ATTAINABILITY PROPERTY FOR A PROBABILISTIC TARGET IN WASSERSTEIN SPACES

Giulia Cavagnari*

Politecnico di Milano
Dipartimento di Matematica "F. Brioschi", Piazza Leonardo da Vinci 32
I-20133 Milano, Italy

Antonio Marigonda

University of Verona
Department of Computer Science, Strada Le Grazie 15
I-37134 Verona, Italy

Abstract. In this paper we establish an attainability result for the minimum time function of a control problem in the space of probability measures endowed with Wasserstein distance. The dynamics is provided by a suitable controlled continuity equation, where we impose a nonlocal nonholonomic constraint on the driving vector field, which is assumed to be a Borel selection of a given set-valued map. This model can be used to describe at a macroscopic level a so-called *multiagent system* made of several possible interacting agents.

1. **Introduction.** We consider a finite-dimensional *multiagent system*, i.e., a system in $\mathbb{R}^d$ where the number of agents is so large that only a *macroscopic* description is available. As usual in this framework, in order to describe the behaviour of the system at a certain time t, we introduce a Borel positive measure μ_t on $\mathbb{R}^d$ whose meaning is the following: given a Borel set $A \subseteq \mathbb{R}^d$ the quantity $\dfrac{\mu_t(A)}{\mu_t(\mathbb{R}^d)}$ represents the fraction of the total number of agents that are present in A at the time t. We will assume that the system is *isolated*, thus the total number of agents remains constant in time. Hence, by normalizing the measure μ_t, we can always assume $\mu_t(\mathbb{R}^d) = 1$, i.e., μ_t is a probability measure for all t.

The macroscopic evolution of the system is thus given by a curve $t \mapsto \mu_t$ in the space of probability measures. Due to the mass-preserving character of the evolution, we can assume that such an evolution is governed by the *continuity equation*

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0,$$

to be satisfied in a distributional sense, where v_t is a suitable time-depending Borel vector field describing the macroscopic mass flux during the evolution.

2020 *Mathematics Subject Classification.* Primary: 34A60; Secondary: 49J15, 49J53, 90C56.
Key words and phrases. optimal transport, differential inclusions, time optimal control, small time local attainability, multi-agent system.
* Corresponding author: Giulia Cavagnari.

It can be easily proved, see e.g. [14], that for a.e. t and μ_t-a.e. $x \in \mathbb{R}^d$ the vector field $v_t(x)$ can be constructed as a weigthed average of the velocities of the agents passing through the point x at time t, where the weights are given by the fraction of the mass carried by each agent w.r.t. the total amount of mass flowing through x at time t. In particular, possibly nonlocal nonholonomic constraints on the agents' motion will reflect into constraints for the possible choices of v_t.

In this paper we consider a situation where each agent is constrained to follow the trajectories of a differential inclusion with a *nonlocal* dependence on the overall configuration of the agents. This fact models the possible *nonlocal interaction* among the agents. Examples of such interactions are quite commmon in the models of pedestrian dynamics, flocks of animals and social dynamics in general.

Due to the potential applications, the literature on control of multi-agent systems is growing quite fast in the recent years. Among the most recent contributions, we mention [8], where the authors investigate a controllability problem for a leader-follower model in a finite-dimensional setting and their aim is to achieve an alignment consensus for a mass of indistinguishable agents when the action of an external policy maker is *sparse*, i.e. concentrated on few individuals. In [23] it is provided a mean-field formulation of the same model through Gamma-convergence techniques.

The relevance of such kind of results is enhanced when dealing with problems involving a considerable number of individuals, in order to circumvent the bounds coming from the curse of dimensionality: indeed, the mean-field limit can be used as a realistic approximation when the number of agents is huge. Results in this direction are provided for example by [22] or the preprint paper [12], where the authors study a Gamma-convergence result for an optimal control problem of a N-particles system subject to a nonlocal dynamics when $N \to +\infty$.

Controllability conditions in the space of probability measures are also analyzed in the preprints [17], [18]. In particular, the aim of the authors is to provide sufficient conditions in order to steer an initial configuration of agents into a desired final one, by acting through a control term on the vector field, under the constraint that the action can be implemented only in a certain fixed space region.

Also the extension of classical *viability theory* to multi-agent systems is attracting an increasing interest in the community. Similarly to the finite-dimensional framework, a subset $\mathcal{K}$ of probability measures is said to be viable for a controlled dynamics if it is possible to keep the evolution confined inside $\mathcal{K}$ by acting with an admissible control when starting with a initial state in $\mathcal{K}$. We refer to [5] for first results in this direction based on a geometric approach (tangent cones to $\mathcal{K}$) and to the preprint [15] for a viscous-type approach to the problem.

It is worth pointing out that a key feature of all these studies, and many others available in the literature, is the combined use of tools, concepts, and techniques from optimal transport theory, measure theory, and from optimal control theory.

In this paper, we deal with a time-optimal control problem. More precisely, given a target set of desired final configurations, we are interested in the minimum time needed to steer the agents to it starting from an initial distribution, and respecting the nonholonomic constraints. In particular, in our measure-theoretic setting, the target set is given in duality with the space $C^0_b(\mathbb{R}^d)$ of continuous and bounded functions as follows. Given a family of observables $\Phi \subseteq C^0_b(\mathbb{R}^d)$, the target set is defined as (see Definition 4.2)

$$\tilde{S}_p^\Phi := \left\{ \mu \in \mathcal{P}_p(\mathbb{R}^d) : \int_{\mathbb{R}^d} \varphi(x)\, d\mu(x) \leq 0 \text{ for all } \varphi \in \Phi \right\},$$

where $(\mathcal{P}_p(\mathbb{R}^d), W_p)$, with $p \geq 1$, is the p-Wasserstein space of probability measures endowed with the metric W_p (see Definition 2.4). Section 4 is entirely devoted to the analysis of topological properties of this class of generalized target sets. In Section 3, we study the set of admissible trajectories $\mathcal{A}_I^p(\mu)$ defined on a time interval $I \subset \mathbb{R}$ and starting from a given initial datum $\mu \in \mathcal{P}_p(\mathbb{R}^d)$, i.e., those absolutely continuous curves in $\mathcal{P}_p(\mathbb{R}^d)$ whose driving velocity field satisfies the nonlocal nonholonomic constraint given by a set-valued map $F : \mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \qquad v_t(x) \in F(\mu_t, x) \text{ for a.e. } t \in I \text{ and } \mu_t\text{-a.e. } x \in \mathbb{R}^d.$$

This equation represents the controlled dynamics of the system. The results obtained, expecially Theorem 3.6 providing Filippov-Gronwall type estimates, are then used in Section 5 where we study the main object of the paper, i.e., the *minimum time function* defined as

$$\tilde{T}_p(\mu) := \inf\{T \geq 0 : \text{ there exists } \boldsymbol{\mu} = \{\mu_t\}_{t \in [0,T]} \in \mathcal{A}_{[0,T]}^p(\mu) \text{ s.t. } \mu_T \in \tilde{S}_p^\Phi\}.$$

More precisely, we first prove the existence of optimal trajectories, lower semicontinuity of the minimum time function and a Dynamic Programming Principle as in the classical case.

