) \leq 1$, we have

$$d_{\mathcal{K}}(\mu) + 1 \geq W_2(\mu, \mu_\varepsilon) + W_2(\mu_\varepsilon, \delta_0) - \varepsilon \geq W_2(\mu, \delta_0) - \varepsilon = m_2^{1/2}(\mu) - \varepsilon.$$

By letting $\varepsilon \rightarrow 0^+$, we have the desired inequality.

The lift of $d_{\mathcal{K}}(\cdot)$ is the convex function $\hat{U} : L^2_{\mathbb{P}}(\Omega) \rightarrow \mathbb{R}$ defined as

$$\hat{U}(X) = \begin{cases} \|X\|_{L^2_{\mathbb{P}}} - 1, & \text{if } \|X\|_{L^2_{\mathbb{P}}} \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The function $\hat{U}(\cdot)$ is C^1 in the open set $D := \{X \in L^2_{\mathbb{P}} : \|X\|_{L^2_{\mathbb{P}}} \neq 1\}$. Thus, if $\psi \in C^1(L^2_{\mathbb{P}}(\Omega))$ is such that $\hat{U} - \psi$ attains a local minimum at $X \in D$ then

$$D\psi(X) = D\hat{U}(X) = \begin{cases} 0, & \text{if } \|X\|_{L^2_{\mathbb{P}}} < 1, \\ \frac{X}{\|X\|_{L^2_{\mathbb{P}}}}, & \text{if } \|X\|_{L^2_{\mathbb{P}}} > 1. \end{cases}$$

Let $\psi \in C^1(L^2_{\mathbb{P}})$ such that $\hat{U} - \psi$ attains a local minimum at $X \in L^2_{\mathbb{P}}$ with $\|X\|_{L^2_{\mathbb{P}}} = 1$. By Propositions 1.2 and 1.5 in [17], we have that

$$D\psi(X) \in \partial\hat{U}(X) := \left\{ \xi \in L^2_{\mathbb{P}}(\Omega) : \hat{U}(Y) - \hat{U}(X) \geq \langle \xi, Y - X \rangle_{L^2_{\mathbb{P}}}, \forall Y \in L^2_{\mathbb{P}} \right\}.$$

Conversely, given $\xi \in \partial \hat{U}(X)$, set $\psi(Y) = \hat{U}(X) + \langle \xi, Y - X \rangle_{L_{\mathbb{P}}^2}$. Then, $\psi \in C^1$, $\hat{U} - \psi$ has a minimum at X , and $D\psi(X) = \xi$.

We want to prove that if $\|X\|_{L_{\mathbb{P}}^2} = 1$, then $\partial \hat{U}(X) = \{\lambda X : \lambda \in [0, 1]\}$.

We prove \supseteq . Given $X, Y \in L_{\mathbb{P}}^2$ with $\|X\|_{L_{\mathbb{P}}^2} = 1$, and $\lambda \in [0, 1]$, it holds

$$\langle \lambda X, Y - X \rangle_{L_{\mathbb{P}}^2} \leq \lambda(\|Y\|_{L_{\mathbb{P}}^2} - 1) \leq \begin{cases} \|Y\|_{L_{\mathbb{P}}^2} - 1 = \hat{U}(Y) - \hat{U}(X), & \text{if } \|Y\|_{L_{\mathbb{P}}^2} \geq 1, \\ 0 = \hat{U}(Y) - \hat{U}(X), & \text{if } \|Y\|_{L_{\mathbb{P}}^2} < 1. \end{cases}$$

Thus, in any case $\langle \lambda X, Y - X \rangle_{L_{\mathbb{P}}^2} \leq \hat{U}(Y) - \hat{U}(X)$, proving \supseteq .

Conversely, we prove \subseteq . Let $X \in L_{\mathbb{P}}^2(\Omega)$, $\|X\|_{L_{\mathbb{P}}^2} = 1$, so $\hat{U}(X) = 0$. Assume that $\xi = \lambda X + \hat{\lambda} Z \in \partial \hat{U}(X)$, with $\|Z\|_{L_{\mathbb{P}}^2} = 1$, $\langle Z, X \rangle_{L_{\mathbb{P}}^2} = 0$ and $\lambda, \hat{\lambda} \in \mathbb{R}$. We want to prove that $\lambda \in [0, 1]$ and $\hat{\lambda} = 0$. Indeed, for all $Y \in L_{\mathbb{P}}^2(\Omega)$ it holds

$$\hat{U}(Y) - \hat{U}(X) \geq \langle \xi, Y - X \rangle_{L_{\mathbb{P}}^2} = \langle \lambda X + \hat{\lambda} Z, Y - X \rangle_{L_{\mathbb{P}}^2} = \hat{\lambda} \langle Y, Z \rangle + \lambda(\langle Y, X \rangle_{L_{\mathbb{P}}^2} - 1).$$

By taking $Y = aX + bZ$, we have $\hat{U}(Y) = \max\{0, \sqrt{|a|^2 + |b|^2} - 1\}$, and so

$$\max\{0, \sqrt{|a|^2 + |b|^2} - 1\} \geq b\hat{\lambda} + \lambda(a - 1).$$

- Choosing $(a, b) = (2, 0)$ leads to $\lambda \leq 1$. Choosing $(a, b) = (1/2, 0)$ leads to $\lambda \geq 0$. Therefore, $0 \leq \lambda \leq 1$.
- Choose $a = 1$. Then for all $b > 0$, we have $\frac{\sqrt{1 + b^2} - 1}{b} \geq \hat{\lambda}$, and by passing to the limit as $b \rightarrow 0^+$ we have $0 \geq \hat{\lambda}$. For all $b < 0$, we have $\frac{\sqrt{1 + b^2} - 1}{b} \leq \hat{\lambda}$, and by passing to the limit as $b \rightarrow 0^-$ we have $0 \leq \hat{\lambda}$. Therefore, $\hat{\lambda} = 0$.

We prove now that \mathcal{X} is invariant for the dynamics F . Thanks to Theorem 6.7, we have to prove that for every $\psi \in C^1(L_{\mathbb{P}}^2)$ such that $\hat{U} - \psi$ attains a local minimum at $X \in L_{\mathbb{P}}^2$ it holds

$$(L + 2)d_{\mathcal{X}}(X \# \mathbb{P}) + H_F^{\text{inv}}(X, D\psi(X)) \geq 0.$$

We distinguish two cases

- when $\|X\|_{L_{\mathbb{P}}^2} < 1$, we have $d_{\mathcal{X}}(X \# \mathbb{P}) = 0$ and $D\psi(X) = 0$, which implies $H_F^{\text{inv}}(X, D\psi(X)) = 0$, so the equation is trivially satisfied.
- when $\|X\|_{L_{\mathbb{P}}^2} \geq 1$, we have $d_{\mathcal{X}}(X \# \mathbb{P}) = \|X\|_{L_{\mathbb{P}}^2} - 1$ and $D\psi(X) = \lambda \frac{X}{\|X\|_{L_{\mathbb{P}}^2}}$, with $\lambda = 1$ if $\|X\|_{L_{\mathbb{P}}^2} > 1$, and $\lambda \in [0, 1]$ otherwise, which implies

$$\begin{aligned} H_F^{\text{inv}}(X, D\psi(X)) &= 1 - \|X\|_{L_{\mathbb{P}}^2} - \frac{1}{2}\lambda \int_{\Omega} \arctan(1 - \|X\|_{L_{\mathbb{P}}^2}) e^{-|X(\omega)|^2} |X(\omega)|^2 d\mathbb{P}(\omega) \\ &\geq 1 - \|X\|_{L_{\mathbb{P}}^2}, \end{aligned}$$