In the case without interactions, by passing to the limit in the Dynamic Programming Principle, in [13] the authors proved that the minimum time function solves in a suitable viscosity sense an Hamilton-Jacobi-Bellman equation in the spaces of measures provided that it is *continuous* (not just l.s.c.), and further development on this theory have been recently done in [30, 26]. We refer the reader to [2, 24, 25] for an introduction to Hamilton-Jacobi equations in Wasserstein spaces.

The aim of this paper is to provide a sufficient condition for the continuity of the minimum time function in this framework, i.e., sufficient conditions granting Small Time Local Attainability (STLA) in the sense of [27]. Indeed, assuming STLA, in Proposition 9 we get the continuity of the minimum time function thanks to the Filippov estimate proved in Theorem 3.6 and the Dynamic Programming Principle. Sufficient conditions for STLA are finally provided in the main Theorem 5.5, assuming geometric properties of the generalized target set $\tilde{S}_p^\Phi$ together with a sort of gradient-descent behavior of the associated family of observables Φ when integrated along admissible trajectories (see Definition 5.4). This condition represents a weakening of the well-known Petrov's condition in the classical framework.

The paper is structured as follows: in Section 2 we fix the notation and review some basic results about measure theory, optimal transport and set-valued analysis, in Section 3 we prove some basic properties of the admissible trajectories in the space of measures, in Section 4 we discuss some geometric properties of the target sets. In Section 5 we state our main result concerning the continuity of the minimum time function, and finally in Section 6 we compare our sufficient condition for STLA with the finite-dimensional one of [27].

2. **Preliminaries and notation.** In this section we review some concepts from measure theory, optimal transport and set-valued analysis. Our main references for this part are [3], [4], and [32].

We will use the following notation.

$B(x,r)$	the open ball of center $x \in X$ and radius r of a normed space X, i.e., $B(x,r) := \{y \in X : \|y-x\|_X < r\}$;
$\overline{K}$	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 function from a Banach space X to Y, endowed with $\|f\|_\infty = \sup\limits_{x\in X} \|f(x)\|_Y$ (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 $C_b^0(X;Y)$;
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 $\mathbb{R}^d$;
$AC([0,T])$	the set of absolutely continuous curves from $[0,T]$ to $\mathbb{R}^d$;
e_t	the evaluation operator $e_t : \mathbb{R}^d \times \Gamma_I \to \mathbb{R}^d$ 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 by $C_b^0(X)$;
$\mathcal{M}(\mathbb{R}^d;\mathbb{R}^d)$	the set of vector-valued Borel measures on $\mathbb{R}^d$ with values in $\mathbb{R}^d$, endowed with the weak* topology induced by $C_c^0(\mathbb{R}^d;\mathbb{R}^d)$;
$\lvert\nu\rvert$	the total variation of a measure $\nu \in \mathcal{M}(\mathbb{R}^d;\mathbb{R}^d)$;
$\ll$	the absolutely continuity relation between measures defined on the same σ-algebra;
$\mathrm{m}_p(\mu)$	the p-moment of a probability measure $\mu \in \mathcal{P}(X)$;
$r\sharp\mu$	the push-forward of the measure μ by the Borel map r;
$\mu \otimes \eta_x$	the product measure of $\mu \in \mathcal{P}(X)$ with the Borel family of measures $\{\eta_x\}_{x\in X}$;
pr_i	the i-th projection map $\mathrm{pr}_i(x_1,\dots,x_N) = x_i$;
$\Pi(\mu,\nu)$	the set of admissible transport plans from μ to ν;
$\Pi_o(\mu,\nu)$	the set of optimal transport plans from μ to ν;
$W_p(\mu,\nu)$	the p-Wasserstein distance between μ and ν;
$\mathcal{P}_p(X)$	the subset of the elements $\mathcal{P}(X)$ with finite p-moment, endowed with the p-Wasserstein distance;
$\mathcal{L}^d$	the Lebesgue measure on $\mathbb{R}^d$;
$\dfrac{\nu}{\mu}$	the Radon-Nikodym derivative of the measure ν w.r.t. the measure μ;
$\mathrm{Lip}(f)$	the Lipschitz constant of a function f.
$\mathcal{A}_I^p(\mu)$	the set of admissible trajectories defined in (1).
$\Upsilon_F(\mu,\boldsymbol{\theta})$	the set defined in Definition 3.4.
K_F	the quantity defined in Hypothesis 3.3 by $K_F := \max\limits_{v\in F(\delta_0,0)} \{\lvert v\rvert\}$.

In this section we give some preliminaries and fix the notation. Our main reference for this part is [3].

Definition 2.1 (Space of probability measures)**.** 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. For any $p \geq 1$, the p-moment of $\mu \in \mathcal{P}(X)$ is defined by $\mathrm{m}_p(\mu) = \int_X \|x\|_X^p \, d\mu(x)$, and we set $\mathcal{P}_p(X) = \{\mu \in \mathcal{P}(X) : \mathrm{m}_p(\mu) < +\infty\}$. For any Borel map $r : X \to Y$ and $\mu \in \mathcal{P}(X)$, we define the *push forward measure* $r\sharp\mu \in \mathcal{P}(Y)$ by setting $r\sharp\mu(B) = \mu(r^{-1}(B))$ for any Borel set B of Y.

Definition 2.2 (Total variation)**.** Let X, Y be Banach spaces, and 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 the set of countable collections $\{B_i\}_{i\in\mathbb{N}}$ of pairwise disjoint Borel sets such that $\bigcup_{i\in\mathbb{N}} B_i = B$.

For the following result see [3, Theorem 5.3.1].

Theorem 2.3 (Disintegration)**.** *Given a measure $\mu \in \mathcal{P}(\mathbb{X})$ and a Borel map $r : \mathbb{X} \to X$, there exists a family of probability measures $\{\mu_x\}_{x\in X} \subseteq \mathcal{P}(\mathbb{X})$, uniquely defined for $r\sharp\mu$-a.e. $x \in X$, such that $\mu_x(\mathbb{X} \setminus r^{-1}(x)) = 0$ for $r\sharp\mu$-a.e. $x \in X$, and for any Borel map $\varphi : \mathbb{X} \to [0, +\infty]$ we have*

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

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.

Definition 2.4 (Transport plans and Wasserstein distance)**.** Let X be a complete separable Banach space, $\mu_1, \mu_2 \in \mathcal{P}(X)$. We define the set of *admissible transport plans* between μ_1 and μ_2 by setting

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

where for $i = 1, 2$, we defined $\mathrm{pr}_i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ by $\mathrm{pr}_i(x_1, x_2) = x_i$. The *inverse* $\boldsymbol{\pi}^{-1}$ of a transport plan $\boldsymbol{\pi} \in \Pi(\mu, \nu)$ is defined by $\boldsymbol{\pi}^{-1} = i\sharp\boldsymbol{\pi} \in \Pi(\nu, \mu)$, where $i(x, y) = (y, x)$ for all $x, y \in X$. The *p-Wasserstein distance* between μ_1 and μ_2 is

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

If $\mu_1, \mu_2 \in \mathcal{P}_p(X)$ then the above infimum is actually a minimum, and we define

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

The space $\mathcal{P}_p(X)$ endowed with the W_p-Wasserstein distance is a complete separable metric space, moreover for all $\mu \in \mathcal{P}_p(X)$ there exists a sequence $\{\mu^N\}_{N\in\mathbb{N}} \subseteq \mathrm{co}\{\delta_x : x \in \mathrm{supp}\,\mu\}$ such that $W_p(\mu^N, \mu) \to 0$ as $N \to +\infty$.