So, also in this case, we have

$$(L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{inv}}(X, D\psi(X)) \geq (L + 2)(\|X\|_{L_{\mathbb{P}}^2} - 1) + 1 - \|X\|_{L_{\mathbb{P}}^2} \geq 0,$$

from which we get the invariance, and thus the viability, of the set \mathcal{X} for the dynamics F . Since all the admissible trajectories for F are also admissible for G , we have that \mathcal{X} is viable for G . We prove now that \mathcal{X} is not invariant for G . Indeed, take $X \in L_{\mathbb{P}}^2(\Omega)$ with $\|X\|_{L_{\mathbb{P}}^2} = 1$. Then, we can consider $\psi \in C^1(L_{\mathbb{P}}^2(\Omega))$ s.t. $\psi(Y) = \|Y\|_{L_{\mathbb{P}}^2}$ in a neighborhood V of X . Given $Y \in V$, we have $\hat{U}(Y) - \psi(Y) = -1$ if $\|Y\|_{L_{\mathbb{P}}^2} \geq 1$ and $\hat{U}(Y) - \psi(Y) = -\|Y\|_{L_{\mathbb{P}}^2} \geq -1$ if $\|Y\|_{L_{\mathbb{P}}^2} < 1$. In particular, $\hat{U}(X) - \psi(X) = -1$, so $\hat{U} - \psi$ attains in V a minimum at X , and $D\psi(X) = X$. Set

$$H_G^{\text{inv}}(Y, D\psi(Y)) = -d_{\mathcal{X}}(Y \sharp \mathbb{P}) - \sup_{\substack{v \in L_{\mathbb{P}}^2(\Omega) \\ v(\cdot) \in G(Y \sharp \mathbb{P}, Y(\cdot))}} \int_{\Omega} \langle v(\omega), D\psi(Y)(\omega) \rangle d\mathbb{P}(\omega),$$

we obtain (recalling that $\|X\|_{L_{\mathbb{P}}^2} = 1$)

$$H_G^{\text{inv}}(X, D\psi(X)) = -\frac{3}{2}\pi \int_{\Omega} \langle X, X \rangle d\mathbb{P}(\omega) = -\frac{3}{2}\pi.$$

Thus,

$$(L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_G^{\text{inv}}(X, D\psi(X)) = -\frac{3}{2}\pi < 0,$$

and therefore $\hat{U}(\cdot)$ is not a supersolution of the invariance equation.

On the other hand, set (see Definition 6.1)

$$H_G^{\text{viab}}(Y, Q) = -d_{\mathcal{X}}(Y \sharp \mathbb{P}) - \inf_{\substack{v \in L_{\mathbb{P}}^2(\Omega) \\ v(\cdot) \in G(Y \sharp \mathbb{P}, Y(\cdot))}} \int_{\Omega} \langle v(\omega), Q(\omega) \rangle d\mathbb{P}(\omega).$$

For every $v \in L_{\mathbb{P}}^2(\Omega)$ with $v(\cdot) \in F(Y \sharp \mathbb{P}, Y(\cdot)) \subseteq G(Y \sharp \mathbb{P}, Y(\cdot))$, it holds

$$H_G^{\text{viab}}(Y, Q) \geq -d_{\mathcal{X}}(Y \sharp \mathbb{P}) - \int_{\Omega} \langle v(\omega), Q(\omega) \rangle d\mathbb{P}(\omega),$$

and by taking the supremum in the right-hand side over the set

$$\{v \in L_{\mathbb{P}}^2(\Omega) : v(\cdot) \in F(Y \sharp \mathbb{P}, Y(\cdot))\},$$

we obtain

$$H_G^{\text{viab}}(Y, Q) \geq H_F^{\text{inv}}(Y, Q),$$

and therefore for every $\psi \in C^1(L_{\mathbb{P}}^2)$ such that $\hat{U} - \psi$ attains a local minimum at $X \in L_{\mathbb{P}}^2$ it holds

$$(L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_G^{\text{viab}}(X, D\psi(X)) \geq (L + 2)d_{\mathcal{X}}(X \sharp \mathbb{P}) + H_F^{\text{inv}}(X, D\psi(X)) \geq 0.$$

Thus, \mathcal{X} is viable for G , as already noticed.

Acknowledgements

G.C. thanks LMBA, Univ Brest, where this research started. G.C. is also indebted to the University of Pavia where this research has been partially carried out, in particular G.C. has been supported by Cariplo Foundation and Regione Lombardia via project *Variational Evolution Problems and Optimal Transport*, and by MIUR PRIN 2015 project *Calculus of Variations*, together with FAR funds of the Department of Mathematics of the University of Pavia. G.C. thanks also the support of the INdAM-GNAMPA Project 2019 *Optimal transport for dynamics with interaction* (“*Trasporto ottimo per dinamiche con interazione*”). M.Q. benefited from the support of the *FMJH Program Gaspard Monge in optimization and operation research* (PGMO 2016-1570H) and of the Air Force Office of Scientific Research under Award Number FA9550-18-1-0254.

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: Essential technical results

Here, we report the proofs of the preliminary results presented in Sect. 3 as well as other technical results which have been significantly used in order to prove the main propositions and theorems of the present paper. In our opinion, these results could also be interesting by themselves.

A.1: Proof of Proposition 3.4

Let $\mu = \{\mu_t\}_{t \in [a, b]}$ be an admissible trajectory defined in $[a, b]$. According to the superposition principle (Theorem 8.2.1 in [2] or Theorem 1 in [16]), there exists $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a, b]})$ such that $\mu_t = e_t \# \eta$ for all $t \in [a, b]$ and, for η -a.e. (x, γ) , $\dot{\gamma}(t) \in F(\mu_t, \gamma(t))$, $\gamma(a) = x$.

Set $|F(\mu_s, x)| = \max\{|\gamma| : \gamma \in F(\mu_s, x)\}$. For η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a,b]}$, we have

$$\begin{aligned} |\gamma(t) - \gamma(s)| &\leq \int_s^t |\dot{\gamma}(\tau)| \, d\tau \leq \int_s^t \left[|F(\mu_s, \gamma(s))| + L W_2(\mu_\tau, \mu_s) + L |\gamma(\tau) - \gamma(s)| \right] \, d\tau \\ &\leq (t-s)(K + L m_2^{1/2}(\mu_s) + L |\gamma(s)|) + L \int_s^t W_2(\mu_\tau, \mu_s) \, d\tau + L \int_s^t |\gamma(\tau) - \gamma(s)| \, d\tau. \end{aligned}$$

Grönwall's inequality yields

$$|(e_t - e_s)(x, \gamma)| = |\gamma(t) - \gamma(s)| \leq g(t, s, |\gamma(s)|),$$

where

$$g(t, s, r) := e^{L(t-s)} \left[(t-s)(K + L m_2^{1/2}(\mu_s) + Lr) + L \int_s^t W_2(\mu_\tau, \mu_s) \, d\tau \right].$$

By taking the L_η^2 -norm, $\|e_t - e_s\|_{L_\eta^2} \leq g(t, s, m_2^{1/2}(\mu_s))$, and so

$$\left\| \frac{e_t - e_s}{t-s} \right\|_{L_\eta^2} \leq \frac{g(t, s, m_2^{1/2}(\mu_s))}{t-s}.$$

By continuity, the right-hand side tends to $K + 2L m_2^{1/2}(\mu_s)$ as $t \rightarrow s^+$, and so $t \mapsto \frac{e_t - e_s}{t-s}$ is uniformly bounded in L_η^2 in a right neighborhood of s .