Remark 1. Recalling formula (5.2.12) in [3], we have

$$W_p(\delta_0, \mu) = m_p^{1/p}(\mu) = \left(\int_{\mathbb{R}^d} |x|^p \, d\mu(x) \right)^{1/p}$$

for all $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. In particular, if $t \mapsto \mu_t$ is W_p-continuous, then $t \mapsto \mathrm{m}_p^{1/p}(\mu_t)$ is continuous.

Definition 2.5 (Set-valued maps)**.** Let X, Y be sets. A *set-valued* map F from X to Y is a map associating to each $x \in X$ a (possible empty) subset $F(x)$ of Y. We will write $F : X \rightrightarrows Y$ to denote a set-valued map from X to Y. The *graph* of a set-valued map F is

$$\operatorname{graph} F := \{(x, y) \in X \times Y : y \in F(x)\} \subseteq X \times Y,$$

while the *domain* of F is $\operatorname{dom} F := \{x \in X : F(x) \neq \emptyset\} \subseteq X$. A *selection* of F is a map $f : \operatorname{dom} F \to Y$ such that $f(x) \in F(x)$ for all $x \in \operatorname{dom} F$. When X, Y are topological spaces, we say that

- F has *closed images* if $F(x)$ is closed in Y for every $x \in X$,
- F has *closed graph* if $\operatorname{graph} F$ is closed in $X \times Y$,
- F is *compact valued* (or that it has *compact images*) if $F(x)$ is compact for every $x \in X$,
- F is *upper semicontinuous* at $x \in X$ if for every open set $V \subseteq Y$ such that $V \supseteq F(x)$ there exists an open neighborhood $U \subseteq X$ of x such that $F(z) \subseteq V$ for all $z \in U$.
- F is *lower semicontinuous* at $x \in X$ if for every open set $V \subseteq Y$ such that $V \cap F(x) \neq \emptyset$ there exists an open neighborhood $U \subseteq X$ of x such that $F(z) \cap V \neq \emptyset$ for all $z \in U$.
- F is continuous at $x \in X$ if it is both lower and upper semicontinuous at x.
- F will be called continuous (resp. lower semicontinuous, upper semicontinuous) if it is continuous (resp. lower semicontinuous, upper semicontinuous) at every $x \in X$.

When Y is a vector space, F is convex valued (or it has *convex images*) if $F(x)$ is convex for every $x \in X$. When X, Y are measurable spaces, we say that F is measurable if $\operatorname{graph} F$ is measurable in $X \times Y$ endowed with the product of σ-algebrae on X and Y. When (X, d) is a metric space and Y is a normed space, given $L > 0$ we say that F is Lipschitz continuous with constant L if for all $x_1, x_2 \in X$

$$F(x_2) \subseteq F(x_1) + L \cdot d(x_1, x_2)\overline{B_Y(0,1)},$$

where the sum and the product of sets are in the Minkowski sense: $A + B = \{a + b : a \in A,\, b \in B\}$ and $\lambda A = \{\lambda a : a \in A\}$ for every $A, B \subseteq Y$, $\lambda \in \mathbb{R}$.

3. **Admissible trajectories.** Given a collection $\boldsymbol{\mu} = \{\mu_h\}_{h \in I} \subseteq \mathcal{P}(X)$ of Borel measures on the measure space X indexed by a parameter $h \in I$, by a slight abuse of notation we will denote by $\boldsymbol{\mu}$ both the set $\boldsymbol{\mu} = \{\mu_h\}_{h \in I} \subseteq \mathcal{P}(X)$ and the function $h \mapsto \mu_h$. In each occurrence, the context will clarify what we are referring to.

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

We say that $\boldsymbol{\mu}$ is an *admissible trajectory driven by* $\boldsymbol{\nu}$ defined on I with underlying dynamics F if

- $|\nu_t| \ll \mu_t$ for a.e. $t \in I$;
- $v_t(x) := \dfrac{\nu_t}{\mu_t}(x) \in F(\mu_t, x)$ 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 on $[0, T] \times \mathbb{R}^d$, equivalently
$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x)\, d\mu_t(x) = \int_{\mathbb{R}^d} \langle \nabla \varphi(x), v_t(x) \rangle\, d\mu_t(x),$$
for a.e. $t \in [0, T]$ and all $\varphi \in C_c^1(\mathbb{R}^d)$.

Define $\mathcal{A}^p_I : \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ by

$$\mathcal{A}^p_I(\mu) := \left\{ \boldsymbol{\mu} = \{\mu_t\}_{t\in I} : \begin{array}{l} \boldsymbol{\mu} \text{ is an admissible trajectory} \\ \text{with } \mu_a = \mu \end{array} \right\}. \tag{1}$$

When I and p are clear by the context, we will omit them.

Definition 3.2 (Concatenation, restriction, extension)**.** Let $I_i = [a_i, b_i]$, $\boldsymbol{\mu}^{(i)} = \{\mu^{(i)}_t\}_{t\in I_i} \in \mathcal{A}^p_{I_i}(\mu^{(i)}_{a_i})$, $i = 1, 2$, be satisfying $\mu^{(1)}_{b_1} = \mu^{(2)}_{a_2}$. We define $I_3 = [a_1, b_1 + b_2 - a_2]$ and $\boldsymbol{\mu}^{(3)} = \{\mu^{(3)}_t\}_{t\in I_3}$ by setting $\mu^{(3)}_t = \mu^{(1)}_t$ for $t \in I_1$ and $\mu^{(3)}_t = \mu^{(2)}_{t+a_2-b_1}$ for $t \in I_3 \setminus I_1$. The curve $\boldsymbol{\mu}^{(3)}$ will be called the *concatenation* of $\boldsymbol{\mu}^{(1)}$ and $\boldsymbol{\mu}^{(2)}$ and will be denoted by $\boldsymbol{\mu}^{(3)} = \boldsymbol{\mu}^{(1)} \odot \boldsymbol{\mu}^{(2)}$. By [16, Lemma 4.4], we have $\boldsymbol{\mu}^{(3)} \in \mathcal{A}^p_{I_3}(\mu^{(1)}_{a_1})$.

Let $I = [a, b]$, $J = [a', b']$ with $J \subseteq I$, and $\boldsymbol{\mu} = \{\mu_t\}_{t\in I} \in \mathcal{A}^p_I(\mu_a)$. The *restriction* $\boldsymbol{\mu}_{|J} = \{\hat{\mu}_t\}_{t\in J}$ of $\boldsymbol{\mu}$ to J is defined by taking $\hat{\mu}_t = \mu_t$ for all $t \in J$ and we have $\boldsymbol{\mu}_{|J} \in \mathcal{A}^p_J(\mu_{a'})$.

Let $\boldsymbol{\mu}^{(i)} \in \mathcal{A}^p_{I_i}(\mu^{(i)})$, $i = 1, 2$. We say that $\boldsymbol{\mu}^{(2)}$ is an *extension* of $\boldsymbol{\mu}^{(1)}$ if $I_2 \supseteq I_1$ and $\boldsymbol{\mu}^{(2)}_{|I_1} = \boldsymbol{\mu}^{(1)}$.

Throughout the paper, we will assume the following

Hypothesis 3.3. The set-valued map $F : \mathcal{P}_p(\mathbb{R}^d)\times\mathbb{R}^d \to \mathbb{R}^d$ has nonempty, convex and compact images, moreover it is Lipschitz continuous with constant $L > 0$ with respect to the metric

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

on $\mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d$. We set $K_F := \max\limits_{v\in F(\delta_0,0)} \{|v|\}$.

Definition 3.4 (Definition of Υ_F)**.** Assume Hypothesis 3.3 for F. Let $\boldsymbol{\theta} = \{\theta_t\}_{t\in[0,T]}$ be a W_p-continuous curve in $\mathcal{P}_p(\mathbb{R}^d)$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Denote by $\Upsilon_F(\mu, \boldsymbol{\theta})$ the set of $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathcal{P}_p(\mathbb{R}^d)$ satisfying the following property: there exists $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ such that

- $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0, T]$, $\mu_0 = \mu$;
- for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^d\times\Gamma_T$ and a.e. $t \in [0, T]$ it holds $\gamma \in AC([0, T])$, $\gamma(0) = x$, $\dot{\gamma}(t) \in F(\theta_t, \gamma(t))$.

We set $M_{\boldsymbol{\theta}} := \sup\limits_{\tau\in[0,T]} \mathrm{m}_p^{1/p}(\theta_\tau)$ and

$$\Xi(\mu, \boldsymbol{\theta}) := \{\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T) : \{e_t\sharp\boldsymbol{\eta}\}_{t\in[0,T]} \in \Upsilon_F(\mu, \boldsymbol{\theta})\}.$$

On the set $X := \mathcal{P}_p(\mathbb{R}^d) \times C^0([0, T]; \mathcal{P}_p(\mathbb{R}^d))$ we define the metric

$$d_X\left(\left(\mu^{(1)}, \boldsymbol{\theta^{(1)}}\right), \left(\mu^{(2)}, \boldsymbol{\theta^{(2)}}\right)\right) = W_p\left(\mu^{(1)}, \mu^{(2)}\right) + \sup_{t\in[0,T]} W_p\left(\theta^{(1)}_t, \theta^{(2)}_t\right),$$

where $\boldsymbol{\theta^{(i)}} = \{\theta^{(i)}_t\}_{t\in[0,T]}$, $i = 1, 2$.

Finally, we define the set-valued map $S^{\boldsymbol{\theta}} : \mathbb{R}^d \rightrightarrows \Gamma_T$ by setting for all $y \in \mathbb{R}^d$

$$S^{\boldsymbol{\theta}}(y) := \{\xi \in AC([0, T]) : \dot{\xi} \in F(\theta_t, \xi(t)) \text{ for a.e. } t \in [0, T],\ \xi(0) = y\}.$$

Since $t \mapsto F(\theta_t, x)$ is continuous for all $x \in \mathbb{R}^d$ and $x \mapsto F(\theta_t, x)$ is Lipschitz continuous by Hypothesis 3.3, with constant L for all $t \in [0, T]$, the set-valued map $S^{\boldsymbol{\theta}}(\cdot)$ is Lipschitz continuous by [4, Corollary 10.4.2].

Lemma 3.5 (Estimates on the moments)**.** *Assume Hypothesis 3.3 for F. Let $\boldsymbol{\theta} = \{\theta_t\}_{t\in[0,T]}$ be a W_p-continuous curve in $\mathcal{P}_p(\mathbb{R}^d)$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Let $\Xi(\mu,\boldsymbol{\theta})$, $\Upsilon_F(\mu,\boldsymbol{\theta})$ and $M_{\boldsymbol{\theta}}$ be as in Definition 3.4, and let $\boldsymbol{\eta} \in \Xi(\mu,\boldsymbol{\theta})$ and $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \Upsilon_F(\mu,\boldsymbol{\theta})$ such that $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0,T]$. Then for all $t,s \in [0,T]$*

$$\begin{aligned}
\mathrm{m}_p^{1/p}(\mu_t) \leq& e^{LT}\left(\mathrm{m}_p^{1/p}(\mu) + K_F T + LTM_{\boldsymbol{\theta}}\right),\\
W_p(\mu_t,\mu_s) \leq& \left(K_F + LM_{\boldsymbol{\theta}} + Le^{LT}\left(\mathrm{m}_p^{1/p}(\mu) + K_F T + LTM_{\boldsymbol{\theta}}\right)\right)\cdot|t-s|,\\
\int_{\mathbb{R}^d\times\Gamma_T} \|\dot\gamma\|^p_{L^\infty([0,1])}\,d\boldsymbol{\eta}(x,\gamma) \leq& \left[K_F + L\left(e^{LT}\left(\mathrm{m}_p^{1/p}(\mu) + K_F T + LTM_{\boldsymbol{\theta}}\right) + M_{\boldsymbol{\theta}}\right)\right]^p.
\end{aligned}$$

Proof. Set $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \Upsilon_F(\mu,\boldsymbol{\theta})$. For $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d\times\Gamma_T$ and a.e. $t \in [0,T]$ we have

$$\dot\gamma(t) \in F(\theta_t,\gamma(t)) \subseteq F(\delta_0,0) + L(|\gamma(t)| + \mathrm{m}_p^{1/p}(\theta_t))\overline{B(0,1)}$$

Thus for all $s,t \in [0,T]$ and a.e. $\tau \in [0,T]$

$$\begin{aligned}
|\dot\gamma(\tau)| \leq& K_F + L(|\gamma(\tau)| + \mathrm{m}_p^{1/p}(\theta_\tau)) \leq K_F + LM_{\boldsymbol{\theta}} + L|\gamma(\tau)|,\\
|\gamma(t)| - |\gamma(0)| \leq& \int_0^t |\dot\gamma(\tau)|\,d\tau \leq (K_F + LM_{\boldsymbol{\theta}})T + L\int_0^t |\gamma(\tau)|\,d\tau,\\
|\gamma(t)-\gamma(s)| \leq& \left|\int_s^t |\dot\gamma(\tau)|\,d\tau\right| \leq (K_F + LM_{\boldsymbol{\theta}})|t-s| + L\left|\int_s^t |\gamma(\tau)|\,d\tau\right|.
\end{aligned}$$

By Grönwall lemma, this implies for all $0 \leq s \leq t \in [0,T]$

$$\begin{aligned}
|\gamma(t)| \leq& e^{Lt}\left(|\gamma(0)| + (K_F + LM_{\boldsymbol{\theta}})T\right),\\
\|\dot\gamma\|_{L^\infty([0,1])} \leq& K_F + L\left(M_{\boldsymbol{\theta}} + e^{LT}\left(|\gamma(0)| + (K_F + LM_{\boldsymbol{\theta}})T\right)\right),\\
|\gamma(t)-\gamma(s)| \leq& \left(K_F + LM_{\boldsymbol{\theta}} + Le^{LT}\left(|\gamma(0)| + (K_F + LM_{\boldsymbol{\theta}})T\right)\right)\cdot|t-s|,
\end{aligned} \tag{2}$$

recalling that $\dot\gamma(s) \in F(\mu_s,\gamma(s))$ for a.e. s.