Let $p \in \mathbb{R}^d$, $t \in]s, b]$. For η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a,b]}$, reasoning as in the first part of the proof of Lemma 6.3, we have

$$\begin{aligned} \langle p, \frac{e_t - e_s}{t-s}(x, \gamma) \rangle &= \frac{1}{t-s} \int_s^t \langle p, \dot{\gamma}(\tau) \rangle \, d\tau \leq \frac{1}{t-s} \int_s^t \sup_{v \in F(\mu_\tau, \gamma(\tau))} \langle p, v \rangle \, d\tau \\ &\leq \sup_{v \in F(\mu_s, \gamma(s))} \langle p, v \rangle + |p| \cdot \frac{L}{t-s} \int_s^t [W_2(\mu_\tau, \mu_s) + |\gamma(\tau) - \gamma(s)|] \, d\tau. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{e_t - e_s}{t-s}(x, \gamma) &\in F(\mu_s, \gamma(s)) \\ &+ \left[\frac{L}{t-s} \int_s^t [W_2(\mu_\tau, \mu_s) + |(e_t - e_s)(x, \gamma)|] \, d\tau \right] \cdot \overline{B(0, 1)}. \end{aligned}$$

By Filippov's theorem (Theorem 8.2.10 in [4]), there exists a Borel map $w : \mathbb{R}^d \times \Gamma_{[a,b]} \rightarrow \mathbb{R}^d$, satisfying $w(x, \gamma) \in F(\mu_s, \gamma(s))$ for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a,b]}$ such that

$$\left| \frac{e_t - e_s}{t-s}(x, \gamma) - w(x, \gamma) \right| \leq \frac{L}{t-s} \int_s^t [W_2(\mu_\tau, \mu_s) + |(e_t - e_s)(x, \gamma)|] \, d\tau.$$

Thus,

$$\left\| \frac{e_t - e_s}{t-s} - w \right\|_{L_\eta^2} \leq \frac{L}{t-s} \int_s^t [W_2(\mu_\tau, \mu_s) + g(\tau, s, m_2^{1/2}(\mu_s))] \, d\tau.$$

7.1. A.2: Proof of Proposition 3.5

For a proof of the nonemptiness of the set $\mathcal{A}_{[a,b]}(\mu)$, we refer the reader to Theorem 1 in [16], where the authors perform a fixed point argument.

Let $\{\mu^{(n)}\}_{n \in \mathbb{N}} \subseteq \mathcal{A}_{[a,b]}(\mu_0)$. By Proposition 3.4(2), for any $n \in \mathbb{N}$ there exists $\eta^{(n)} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ such that $e_t \# \eta^{(n)} = \mu_t^{(n)}$ for all $t \in [a, b]$. Moreover, for $s \in [a, b]$ with $s < t$, we have

$$\begin{aligned} \|e_t - e_s\|_{L^2_{\eta^{(n)}}} &\leq e^{L(t-s)} \left[(t-s)(K + 2Lm_2^{1/2}(\mu_s^{(n)})) + \right. \\ &\quad \left. + L \int_s^t W_2(\mu_\tau^{(n)}, \mu_s^{(n)}) d\tau \right] =: h(t, s). \end{aligned}$$

Notice that $W_2(\mu_t^{(n)}, \mu_s^{(n)}) \leq \|e_t - e_s\|_{L^2_{\eta^{(n)}}}$. Indeed, it suffices to consider the admissible plan $\sigma := (e_t, e_s) \# \eta^{(n)} \in \Pi(\mu_t^{(n)}, \mu_s^{(n)})$. Thus,

$$W_2(\mu_t^{(n)}, \mu_s^{(n)}) \leq h(t, s),$$

and Grönwall's inequality yields

$$W_2(\mu_t^{(n)}, \mu_s^{(n)}) \leq e^{L(t-s)(1+e^{L(t-s)})} (t-s)(K + 2Lm_2^{1/2}(\mu_s^{(n)})). \quad (\text{A.1})$$

By taking $s = a$, there exists $\tilde{C} > 0$ such that

$$W_2(\mu_t^{(n)}, \mu_0) \leq \tilde{C}(1 + m_2^{1/2}(\mu_0)) \quad \text{for any } t \in [a, b],$$

and since $m_2(\mu_0) < +\infty$, then we obtain uniform boundedness of $\{\mu^{(n)}\}_{n \in \mathbb{N}}$.

Moreover, by the triangle inequality and by recalling that $m_2^{1/2}(\mu) = W_2(\mu, \delta_0)$ by definition, we get for any $s \in [a, b]$

$$m_2^{1/2}(\mu_s^{(n)}) \leq W_2(\mu_s^{(n)}, \mu_0) + m_2^{1/2}(\mu_0) \leq (1 + \tilde{C})(1 + m_2^{1/2}(\mu_0)). \quad (\text{A.2})$$

Thus, combining the previous estimate with (A.1), there exists $\tilde{K} > 0$ such that

$$W_2(\mu_t^{(n)}, \mu_s^{(n)}) \leq \tilde{K}(t-s),$$

and hence $\mu^{(n)}$ are continuous for any $n \in \mathbb{N}$, with uniformly bounded Lipschitz constants. By the Ascoli–Arzelà Theorem, we conclude that, up to an unlabeled subsequence, there exists $\mu = \{\mu_t\}_{t \in [a,b]} \in \text{AC}([a, b]; \mathcal{P}_2(\mathbb{R}^d))$ such that $\sup_{t \in [a,b]} W_2(\mu_t^{(n)}, \mu_t) \rightarrow 0$ as $n \rightarrow +\infty$.

We now prove the admissibility of μ . Notice that, by (3.1) and (A.2), we have

$$\sup_{n \in \mathbb{N}} \int_{\mathbb{R}^d \times \Gamma_{[a,b]}} [|\gamma| + |\gamma(a)| + \|\dot{\gamma}\|_{L^\infty([a,b])}] d\eta^{(n)}(x, \gamma) \leq C' m_2^{1/2}(\mu_0) + C'', \quad (\text{A.3})$$

for some constants $C', C'' > 0$. Moreover, the map

$$(x, \gamma) \mapsto \begin{cases} |x| + |\gamma(a)| + \|\dot{\gamma}\|_{L^\infty([a,b])}, & \text{if } \gamma \in \text{Lip}([a, b]), \\ +\infty, & \text{otherwise,} \end{cases}$$

has compact sublevels in $\mathbb{R}^d \times \Gamma_{[a,b]}$. Thus, by Remark 5.1.5 in [2], there exists $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ such that $\eta^{(n)}$ narrowly converges to η , up to (unrelabeled) subsequences. By Proposition 5.1.8 in [2], for any $(x, \gamma) \in \text{supp } \eta$ there exists $\{(x_n, \gamma_n)\}_{n \in \mathbb{N}} \subseteq \text{supp } \eta^{(n)}$ s.t. $x_n \rightarrow x, \gamma_n \rightrightarrows \gamma \in C([a, b]; \mathbb{R}^d)$.

Claim: if $(x, \gamma) \in \text{supp } \eta$, then γ is a Lipschitz continuous solution of

$$\begin{cases} \dot{\gamma}(t) \in F(\mu_t, \gamma(t)), & \text{for a.e. } t \in [a, b], \\ \gamma(a) = x. \end{cases} \tag{A.4}$$

Indeed, let $\mathcal{N} := \bigcup_{n \in \mathbb{N}} \mathcal{N}^{(n)} \subseteq [a, b]$, where

$$\mathcal{N}^{(n)} := \{s \in [a, b] : \nexists \dot{\gamma}_n(s) \text{ or } \dot{\gamma}_n(s) \notin F(\mu_s^{(n)}, \gamma_n(s))\}.$$

Notice that for any $n \in \mathbb{N}, \mathcal{N}^{(n)}$ is a negligible set w.r.t. the Lebesgue's measure, hence so is \mathcal{N} . Take $t \in [a, b] \setminus \mathcal{N}$. By Proposition 3.4, we have that $\dot{\gamma}_n(t) \in F(\mu_t^{(n)}, \gamma_n(t))$, and by assumption **(F₁)**, for any $\varepsilon > 0$ there exists \bar{n} s.t. for any $n \geq \bar{n}$

$$\dot{\gamma}_n(t) \in F(\mu_t^{(n)}, \gamma_n(t)) \subseteq F(\mu_t, \gamma(t)) + \varepsilon \overline{B(0, R)}.$$

In particular, by continuity of $t \mapsto F(\mu_t, \gamma(t))$ in $[a, b]$ we have that $\{\dot{\gamma}_n(t)\}_{n \in \mathbb{N}}$ is uniformly bounded for a.e. $t \in [a, b]$. Hence, γ_n are continuous for any $n \in \mathbb{N}$, with uniformly bounded Lipschitz constants, and $\{\gamma_n\}_{n \in \mathbb{N}}$ is uniformly bounded. By the Ascoli–Arzelà Theorem, we get that γ is a Lipschitz curve.