We conclude by taking the $L^p_{\boldsymbol{\eta}}$ norm of the above inequalities and using the triangular inequality. □

Proposition 1 (Upper semicontinuity of the solution map)**.** *Set $X := \mathcal{P}_p(\mathbb{R}^d) \times C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$.*

The set-valued map $\Upsilon_F : X \rightrightarrows C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$, defined in Definition 3.4, is upper semicontinuous with compact nonempty images.

Proof. We prove first that $\Upsilon_F(\mu,\boldsymbol{\theta}) \neq \emptyset$ for all $(\mu,\boldsymbol{\theta}) \in X$. Consider now the set-valued map $S^{\boldsymbol{\theta}}(\cdot)$ defined as in Definition 3.4. Since it is Lipschitz continuous, it has a Borel selection. Thus let $h_0 : \mathbb{R}^d \to AC([0,T])$ be a Borel map such that $h_0(x) \in S^{\boldsymbol{\theta}}(x)$ for every $x \in \mathbb{R}^d$. Define $\boldsymbol{\eta} = \mu\otimes\delta_{h_0(x)}$, $\mu_t = e_t\sharp\boldsymbol{\eta}$, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$. Then, by construction, we have $\boldsymbol{\mu} \in \Upsilon_F(\mu,\boldsymbol{\theta})$.

Let now $\left\{\left(\mu^{(n)},\boldsymbol{\theta}^{(n)}\right)\right\}_{n\in\mathbb{N}} \subseteq X$ be a sequence d_X-converging to $(\mu,\boldsymbol{\theta}) \in X$, and $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}} \subseteq C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$, $\{\boldsymbol{\eta}^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(\mathbb{R}^d\times\Gamma_T)$ be such that

- $\boldsymbol{\theta}^{(n)} = \{\theta_t^{(n)}\}_{t\in[0,T]}$, $\boldsymbol{\theta} = \{\theta_t\}_{t\in[0,T]}$;
- $\boldsymbol{\mu}^{(n)} \in \Upsilon_F(\mu^{(n)},\boldsymbol{\theta}^{(n)})$ and $\boldsymbol{\eta}^{(n)} \in \Xi(\mu^{(n)},\boldsymbol{\theta}^{(n)})$ for all $n \in \mathbb{N}$;
- $\boldsymbol{\mu}^{(n)} = \{\mu_t^{(n)}\}_{t\in[0,T]}$ with $\mu_t^{(n)} = e_t\sharp\boldsymbol{\eta}^{(n)}$ for all $t \in [0,T]$ and $n \in \mathbb{N}$,

where $\Xi(\cdot,\cdot)$ is defined as in Definition 3.4. We prove that the sequence $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ has always cluster points, and all the cluster points are contained in $\Upsilon_F(\mu,\boldsymbol{\theta})$. This will imply in particular that $\Upsilon_F(\cdot)$ has compact images (by taking constant sequences $(\mu^{(n)},\boldsymbol{\theta}^{(n)})\equiv(\mu,\boldsymbol{\theta})$).

For n sufficiently large, we have $\mathrm{m}_p^{1/p}(\mu^{(n)})\le \mathrm{m}_p^{1/p}(\mu)+1$ and $M_{\boldsymbol{\theta}^{(n)}}\le M_{\boldsymbol{\theta}}+1$, recalling the definition of the convergence in X and the definition for $M_{\boldsymbol{\theta}}$ given in Definition 3.4. Thus, by applying the estimates of Lemma 3.5, we have

$$\begin{aligned}
&\mathrm{m}_p^{1/p}(\mu_t^{(n)})\le e^{LT}\left(m_p^{1/p}(\mu)+1+K_FT+LTM_{\boldsymbol{\theta}}+LT\right),\\
&W_p(\mu_t^{(n)},\mu_s^{(n)})\le\\
&\le\left(K_F+LM_{\boldsymbol{\theta}}+L+Le^{LT}\left(\mathrm{m}_p^{1/p}(\mu)+1+K_FT+LTM_{\boldsymbol{\theta}}+LT\right)\right)\cdot|t-s|,\\
&\int_{\mathbb{R}^d\times\Gamma_T}\|\dot\gamma\|^p_{L^\infty([0,1])}\,d\boldsymbol{\eta}^{(n)}(x,\gamma)\le\\
&\le\left[K_F+L\left(e^{LT}\left(\mathrm{m}_p^{1/p}(\mu)+1+K_FT+LTM_{\boldsymbol{\theta}}+LT\right)+M_{\boldsymbol{\theta}}+1\right)\right]^p.
\end{aligned}$$

In particular

- $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ is equicontinuous;
- for all $t\in[0,T]$, we have that $\{\mu_t^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact in $\mathscr{P}_p(\mathbb{R}^d)$, since it has p-moment uniformly bounded.

Thus $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact in $C^0([0,T];\mathscr{P}_p(\mathbb{R}^d))$ by Ascoli-Arzelà theorem. Up to passing to a subsequence, we may assume that there exists $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}\in C^0([0,T];\mathscr{P}_p(\mathbb{R}^d))$ such that

$$\lim_{n\to+\infty}\sup_{t\in[0,T]}W_p(\mu_t,\mu_t^{(n)})=0.$$

We notice also that the functional $\Psi:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}\cup\{+\infty\}$

$$\Psi(x,\gamma):=\begin{cases}(|x|+|\gamma(0)|+\|\dot\gamma\|_{L^\infty})^p, & \text{if } \gamma\in\mathrm{Lip}([0,T];\mathbb{R}^d),\\ \\ +\infty, & \text{otherwise,}\end{cases}$$

has compact sublevels in $\mathbb{R}^d\times\Gamma_T$ by Ascoli-Arzelà theorem. Since

$$\begin{aligned}
&\sup_{n\in\mathbb{N}}\int_{\mathbb{R}^d\times\Gamma_T}\Psi(x,\gamma)\,d\boldsymbol{\eta}^{(n)}(x,\gamma)\le\\
&\qquad\le 2^{p-1}\sup_{n\in\mathbb{N}}\left[2\mathrm{m}_p(\mu^{(n)})+\int_{\mathbb{R}^d\times\Gamma_T}\|\dot\gamma\|^p_{L^\infty([0,1])}\,d\boldsymbol{\eta}^{(n)}(x,\gamma)\right]<+\infty,
\end{aligned}$$

we have that $\{\boldsymbol{\eta}^{(n)}\}_{n\in\mathbb{N}}$ is tight in $\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$. Thus, up to passing to a subsequence, we may assume also that there exists $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ such that $\boldsymbol{\eta}^{(n)}\rightharpoonup\boldsymbol{\eta}$ narrowly. By the continuity of $e_t:\mathbb{R}^d\times\Gamma_T\to\mathbb{R}^d$, we have $\mu_t=e_t\sharp\boldsymbol{\eta}$. By [3, Proposition 5.1.8], for $\boldsymbol{\eta}$-a.e. $(x,\gamma)\in\mathbb{R}^d\times\Gamma_T$ there exists a sequence $\{(x_n,\gamma_n)\}_{n\in\mathbb{N}}$ such that $x_n=\gamma_n(0)$, $\gamma_n\in AC([0,T])$, $\dot\gamma_n(t)\in F(\theta_t^{(n)},\gamma_n(t))$ for a.e. $t\in[0,T]$ and for all $n\in\mathbb{N}$ with $x_n\to x$, $\|\gamma_n-\gamma\|_\infty\to 0$ as $n\to+\infty$.

By (2) and recalling that $M_{\boldsymbol{\theta}^{(n)}}\le M_{\boldsymbol{\theta}}+1$ and $|x_n|\le|x|+1$ for n sufficiently large, we have $n\in\mathbb{N}$,

$$\|\dot\gamma_n\|_{L^\infty([0,1])}\le K_F+L\left(M_{\boldsymbol{\theta}}+1+e^{LT}\left(|x|+1+(K_F+LM_{\boldsymbol{\theta}}+L)T\right)\right).$$

In particular, by Ascoli-Arzelà Theorem, we have that γ is Lipschitz continuous. For a.e. $t,\tau\in[0,T]$ we have also

$$\begin{aligned}
&F(\theta_\tau^{(n)},\gamma_n(\tau))\subseteq F(\theta_t,\gamma(t))+\\
&\quad+L\left(W_p(\theta_\tau^{(n)},\theta_\tau)+W_p(\theta_\tau,\theta_t)+|\gamma_n(\tau)-\gamma(\tau)|+|\gamma(t)-\gamma(\tau)|\right)\overline{B(0,1)}\\
&\subseteq F(\theta_t,\gamma(t))+L\left(\sup_{\tau\in[0,T]}W_p(\theta_\tau^{(n)},\theta_\tau)+W_p(\theta_\tau,\theta_t)+\|\gamma_n-\gamma\|_\infty+\mathrm{Lip}(\gamma)\cdot|t-\tau|\right)\overline{B(0,1)}
\end{aligned}$$

For every $\varepsilon>0$ there is $n_\varepsilon\in\mathbb{N}$ such that if $n>n_\varepsilon$ we have for a.e. $t,\tau\in[0,T]$

$$\dot\gamma_n(\tau)\in F(\theta_t,\gamma(t))+L(\varepsilon+W_p(\theta_\tau,\theta_t)+\mathrm{Lip}(\gamma)|t-\tau|)\overline{B(0,1)}.$$

In particular, let $t\in[0,T]$ be a differentiability point of γ_n. We have for all $z\in\mathbb{R}^d$, $s\in[0,T]$, $s\neq t$, and $n>n_\varepsilon$

$$\begin{aligned}
&\langle\frac{\gamma_n(s)-\gamma_n(t)}{s-t},z\rangle=\frac{1}{s-t}\int_t^s\langle z,\dot\gamma_n(\tau)\rangle\,d\tau\\
&\leq\sup_{v\in F(\theta_t,\gamma(t))}\langle z,v\rangle+L\varepsilon|z|+L|z|\mathrm{Lip}(\gamma)\frac{1}{s-t}\int_t^s|t-\tau|\,d\tau+L|z|\frac{1}{s-t}\int_s^t W_p(\theta_\tau,\theta_t)\,d\tau
\end{aligned}$$

By letting $n\to+\infty$ and $s\to t$ we conclude that $\dot\gamma(t)\in F(\theta_t,\gamma(t))$ since $F(\theta_t,\gamma(t))$ is closed and convex. Hence $\boldsymbol{\mu}\in\Upsilon_F(\mu,\boldsymbol{\theta})$, which completes the proof. $\square$

Proposition 2 (Superposition Principle)**.** *Assume Hypothesis 3.3 for F, and let $T>0$. Then $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory if and only if $\boldsymbol{\mu}\in\Upsilon_F(\mu_0,\boldsymbol{\mu})$, with $\Upsilon_F(\cdot,\cdot)$ defined in Definition 3.4.*

Proof.

1. *Sufficience.* Assume that $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}\in\Upsilon_F(\mu_0,\boldsymbol{\mu})$. Let $\Xi(\cdot,\cdot)$ be as in Definition 3.4. Then there exists $\boldsymbol{\eta}\in\Xi(\mu_0,\boldsymbol{\mu})$ such that $\mu_t=e_t\sharp\boldsymbol{\eta}$. Set

$$\mathcal{N}:=\left\{(t,x,\gamma)\in[0,T]\times\mathbb{R}^d\times\Gamma_T:\nexists\dot\gamma(t)\text{ or }\dot\gamma(t)\notin F(\mu_t,\gamma(t))\text{ or }\gamma(0)\neq x\right\}.$$

Since $\mathcal{L}^1\otimes\boldsymbol{\eta}\,(\mathcal{N})=0$, for $\boldsymbol{\eta}$-a.e. $(x,\gamma)\in\mathbb{R}^d\times\Gamma_T$ and a.e. $t\in[0,T]$ we have that $\dot\gamma(t)$ exists and belongs to $F(\mu_t,\gamma(t))$, and $\gamma(0)=x$. Given $\varphi\in C^1_c(\mathbb{R}^d)$, we have

$$\left|\int_{\mathbb{R}^d}\varphi(x)\,d\mu_t(x)-\int_{\mathbb{R}^d}\varphi(x)\,d\mu_s(x)\right|\leq\|\nabla\varphi\|_\infty\int_{\mathbb{R}^d\times\Gamma_T}|\gamma(t)-\gamma(s)|\,d\boldsymbol{\eta}(x,\gamma).$$

According to (2), this implies that

$$t\mapsto\int_{\mathbb{R}^d}\varphi(x)\,d\mu_t(x)$$

is Lipschitz continuous. Hence its distributional derivative is in L^∞ and coincides with the pointwise derivative almost everywhere. Thus, in the sense of distributions in $]0,T[$, we obtain for all $\varphi\in C^1_c(\mathbb{R}^d)$

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(x)\,d\mu_t(x)=&\frac{d}{dt}\int_{\mathbb{R}^d\times\Gamma_T}\varphi(\gamma(t))\,d\boldsymbol{\eta}(x,\gamma)=\int_{\mathbb{R}^d\times\Gamma_T}\langle\nabla\varphi(\gamma(t)),\dot\gamma(t)\rangle\,d\boldsymbol{\eta}(x,\gamma)\\
=&\int_{\mathbb{R}^d}\langle\nabla\varphi(y),\int_{e_t^{-1}(y)}\dot\gamma(t)\,d\eta_{t,y}(x,\gamma)\rangle\,d\mu_t(y),
\end{aligned}$$

where we disintegrated $\boldsymbol{\eta}$ w.r.t. e_t obtaining $\boldsymbol{\eta}=\mu_t\otimes\eta_{t,y}$ and used the fact that $\|\nabla\varphi\|_\infty$ is bounded, and that the map $\gamma\mapsto\|\dot\gamma\|_{L^\infty}$ is in $L^1_{\boldsymbol{\eta}}$ due to the uniform bound on the moments. By Jensen's inequality, we have

$$d_{F(\mu_t,y)}\left(\int_{e_t^{-1}(y)} \dot\gamma(t)\, d\eta_{t,y}(x,\gamma)\right) \le \int_{e_t^{-1}(y)} d_{F(\mu_t,y)}\left(\dot\gamma(t)\right)\, d\eta_{t,y}(x,\gamma) = 0,$$

and so for μ_t-a.e. $y \in \mathbb{R}^d$ and a.e. $t \in [0,T]$ we have

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

hence $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ with $\nu_t = v_t\mu_t$ for a.e. $t \in [0,T]$.