We now prove that γ solves (A.4). Take any $v \in \mathbb{R}^d$, and denote by $\sigma_A(v) := \sup_{z \in A} \langle v, z \rangle$ the support function of $A \subseteq \mathbb{R}^d$ at v . For any $a \leq s < t \leq b$, we have

$$\begin{aligned} \langle v, \frac{\gamma_n(t) - \gamma_n(s)}{t - s} \rangle &= \frac{1}{t - s} \int_s^t \langle v, \dot{\gamma}_n(\tau) \rangle d\tau \leq \frac{1}{t - s} \int_s^t \sigma_{F(\mu_\tau^{(n)}, \gamma_n(\tau))}(v) d\tau \\ &\leq \frac{1}{t - s} \int_s^t \left(\sigma_{F(\mu_\tau, \gamma(\tau))}(v) + L|v| \left[W_2(\mu_\tau^{(n)}, \mu_\tau) + |\gamma_n(\tau) - \gamma(\tau)| \right] \right) d\tau, \end{aligned}$$

where we used the Lipschitz continuity of F coming from **(F₂)**. By uniform in time convergence, passing to the limit as $n \rightarrow +\infty$, we have

$$\begin{aligned} \langle v, \frac{\gamma(t) - \gamma(s)}{t - s} \rangle &\leq \frac{1}{t - s} \int_s^t \sigma_{F(\mu_\tau, \gamma(\tau))}(v) d\tau \\ &\leq \frac{1}{t - s} \int_s^t \left(\sigma_{F(\mu_s, \gamma(s))}(v) + L|v| \left[W_2(\mu_\tau, \mu_s) + |\gamma(\tau) - \gamma(s)| \right] \right) d\tau. \end{aligned}$$

Thus, for a.e. s , passing to the limit as $t \rightarrow s$, we get $\langle v, \dot{\gamma}(s) \rangle \leq \sigma_{F(\mu_s, \gamma(s))}(v)$ for any $v \in \mathbb{R}^d$, whence $\dot{\gamma}(s) \in F(\mu_s, \gamma(s))$ as claimed.

Observe that, by continuity of e_t and uniqueness of the narrow limit, we have that $\mu_t^{(n)} = e_t \# \eta^{(n)}$ narrowly converges to $\mu_t = e_t \# \eta$ for any $t \in [a, b]$, up to subsequences (see Lemma 5.2.1 in [2]). The rest of the proof is an adaptation of the proof of Theorem 1 in [13]. In order to conclude the admissibility of μ , we notice that $t \mapsto \mu_t$ is a Lipschitz continuous map, indeed

$$\begin{aligned} W_2^2(\mu_t, \mu_s) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d(e_t, e_s) \# \eta \\ &= \int_{\Gamma_{[a,b]}} |\gamma(t) - \gamma(s)|^2 d\eta(x, \gamma) \leq \tilde{C} |t - s|^2, \end{aligned}$$

by Lipschitz continuity of γ in the support of η . According to Theorem 3.5 in [3], the map $t \mapsto \mu_t$ is differentiable almost everywhere in $[a, b]$, and for all $\varphi \in C_c^1(\mathbb{R}^d)$

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) d\mu_t(x) &= \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(\gamma(t)) d\eta(x, \gamma) = \\ \iint_{\mathbb{R}^d \times \Gamma_{[a,b]}} \nabla \varphi(\gamma(t)) \cdot \dot{\gamma}(t) d\eta(x, \gamma) &= \int_{\mathbb{R}^d} \nabla \varphi(y) \int_{e_t^{-1}(y)} \dot{\gamma}(t) d\eta_{t,y}(x, \gamma) d\mu_t(y), \end{aligned}$$

where $\{\eta_{t,y}\}_{y \in \mathbb{R}^d} \subseteq \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ is the disintegration of η w.r.t. the evaluation operator e_t , i.e., $\eta = \mu_t \otimes \eta_{t,y}$. Finally, notice that the vector field

$$v_t(y) := \int_{e_t^{-1}(y)} \dot{\gamma}(t) d\eta_{t,y}(x, \gamma)$$

is well-defined for a.e. $t \in [a, b]$ and μ_t -a.e. $y \in \mathbb{R}^d$, moreover, by convexity of $F(\mu_t, y)$, we can use Jensen's inequality to get that $v_t(y) \in F(\mu_t, y)$ for a.e. t and μ_t -a.e. y . Hence the conclusion.

7.2. A.3: Technical results

Corollary A.1. *Assume $(F_1) - (F_2)$. Let $\mu = \{\mu_t\}_{t \in [a,b]}$ be an admissible trajectory, with $0 \leq a < b < +\infty$. Then, there exists a family of random variables $\{X_t\}_{t \in [a,b]} \subseteq L_{\mathbb{P}}^2(\Omega)$ such that $X_t \# \mathbb{P} = \mu_t$ for all $t \in [a, b]$, and*

$$\begin{aligned} \frac{X_t - X_s}{t - s}(\omega) &\in F(X_s \# \mathbb{P}, X_s(\omega)) + \left[\frac{L}{t - s} \int_s^t [W_2(\mu_\tau, \mu_s) + |X_\tau(\omega) - X_s(\omega)|] d\tau \right] \cdot \overline{B(0, 1)}, \\ \|X_t - X_s\|_{L_{\mathbb{P}}^2} &\leq e^{L(t-s)} \left[(t - s)(K + 2Lm_2^{1/2}(\mu_s)) + L \int_s^t W_2(\mu_\tau, \mu_s) d\tau \right], \end{aligned}$$

for all $t, s \in [a, b]$, with $s < t$.

In particular, for every $p(\cdot) \in L_{\mathbb{P}}^2(\Omega)$ we have

$$\liminf_{t \rightarrow s^+} \langle p, \frac{X_t - X_s}{t - s} \rangle_{L_{\mathbb{P}}^2} \geq \int_{\Omega} \inf_{v \in F(X_s \# \mathbb{P}, X_s(\omega))} \langle p(\omega), v \rangle d\mathbb{P}(\omega), \quad (\text{A.5})$$

$$\limsup_{t \rightarrow s^+} \langle p, \frac{X_t - X_s}{t - s} \rangle_{L_{\mathbb{P}}^2} \leq \int_{\Omega} \sup_{v \in F(X_s \# \mathbb{P}, X_s(\omega))} \langle p(\omega), v \rangle d\mathbb{P}(\omega). \quad (\text{A.6})$$