2. *Necessity.* Assume that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. Set $v_t(x) = \dfrac{\nu_t}{\mu_t}(x) \in F(\mu_t, x)$ for μ_t-a.e. $x \in \mathbb{R}^d$ and a.e. $t \in [0,T]$. Filippov's Theorem (see e.g. [4, Theorem 8.2.10]) implies that there exists a Borel selection $\xi(\cdot)$ of $F(\delta_0, 0)$, such that

$$|v_t(x) - \xi(x)| = d_{F(\delta_0,0)}(v_t(x))$$

for all $x \in \mathbb{R}^d$, and so we have

$$\begin{aligned}\left(\int_0^T \int_{\mathbb{R}^d} |v_t(x)|^p\, d\mu_t(x)\, dt\right)^{1/p} \le& \int_0^T \left(\int_{\mathbb{R}^d} |v_t(x)-\xi(x)|^p\, d\mu_t(x)\right)^{1/p} dt + \\ &+ \int_0^T \left(\int_{\mathbb{R}^d} |\xi(x)|^p\, d\mu_t(x)\right)^{1/p} dt \\ =& \int_0^T \left(\int_{\mathbb{R}^d} d^p_{F(\delta_0,0)}(v_t(x))\, d\mu_t(x)\right)^{1/p} dt + TK_F \\ \le& 2^{p-1}L \int_0^T \left[W_p(\delta_0,\mu_t) + \mathrm{m}_p^{1/p}(\mu_t)\right] dt + TK_F \\ \le& 2^p L \int_0^T \mathrm{m}_p^{1/p}(\mu_t)\, dt + TK_F < +\infty.\end{aligned}$$

By [3, Theorem 8.2.1], there exists $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = e_t \sharp \boldsymbol{\eta}$ for all $t \in [0,T]$ and $\boldsymbol{\eta}$ is concentrated on $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ with $\gamma \in AC([0,T])$, $\dot\gamma(t) = v_t(\gamma(t)) \in F(\mu_t, \gamma(t))$ for a.e. $t \in [0,T]$ and $\gamma(0) = x$. Thus $\boldsymbol{\mu} \in \Upsilon_F(\mu_0, \boldsymbol{\mu})$.

□

Remark 2. Notice that by Definition of Υ_F in Definition 3.4, Proposition 2 gives in fact a Superposition Principle along the line of [3, Theorem 8.2.1] adapted to nonlocal differential inclusions. Indeed, under the given assumptions, it states that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory if and only if there exists $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ such that

- $\mu_t = e_t \sharp \boldsymbol{\eta}$ for all $t \in [0,T]$, $\mu_{t=0} = \mu_0$;
- for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ and a.e. $t \in [0,T]$ it holds $\gamma \in AC([0,T])$, $\gamma(0) = x$, $\dot\gamma(t) \in F(\mu_t, \gamma(t))$.

In this case, we say that $\boldsymbol{\mu}$ is *represented* by $\boldsymbol{\eta}$.

Corollary 1 (Existence of admissible trajectories)**.** *Assume Hypothesis 3.3 for F and let $T > 0$. The set-valued map $\mathcal{A} : \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ is upper semicontinuous with nonempty compact images.*

Proof. Let $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Given $R > 0$, define

$$\mathcal{C}(R) := \Big\{ \boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathcal{P}_p(\mathbb{R}^d) : \mu_0 = \mu, \text{ and for all } t,s \in [0,T]\ \mathrm{m}_p^{1/p}(\mu_t) \le R,$$
$$W_p(\mu_t,\mu_s) \le \Big(K_F + LR + Le^{LT}\Big(\mathrm{m}_p^{1/p}(\mu) + K_F T + LTR\Big)\Big) \cdot |t-s| \Big\}.$$

Recalling that the concatenation of solutions of the continuity equation is again a solution of the continuity equation driven by the time concatenation of the vector fields (see [16, Lemma 4.4]), in order to prove that $\mathcal{A}(\mu) \neq \emptyset$ it is not restrictive to assume $LT < 1/2$. In particular, we have $1 - e^{LT}LT > 0$. Define

$$R := \frac{e^{LT}(\mathrm{m}_p^{1/p}(\mu) + K_F T)}{1 - e^{LT}LT} \ge \mathrm{m}_p^{1/p}(\mu).$$

Notice that $\mathcal{C}(R) \neq \emptyset$, since it contains the constant curve $\mu_t \equiv \mu$ for all $t \in [0,T]$, it is convex, and it is compact in $C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ by Ascoli-Arzelà theorem. Moreover, defining $\Upsilon_F(\cdot,\cdot)$ as in Definition 3.4, we have $\Upsilon_F(\mu,\mathcal{C}(R)) \subseteq \mathcal{C}(R)$ by Lemma 3.5 and the choice of R. By Kakutani-Ky Fan Theorem (see e.g. [19, Theorem 1]) we have that there exists $\boldsymbol{\mu} \in \mathcal{C}(R)$ such that $\boldsymbol{\mu} \in \Upsilon_F(\mu,\boldsymbol{\mu})$, i.e., by Proposition 2, $\boldsymbol{\mu}$ is an admissible trajectory starting from μ. All the other properties of $\mathcal{A}(\cdot)$ trivially follows from the fact that $\Upsilon_F(\cdot)$ is upper semicontinuous with nonempty compact images. □

Remark 3. An alternative proof of existence of admissible trajectories, i.e. $\mathcal{A}(\mu) \neq \emptyset$, can be found for example in [31, Theorem 6.1] where the author provides sufficient conditions in order to ensure existence (and uniqueness) of solutions of a continuity equation for some given non-local vector field.

Theorem 3.6 (Filippov-type estimate for the set of admissible trajectories)**.** *Assume Hypothesis 3.3 for F. Let $T > 0$, $\mu^{(A)}, \mu^{(B)} \in \mathcal{P}_p(\mathbb{R}^d)$ be given. Let $\boldsymbol{\mu^{(A)}} = \{\mu_t^{(A)}\}_{t\in[0,T]}$ be an admissible trajectory satisfying $\mu_0^{(A)} = \mu^{(A)}$.*

Then there exists an admissible trajectory $\boldsymbol{\mu^{(B)}} = \{\mu_t^{(B)}\}_{t\in[0,T]}$ satisfying $\mu_0^{(B)} = \mu^{(B)}$ such that

$$W_p(\mu_t^{(A)}, \mu_t^{(B)}) \le 2^{\frac{p-1}{p}} e^{L(2+Le^{LT})T} \cdot W_p(\mu^{(A)}, \mu^{(B)}) \quad \text{for all } t \in [0,T].$$

In particular, the set-valued map $\mathcal{A} : \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ is Lipschitz continuous.

Proof. Let $\boldsymbol{\pi} \in \Pi_o^p(\mu^{(A)}, \mu^{(B)})$ be an optimal transport plan between $\mu^{(A)}$ and $\mu^{(B)}$ for the p-Wasserstein distance. By disintegrating $\boldsymbol{\pi}$ w.r.t. $\mathrm{pr}_1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, defined by $\mathrm{pr}_1(x,y) = x$, we have a Borel collection of measures $\{\pi_x\}_{x\in\mathbb{R}^d} \subseteq \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, uniquely defined for $\mu^{(A)}$-a.e. $x \in \mathbb{R}^d$, such that $\boldsymbol{\pi} = \mu^{(A)} \otimes \pi_x$.