Proof. Let $\mu = \{\mu_t\}_{t \in [a,b]}$ be an admissible trajectory defined in $[a, b]$, and $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ be as in Proposition 3.4, with $\mu_t = e_t \# \eta$ for $t \in [a, b]$. In particular, see, e.g., Lemma 5.29 in [11], there exists a Borel map $\mathcal{V} : \Omega \rightarrow \mathbb{R}^d \times \Gamma_{[a,b]}$ such that $\eta = \mathcal{V} \# \mathbb{P}$, and thus $\mu_t = X_t \# \mathbb{P}$ for all $t \in [a, b]$, where $X_t = e_t \circ \mathcal{V}$. Evaluating the estimates obtained in Proposition 3.4 for $(x, \gamma) = \mathcal{V}(\omega)$, and recalling that $X_t = e_t \circ \mathcal{V}$, $X_s = e_s \circ \mathcal{V}$, we obtain

$$\begin{aligned} \frac{X_t - X_s}{t - s}(\omega) &\in F(X_s \# \mathbb{P}, X_s(\omega)) + \\ &+ \left[\frac{L}{t - s} \int_s^t [W_2(\mu_\tau, \mu_s) + |X_\tau(\omega) - X_s(\omega)|] \, d\tau \right] \cdot \overline{B(0, 1)}, \\ \|X_t - X_s\|_{L^2_{\mathbb{P}}} &\leq e^{L(t-s)} \left[(t - s)(K + 2Lm_2^{1/2}(\mu_s)) + L \int_s^t W_2(\mu_\tau, \mu_s) \, d\tau \right]. \end{aligned}$$

Thus, for every $p(\cdot) \in L^2_{\mathbb{P}}(\Omega)$, we have

$$\begin{aligned} \left\langle p, \frac{X_t - X_s}{t - s} \right\rangle_{L^2_{\mathbb{P}}} &= \int_{\Omega} \left\langle p(\omega), \frac{X_t - X_s}{t - s}(\omega) \right\rangle \, d\mathbb{P}(\omega) \\ &\geq \int_{\Omega} \inf_{v \in F(X_s \# \mathbb{P}, X_s(\omega))} \langle p(\omega), v \rangle \, d\mathbb{P}(\omega) + \\ &\quad - \frac{L}{t - s} \int_{\Omega} |p(\omega)| \int_s^t [W_2(\mu_\tau, \mu_s) + |X_\tau(\omega) - X_s(\omega)|] \, d\tau \, d\mathbb{P}(\omega) \quad (\text{A.7}) \\ &\geq \int_{\Omega} \inf_{v \in F(X_s \# \mathbb{P}, X_s(\omega))} \langle p(\omega), v \rangle \, d\mathbb{P}(\omega) + \\ &\quad - \|p\|_{L^2_{\mathbb{P}}} \cdot \frac{L}{t - s} \int_s^t [W_2(\mu_\tau, \mu_s) + \|X_\tau - X_s\|_{L^2_{\mathbb{P}}}] \, d\tau. \end{aligned}$$

By taking the liminf as $t \rightarrow s^+$, and using the estimate on $\|X_\tau - X_s\|_{L^2_{\mathbb{P}}}$, we obtain

$$\liminf_{t \rightarrow s^+} \left\langle p, \frac{X_t - X_s}{t - s} \right\rangle_{L^2_{\mathbb{P}}} \geq \int_{\Omega} \inf_{v \in F(X_s \# \mathbb{P}, X_s(\omega))} \langle p(\omega), v \rangle \, d\mathbb{P}(\omega).$$

In the same way, we prove the inequality for the limsup. □

We recall the following well-known result, used to prove Corollary A.3.

Lemma A.2. (Lemma 5.23 p. 379 in [11]) *Let \mathbb{P} be an atomless Borel probability measure on Ω , $X, Y \in L^2_{\mathbb{P}}(\Omega; \mathbb{R}^d)$ two random variables with the same law, i.e., $X \# \mathbb{P} = Y \# \mathbb{P}$. Then for any $\varepsilon > 0$, there exist two Borel measurable maps $r, r^{-1} : \Omega \rightarrow \Omega$ such that*

- r and r^{-1} are measure-preserving, i.e., $r \# \mathbb{P} = r^{-1} \# \mathbb{P} = \mathbb{P}$;
- $\mathbb{P}(\{\omega \in \Omega : r \circ r^{-1}(\omega) = r^{-1} \circ r(\omega) = \omega\}) = 1$;
- $\mathbb{P}(\{\omega \in \Omega : |X(\omega) - Y \circ r(\omega)| \leq \varepsilon\}) = 1$.

In particular, we have $\|X - Y \circ r\|_{L^2_{\mathbb{P}}} \leq \varepsilon$.

Corollary A.3. *Assume (F₁) – (F₂). Let $\mu = \{\mu_t\}_{t \in [a,b]}$ be an admissible trajectory, and $X \in L^2_{\mathbb{P}}(\Omega)$ such that $X \# \mathbb{P} = \mu_a$. Let $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ such that $\mu_t = e_t \# \eta$ for any $t \in [a, b]$. Then, for every $\varepsilon > 0$ there exists a family of random variables $\{Y_t\}_{t \in [a,b]} \subseteq L^2_{\mathbb{P}}(\Omega)$ such that*

- (1) $Y_t \# \mathbb{P} = \mu_t$ for all $t \in [a, b]$;
- (2) $\|Y_a - X\|_{L^2_{\mathbb{P}}} \leq \varepsilon$, and so it holds $\|Y_t - X\|_{L^2_{\mathbb{P}}} \leq \|Y_t - Y_a\|_{L^2_{\mathbb{P}}} + \varepsilon$;
- (3) for every $t \in [a, b]$ and for every $p(\cdot) \in L^2_{\mathbb{P}}(\Omega)$ we have

$$\begin{aligned} \langle p, \frac{Y_t - Y_a}{t - a} \rangle_{L^2_{\mathbb{P}}} &\geq \int_{\Omega} \inf_{v \in F(X \# \mathbb{P}, X(\omega))} \langle p(\omega), v \rangle d\mathbb{P}(\omega) - (\widehat{\varpi}(t) + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}}, \\ \langle p, \frac{Y_t - Y_a}{t - a} \rangle_{L^2_{\mathbb{P}}} &\leq \int_{\Omega} \sup_{v \in F(X \# \mathbb{P}, X(\omega))} \langle p(\omega), v \rangle d\mathbb{P}(\omega) + (\widehat{\varpi}(t) + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}}, \end{aligned} \quad (\text{A.8})$$

where

$$\widehat{\varpi}(t) := \frac{L}{t - a} \int_a^t \left[W_2(\mu_\tau, \mu_a) + \|e_\tau - e_a\|_{L^2_{\eta}} \right] d\tau.$$

Proof. Fix $\varepsilon > 0$. Let $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a,b]})$ represent μ , i.e., $\mu_t = e_t \# \eta$. Since \mathbb{P} is an atomless Borel probability measure on a Polish space, as already noticed, there exists a Borel map $\mathcal{V} : \Omega \rightarrow \mathbb{R}^d \times \Gamma_{[a,b]}$ such that $\eta = \mathcal{V} \# \mathbb{P}$. Set $X_t = e_t \circ \mathcal{V}$ for all $t \in [a, b]$. Notice that for every measure-preserving map $r : \Omega \rightarrow \Omega$, we have $\eta = (\mathcal{V} \circ r) \# \mathbb{P}$, since $r \# \mathbb{P} = \mathbb{P}$. Moreover, $(X_t \circ r) \# \mathbb{P} = X_t \# \mathbb{P} = \mu_t$ for all $t \in [a, b]$. By Proposition 3.4, for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a,b]}$ it holds

$$\frac{e_t - e_a}{t - a}(x, \gamma) \in F(\mu_a, e_a(x, \gamma)) + \varpi(t, x, \gamma) \cdot \overline{B(0, 1)}.$$

where

$$\varpi(t, x, \gamma) := \frac{L}{t - a} \int_a^t \left[W_2(\mu_\tau, \mu_a) + |(e_\tau - e_a)(x, \gamma)| \right] d\tau,$$

and so $\varpi(t, x, \gamma) \rightarrow 0^+$ as $t \rightarrow a^+$.

Evaluating at $(x, \gamma) = \mathcal{V} \circ r(\omega)$, and recalling that $X_t = e_t \circ \mathcal{V}$, we obtain

$$\begin{aligned} \left(\frac{X_t - X_a}{t - a} \right) \circ r(\omega) &\in F(\mu_a, X_a \circ r(\omega)) + \varpi(t, \mathcal{V} \circ r(\omega)) \overline{B(0, 1)} \\ &\subseteq F(\mu_a, X(\omega)) + [\varpi(t, \mathcal{V} \circ r(\omega)) \\ &\quad + L|X(\omega) - X_a \circ r(\omega)|] \cdot \overline{B(0, 1)}. \end{aligned}$$

Since $X \# \mathbb{P} = X_a \# \mathbb{P} = \mu_a$, by Lemma A.2 for any $\varepsilon > 0$ we can choose a measure-preserving map $r = r_\varepsilon$ such that $|X(\omega) - X_a \circ r(\omega)| \leq \varepsilon$ for \mathbb{P} -a.e. $\omega \in \Omega$. So we have

$$\left(\frac{X_t - X_a}{t - a} \right) \circ r_\varepsilon(\omega) \in F(X \# \mathbb{P}, X(\omega)) + (\varpi(t, \mathcal{V} \circ r_\varepsilon(\omega)) + L\varepsilon) \cdot \overline{B(0, 1)}.$$

Let $Y_t = X_t \circ r_\varepsilon$ for all $t \in [a, b]$. Then, as seen in the proof of Corollary A.1, we have

$$\begin{aligned} \langle p, \frac{Y_t - Y_a}{t - a} \rangle_{L^2_{\mathbb{P}}} &\geq \int_{\Omega} \inf_{v \in F(X_{\sharp\mathbb{P}}, X(\omega))} \langle v, p(\omega) \rangle d\mathbb{P}(\omega) - (\|\varpi(t, \mathcal{V} \circ r_\varepsilon(\cdot))\|_{L^2_{\mathbb{P}}} + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}}, \\ \langle p, \frac{Y_t - Y_a}{t - a} \rangle_{L^2_{\mathbb{P}}} &\leq \int_{\Omega} \sup_{v \in F(X_{\sharp\mathbb{P}}, X(\omega))} \langle v, p(\omega) \rangle d\mathbb{P}(\omega) + (\|\varpi(t, \mathcal{V} \circ r_\varepsilon(\cdot))\|_{L^2_{\mathbb{P}}} + L\varepsilon) \|p\|_{L^2_{\mathbb{P}}}, \end{aligned}$$

for every $p \in L^2_{\mathbb{P}}(\Omega)$. Notice that

- (1) $Y_t \sharp\mathbb{P} = X_t \circ r_\varepsilon \sharp\mathbb{P} = X_t \sharp\mathbb{P} = \mu_t$ since $r_\varepsilon \sharp\mathbb{P} = \mathbb{P}$;
- (2) we have

$$\|Y_t - X\|_{L^2_{\mathbb{P}}} \leq \|Y_t - Y_a\|_{L^2_{\mathbb{P}}} + \|Y_a - X\|_{L^2_{\mathbb{P}}} \leq \|Y_t - Y_a\|_{L^2_{\mathbb{P}}} + \varepsilon;$$

- (3) it holds

$$\|\varpi(t, \mathcal{V} \circ r(\cdot))\|_{L^2_{\mathbb{P}}} \leq \widehat{\varpi}(t) = \frac{L}{t - a} \int_a^t \left[W_2(\mu_\tau, \mu_a) + \|e_\tau - e_a\|_{L^2_{\mathbb{H}}} \right] d\tau.$$

From here follows the conclusion. \square

Lemma A.4. Assume $(F_1) - (F_2)$. Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $a \geq 0$ be fixed, and consider a continuous selection $v(\cdot)$ of $F(\mu, \cdot)$. Then, there exist $T > a$ and $\widehat{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_{[a, T]})$ such that, if we set $\widehat{\theta}_t = e_t \sharp \widehat{\eta}$ for all $t \in [a, T]$ and $\widehat{\theta} = \{\widehat{\theta}_t\}_{t \in [a, T]}$,

- a. $\widehat{\theta}$ is an admissible trajectory with $\widehat{\theta}_a = \mu$ defined on $[a, T]$;
- b. for $\widehat{\eta}$ -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_{[a, T]}$ we have $\gamma \in C^1([a, T])$ with $\dot{\gamma}(a) = v(x)$, and

$$\begin{cases} \dot{\gamma}(t) \in F(\widehat{\theta}_t, \gamma(t)), & \text{for a.e. } t \in [a, T], \\ \gamma(a) = x; \end{cases}$$

- c. $\frac{e_t - e_a}{t - a} \rightarrow v \circ e_a$ in $L^2_{\widehat{\eta}}$ as $t \rightarrow a^+$.

Proof. Without loss of generality, we set $a = 0$. According to Theorem 9.7.2 in [4], there exists a continuous map $f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \overline{B(0, 1)} \rightarrow \mathbb{R}^d$ and a constant c independent on F such that for all $(\theta, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ it holds

- $F(\theta, x) = \{f(\theta, x, u) : u \in \overline{B(0, 1)}\}$.
- for every $u \in \overline{B(0, 1)}$, the map $(\theta, x) \mapsto f(\theta, x, u)$ is Lipschitz continuous with Lipschitz constant less than $c \cdot \text{Lip}(F)$,
- $|f(\theta, x, u) - f(\theta, x, v)| \leq c \cdot |F(\theta, x)| \cdot |u - v|$.

In particular, for all $(\theta, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ we have

$$|f(\theta, x, u) - f(\theta, x, v)| \leq c \cdot \left[K + \text{Lip}(F) \cdot \left(m_2^{1/2}(\theta) + |x| \right) \right] \cdot |u - v|.$$

Let

$$L = \text{Lip}(F), \quad K = |F(\delta_0, 0)|, \quad L' > 2 \left[2cL(1 + cLT e^{cLT}) m_2^{1/2}(\mu) + cLT K e^{cLT} + K \right],$$

and let $T > 0$ that will be fixed later. Given $(x, u) \in \mathbb{R}^d \times \overline{B(0, 1)}$ and a Lipschitz curve $\theta = \{\theta_t\}_{t \in [0, T]} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ with $\theta_0 = \mu$ and $\text{Lip}(\theta) < L'$, denote by $\gamma_{\theta, x, u}(\cdot)$ the unique solution of

$$\dot{\gamma}(t) = f(\theta_t, \gamma(t), u), \quad \gamma(0) = x.$$

For any $(\hat{\mu}, q, r) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R} \times \mathbb{R}$ define

$$\begin{aligned} C_1(\hat{\mu}, q, r) &:= \frac{cLL'}{2}r^2 + cLr\mathfrak{m}_2^{1/2}(\hat{\mu}) + cLrq + rK, \\ C_2(\hat{\mu}, q, r) &:= K + L \left(L'r + \mathfrak{m}_2^{1/2}(\hat{\mu}) + q + e^{cLr} C_1(\hat{\mu}, q, r) \right), \\ C_3(\hat{\mu}, q, r) &:= e^{cLr} (1 + crC_2(\hat{\mu}, q, r)). \end{aligned}$$

For all $0 \leq s \leq t \leq T$, $y \in \mathbb{R}^d$, we have

$$\begin{aligned} & \int_s^t |f(\theta_\tau, y, u)| \, d\tau \leq \\ & \leq \int_s^t |f(\theta_\tau, y, u) - f(\theta_s, y, u)| \, d\tau + \int_s^t |f(\theta_s, y, u) - f(\delta_0, 0, u)| \, d\tau + (t-s)|f(\delta_0, 0, u)| \\ & \leq cL \int_s^t W_2(\theta_\tau, \theta_s) \, d\tau + cL(t-s)\mathfrak{m}_2^{1/2}(\theta_s) + cL(t-s)|y| + (t-s)K \leq C_1(\theta_s, |y|, t-s), \end{aligned}$$

where we used Lipschitz-in-time continuity of θ . Since

$$\begin{aligned} |\gamma_{\theta, x, u}(t) - \gamma_{\theta, x, u}(s)| & \leq \int_s^t |f(\theta_\tau, \gamma_{\theta, x, u}(\tau), u) - f(\theta_\tau, \gamma_{\theta, x, u}(s), u)| \, d\tau + \\ & + \int_s^t |f(\theta_\tau, \gamma_{\theta, x, u}(s), u)| \, d\tau \leq cL \int_s^t |\gamma_{\theta, x, u}(\tau) - \gamma_{\theta, x, u}(s)| \, d\tau + C_1(\theta_s, |\gamma_{\theta, x, u}(s)|, t-s), \end{aligned}$$

by Grönwall's inequality,

$$|\gamma_{\theta, x, u}(t) - \gamma_{\theta, x, u}(s)| \leq e^{cL(t-s)} C_1(\theta_s, |\gamma_{\theta, x, u}(s)|, t-s).$$

Choosing $s = 0$, for all $t \in [0, T]$, we have

$$|\gamma_{\theta, x, u}(t)| \leq |x| + e^{cLT} C_1(\mu, |x|, T). \quad (\text{A.9})$$

This implies also

$$\begin{aligned} |F(\theta_t, \gamma_{\theta, x, u}(t))| & \leq K + L (W_2(\theta_t, \delta_0) + |\gamma_{\theta, x, u}(t)|) \\ & \leq K + L(W_2(\theta_t, \mu) + \mathfrak{m}_2^{1/2}(\mu) + |\gamma_{\theta, x, u}(t)|) \leq C_2(\mu, |x|, T). \end{aligned}$$

Notice that the map $(x, u) \mapsto (x, \gamma_{\theta, x, u})$ is locally Lipschitz continuous. Indeed,

$$\begin{aligned}
 & |\gamma_{\theta, x_1, u}(t) - \gamma_{\theta, x_2, v}(t)| \leq \\
 & \leq |x_1 - x_2| + \int_0^t |f(\theta_\tau, \gamma_{\theta, x_1, u}(\tau), u) - f(\theta_\tau, \gamma_{\theta, x_2, v}(\tau), v)| \, d\tau \\
 & \leq |x_1 - x_2| + \int_0^t |f(\theta_\tau, \gamma_{\theta, x_1, u}(\tau), u) - f(\theta_\tau, \gamma_{\theta, x_1, u}(\tau), v)| \, d\tau + \\
 & \quad + \int_0^t |f(\theta_\tau, \gamma_{\theta, x_1, u}(\tau), v) - f(\theta_\tau, \gamma_{\theta, x_2, v}(\tau), v)| \, d\tau \\
 & \leq |x_1 - x_2| + c \int_0^t |F(\theta_\tau, \gamma_{\theta, x_1, u}(\tau))| \, d\tau \cdot |u - v| + \\
 & \quad + cL \int_0^t |\gamma_{\theta, x_1, u}(\tau) - \gamma_{\theta, x_2, v}(\tau)| \, d\tau \\
 & \leq |x_1 - x_2| + ctC_2(\mu, |x_1|, T) \cdot |u - v| + cL \int_0^t |\gamma_{\theta, x_1, u}(\tau) - \gamma_{\theta, x_2, v}(\tau)| \, d\tau.
 \end{aligned}$$

By Grönwall's inequality, for all $t \in [0, T]$ we have

$$\begin{aligned}
 |\gamma_{\theta, x_1, u}(t) - \gamma_{\theta, x_2, v}(t)| & \leq e^{cLT} (1 + cTC_2(\mu, |x_1|, T)) (|x_1 - x_2| + |u - v|) \\
 & = C_3(\mu, |x_1|, T) (|x_1 - x_2| + |u - v|).
 \end{aligned}$$

This provides the Lipschitz continuity on all bounded subsets of $\mathbb{R}^d \times \overline{B(0, 1)}$ by the continuity of $C_3(\mu, \cdot, T)$.

By Filippov's Theorem (see Theorem 8.2.10 in [4]), there exists a Borel map $u : \mathbb{R}^d \rightarrow \overline{B(0, 1)}$ such that $v(x) = f(\mu, x, u(x))$ for all $x \in \mathbb{R}^d$. The map $x \mapsto \gamma_{\theta, x, u(x)}$ is a composition of Borel maps, so it is Borel, and we define $\eta^\theta = \mu \otimes \delta_{\gamma_{\theta, x, u(x)}}$. We have by construction that

- $e_0 \# \eta^\theta = \mu$;
- for η^θ -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$ we have $\gamma \in C^1([0, T])$ with $\gamma(0) = x$ and $\dot{\gamma}(0) = v(x)$.

We want to show now that $t \mapsto e_t \# \eta^\theta$ is Lipschitz continuous with constant less than L' . Indeed, given $0 \leq s \leq t \leq T$ we have

$$\begin{aligned}
 W_2(e_t \# \eta^\theta, e_s \# \eta^\theta) & \leq \|e_t - e_s\|_{L^2_{\eta^\theta}} = \left(\int_{\mathbb{R}^d \times \Gamma_T} |\gamma(t) - \gamma(s)|^2 \, d\eta^\theta(x, \gamma) \right)^{1/2} \\
 & \leq e^{cLT} \left(\int_{\mathbb{R}^d \times \Gamma_T} |C_1(\theta_s, |\gamma(s)|, t - s)|^2 \, d\eta^\theta(x, \gamma) \right)^{1/2} \\
 & \leq \frac{cLL'}{2} (t - s)^2 + cL(t - s)m_2^{1/2}(\theta_s) + cL(t - s) \left(\int_{\mathbb{R}^d \times \Gamma_T} |\gamma(s)|^2 \, d\eta^\theta(x, \gamma) \right)^{1/2} + (t - s)K.
 \end{aligned}$$

Notice that

$$m_2^{1/2}(\theta_s) = W_2(\theta_s, \delta_0) \leq W_2(\theta_s, \theta_0) + W_2(\theta_0, \delta_0) \leq L'T + m_2^{1/2}(\mu).$$

Moreover, by (A.9), we have

$$\begin{aligned}
 & \left(\int_{\mathbb{R}^d \times \Gamma_T} |\gamma(s)|^2 d\eta^\theta(x, \gamma) \right)^{1/2} \leq m_2^{1/2}(\mu) + e^{cLT} \left(\int_{\mathbb{R}^d \times \Gamma_T} |C_1(\mu, |x|, T)|^2 d\eta^\theta(x, \gamma) \right)^{1/2} \\
 & \leq m_2^{1/2}(\mu) + e^{cLT} \left[\frac{cLL'}{2} T^2 + 2cLT m_2^{1/2}(\mu) + TK \right] \\
 & \leq \frac{cLL'}{2} T^2 e^{cLT} + (1 + 2cLT e^{cLT}) m_2^{1/2}(\mu) + TK e^{cLT} =: \frac{cLL'}{2} T^2 e^{cLT} + B(\mu, T).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 & W_2(e_t \# \eta^\theta, e_s \# \eta^\theta) \\
 & \leq |t - s| \left(\frac{cLL'}{2} T + cL(L'T + m_2^{1/2}(\mu)) + cL \left(\frac{cLL'}{2} T^2 e^{cLT} + B(\mu, T) \right) + K \right) \\
 & \leq |t - s| \left[L' \left(\frac{3}{2} cLT + \frac{c^2 L^2}{2} T^2 e^{cLT} \right) + 2cL(1 + cLT e^{cLT}) m_2^{1/2}(\mu) + cLTK e^{cLT} + K \right] \\
 & =: S(\mu, L', T) |t - s|.
 \end{aligned} \tag{A.10}$$

In particular, since we choose

$$\frac{L'}{2} > 2cL(1 + cLT e^{cLT}) m_2^{1/2}(\mu) + cLTK e^{cLT} + K,$$

there is $T > 0$ such that $S(\mu, L', T) < L'$, and so $t \mapsto e_t \# \eta^\theta$ is Lipschitz continuous with constant less than L' .

Define by recurrence a sequence of curves $\{\theta^{(n)} = \{\theta_t^{(n)}\}_{t \in [0, T]}\}_{n \in \mathbb{N}}$ and of measures $\{\eta^{(n)}\}_{n \in \mathbb{N}}$ by setting $\theta_t^{(0)} = \mu$ for all $t \in [0, T]$. Supposing that we have defined $\theta^{(n)}$, then we define $\eta^{(n)} = \mu \otimes \delta_{\gamma_{\theta^{(n)}, x, u(x)}}$ and $\theta^{(n+1)}$ by setting $\theta_t^{(n+1)} = e_t \# \eta^{(n)}$. Notice that, by construction, for all $n \in \mathbb{N}$, $\theta_0^{(n)} = \mu$ and $\theta^{(n)}$ is a Lipschitz continuous curve with $\text{Lip}(\theta^{(n)}) < L'$, by (A.10).

Since we have the same estimate as in (A.3), then there exists $\eta \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ and a subsequence $\eta^{(n_k)}$ such that $\eta^{(n_k)}$ narrowly converges toward η . As already observed, we also have that $\theta^{(n_k)}$ is a family of uniformly bounded and continuous curves, with uniformly bounded Lipschitz constants. Thus, it has a subsequence which is uniformly convergent to a Lipschitz curve $\theta = \{\theta_t\}_{t \in [0, T]}$. We now follow the same reasoning as in the last part of the proof of Proposition 3.5 with μ and $\mu^{(n)}$ replaced, respectively, by θ and $\theta^{(n_k)}$. For η -a.e. (x, γ) , we get that $\dot{\gamma}(t) \in F(\theta_t, \gamma(t))$ for a.e. $t \in [0, T]$, $\gamma(0) = x$, $\gamma \in C^1([0, T])$ and $\dot{\gamma}(0) = v(x)$. Thus, θ is an admissible trajectory and $\frac{e_t - e_0}{t}(x, \gamma) \rightarrow v(x)$, as $t \rightarrow 0^+$, for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$. We also notice that $v(x) = v \circ e_0(x, \gamma)$ for η -a.e. $(x, \gamma) \in \mathbb{R}^d \times \Gamma_T$. Finally, recalling the estimates on the admissible trajectories provided in Proposition 3.4, by the Dominated Convergence Theorem we have that the convergence is actually in L^2_η . \square

REFERENCES

- [1] G. Albi, N. Bellomo, L. Fermo, S.-Y. Ha, J. Kim, L. Pareschi, D. Poyato, and J. Soler, Vehicular traffic, crowds, and swarms: From kinetic theory and multiscale methods to applications and research perspectives, *Mathematical Models and Methods in Applied Sciences* **29** (2019), no. 10, 1901–2005, <https://doi.org/10.1142/S0218202519500374>.
- [2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
- [3] Luigi Ambrosio and Bernd Kirchheim, Rectifiable sets in metric and Banach spaces, *Math. Ann.* **318** (2000), no. 3, 527–555, <https://doi.org/10.1007/s002080000122>.
- [4] Jean-Pierre Aubin and Hélène Frankowska, *Set-valued analysis*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2009. Reprint of the 1990 edition.
- [5] Yurii Averboukh, *Viability theorem for deterministic mean field type control systems*, *Set-Valued Var. Anal.* **26** (2018), no. 4, 993–1008.
- [6] Martino Bardi and Italo Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, *Systems & Control: Foundations & Applications*, Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [7] Alain Bensoussan, Jens Frehse, and Phillip Yam, *Mean field games and mean field type control theory*, SpringerBriefs in Mathematics, Springer, New York, 2013.
- [8] Jonathan M. Borwein and Qiji J. Zhu, *Techniques of variational analysis*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, vol. 20, Springer-Verlag, New York, 2005.
- [9] Rainer Buckdahn, Shige Peng, Marc Quincampoix, and Catherine Rainer, Existence of stochastic control under state constraints, *C. R. Acad. Sci. Paris Sér. I Math.* **327** (1998), no. 1, 17–22.
- [10] Pierre Cardaliaguet, *Notes on Mean Field Games*, 2013. From P.-L. Lions' lectures at College de France, link to the notes.
- [11] René Carmona and François Delarue, *Probabilistic Theory of Mean Field Games with Applications, vol I*, *Probability Theory and Stochastic Modeling*, vol. 84, Springer International Publishing, 2018.
- [12] G. Cagnani, S. Lisini, C. Orrieri, and G. Savaré, *Lagrangian, Eulerian and Kantorovich formulations of multi-agent optimal control problems: equivalence and Gamma-convergence*, Preprint: [arXiv:2011.07117](https://arxiv.org/abs/2011.07117) [math.OC].
- [13] Giulia Cagnani, Antonio Marigonda, and Benedetto Piccoli, *Superposition principle for differential inclusions*, *Large-scale scientific computing*, *Lecture Notes in Comput. Sci.*, vol. 10665, Springer, Cham, 2018, pp. 201–209.
- [14] Francis Clarke, *Functional analysis, calculus of variations and optimal control*, *Graduate Texts in Mathematics*, vol. 264, Springer, London, 2013.
- [15] M. Fornasier, S. Lisini, C. Orrieri, and G. Savaré, *Mean-field optimal control as gammadimit of finite agent controls*, *European J. Appl. Math.* **30** (2019), no. 6, 1153–1186, <https://doi.org/10.1017/s0956792519000044>.
- [16] Chloé Jimenez, Antonio Marigonda, and Marc Quincampoix, *Optimal Control of Multiagent Systems in the Wasserstein Space*, *Calculus of Variations and Partial Differential Equations* **59** (2020), <https://doi.org/10.1007/s00526-020-1718-6>.
- [17] Alexander Kruger, *On Fréchet Subdifferentials*, *Journal of Mathematical Sciences* **116** (2003), no. 3, 3325–3358, <https://doi.org/10.1023/A:1023673105317>.
- [18] Antonio Marigonda and Marc Quincampoix, *Mayer control problem with probabilistic uncertainty on initial positions*, *J. Differential Equations* **264** (2018), no. 5, 3212–3252.
- [19] Nikolay Pogodaev, *Optimal control of continuity equations*, *NoDEA Nonlinear Differential Equations Appl.* **23** (2016), no. 2, Art. 21, 24, <https://doi.org/10.1007/s00030-016-0357-2>.

Compatibility of state constraints

Giulia Cavagnari
Dipartimento di Matematica “F.
Brioschi”
Politecnico di Milano
Piazza Leonardo da Vinci 32
20133 Milan
Italy
E-mail: giulia.cavagnari@polimi.it

Antonio Marigonda
Department of Computer Science
University of Verona
Strada Le Grazie 15
37134 Verona
Italy
E-mail: antonio.marigonda@univr.it

Marc Quincampoix
Laboratoire de Mathématiques
de Bretagne Atlantique, UMR CNRS 6205
Univ Brest
6, Avenue Victor Le Gorgeu
29200 Brest
France
E-mail: marc.quincampoix@univ-brest.fr

Accepted: 26 May 2021