According to Proposition 2, there exists $\boldsymbol{\eta^{(A)}} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ concentrated on pairs $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ with $\gamma \in AC([0,T])$, $\dot\gamma(t) \in F(e_t\sharp\boldsymbol{\eta^{(A)}}, \gamma(t))$ for a.e. $t \in [0,T]$ and $\gamma(0) = x$ such that $\mu_t^{(A)} = e_t\sharp\boldsymbol{\eta^{(A)}}$.

Let $\boldsymbol{\theta} \in C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$, and define the set-valued map $S^{\boldsymbol{\theta}}(\cdot)$ as in Definition 3.4. Define the set-valued map $R^{\boldsymbol{\theta}} : \mathbb{R}^d \times \operatorname{supp} \boldsymbol{\eta}^{(\boldsymbol{A})} \rightrightarrows \Gamma_T$ by

$$R^{\boldsymbol{\theta}}(y,x,\gamma) := \left\{ \xi \in S^{\boldsymbol{\theta}}(y) : \begin{array}{l} |\gamma(t)-\xi(t)| \le e^{LT}|\gamma(0)-\xi(0)| + L(e^{LT}+1)\int_0^t W_p(\mu_\tau^{(A)},\theta_\tau)\,d\tau \\ \text{for all } t\in[0,T] \end{array} \right\}.$$

Notice that this map has closed domain, closed graph, and compact values since $R^{\boldsymbol{\theta}}(y,x,\gamma) \subseteq S^{\boldsymbol{\theta}}(y)$, thus it is upper semicontinuous, hence Borel measurable.

We prove that it has nonempty images. Given a point $(y,x,\gamma) \in \mathbb{R}^d \times \operatorname{supp} \boldsymbol{\eta}^{(\boldsymbol{A})}$, there are sequences $\{x_n\}_{n\in\mathbb{N}}$ converging to x and $\{\gamma_n\}_{n\in\mathbb{N}} \subseteq AC([0,T])$ uniformly converging to γ such that $x_n = \gamma_n(0)$ and $\dot\gamma_n(t) \in F(\mu_t^{(A)},\gamma_n(t))$ for a.e. $t\in[0,T]$. According to Filippov's theorem (see [4, Theorem 10.4.1]), for every $n \in \mathbb{N}$ there exists $\xi_n \in S^{\boldsymbol{\theta}}(y)$ such that

$$\begin{aligned} |\gamma_n(t)-\xi_n(t)| \le & e^{LT}|\gamma_n(0)-\xi_n(0)| + (Le^{LT}+1)\int_0^t d_{F(\theta_t,\gamma_n(\tau))}(\dot\gamma_n(\tau))\,d\tau \\ \le & e^{LT}|\gamma_n(0)-\xi_n(0)| + L(Le^{LT}+1)\int_0^t W_p(\theta_t,\mu_t^{(A)})\,d\tau, \end{aligned}$$

recalling the Lipschitz continuity of $F(\cdot)$ and the choice of γ_n. By compactness of $S^{\boldsymbol{\theta}}(y)$, up to passing to a subsequence, we may assume that $\{\xi_n\}_{n\in\mathbb{N}}$ uniformly converges to $\xi \in S^{\boldsymbol{\theta}}(y)$ and, by construction, we have $\xi \in R^{\boldsymbol{\theta}}(y,x,\gamma)$, hence $R^{\boldsymbol{\theta}}(\cdot)$ is Borel measurable with closed domain and nonempty images, thus it admits a Borel selection $h_{\boldsymbol{\theta}} : \mathbb{R}^d \times \operatorname{supp} \boldsymbol{\eta}^{(\boldsymbol{A})} \to \Gamma_T$. We extend $h_{\boldsymbol{\theta}}(\cdot)$ to a Borel map defined on the whole of $\mathbb{R}^d \times \mathbb{R}^d \times \Gamma_T \to \Gamma_T$ by setting $h_{\boldsymbol{\theta}}(y,x,\gamma) = \gamma$ if $(x,\gamma) \notin \operatorname{supp} \boldsymbol{\eta}^{(\boldsymbol{A})}$.

Define $\boldsymbol{\eta}^{\boldsymbol{\theta}} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$ by

$$\int_{\mathbb{R}^d\times\Gamma_T} \varphi(y,\xi)\,d\boldsymbol{\eta}^{\boldsymbol{\theta}}(y,\xi) = \int_{\mathbb{R}^d\times\Gamma_T} \left[\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x,h_{\boldsymbol{\theta}}(y,x,\gamma))\,d\pi_x(x,y)\right] d\boldsymbol{\eta}^{(\boldsymbol{A})}(x,\gamma),$$

and set $\boldsymbol{\mu}^{\boldsymbol{\theta}} = \{\mu_t^\theta\}_{t\in[0,T]}$ where $\mu_t^\theta = e_t\sharp\boldsymbol{\eta}^{\boldsymbol{\theta}}$ for all $t\in[0,T]$.

We have, by construction,

$$\operatorname{supp} \boldsymbol{\eta}^{\boldsymbol{\theta}} \subseteq \left\{(y,\xi) \in \mathbb{R}^d \times \Gamma_T : \xi \in S^{\boldsymbol{\theta}}(y)\right\}.$$

Notice that

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x)\,d\mu_0^\theta(x) &= \int_{\mathbb{R}^d\times\Gamma_T}\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi\left(h_{\boldsymbol{\theta}}(y,x,\gamma)(0)\right)\,d\pi_x(x,y)\,d\boldsymbol{\eta}^{(\boldsymbol{A})}(x,\gamma) \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(y)\,d\pi_x(x,y)\,d\mu^{(A)}(x) \\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(y)\,d\boldsymbol{\pi}(x,y) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(y)\,d\mu^{(B)}(y). \end{aligned}$$

Thus $\boldsymbol{\mu}^{\boldsymbol{\theta}} \in \Upsilon_F(\mu^{(B)},\boldsymbol{\theta})$, where $\Upsilon_F(\cdot,\cdot)$ is as in Definition 3.4.

We have

$$\begin{aligned} W_p(\mu_t^{(A)},\mu_t^\theta) &\le \left(\int_{\mathbb{R}^d\times\Gamma_T}\int_{\mathbb{R}^d\times\mathbb{R}^d} |\gamma(t)-h_{\boldsymbol{\theta}}(y,x,\gamma)(t)|^p\,d\pi_x(x,y)\,d\boldsymbol{\eta}^{(\boldsymbol{A})}(x,\gamma)\right)^{1/p} \\ &\le \left(\int_{\mathbb{R}^d\times\Gamma_T}\int_{\mathbb{R}^d\times\mathbb{R}^d} \left[e^{LT}|x-y| + L(Le^{LT}+1)\int_0^t W_p(\mu_\tau^{(A)},\theta_\tau)\,d\tau\right]^p d\pi_x(x,y)\,d\boldsymbol{\eta}^{(\boldsymbol{A})}(x,\gamma)\right)^{1/p} \\ &\le 2^{\frac{p-1}{p}}\left[e^{LT}W_p(\mu^{(A)},\mu^{(B)}) + L(Le^{LT}+1)\int_0^t W_p(\mu_\tau^{(A)},\theta_\tau)\,d\tau\right]. \end{aligned}$$

Thus, since $W_p(\mu_t^{(A)},\mu_t^\theta) \geq W_p(\delta_0,\mu_t^\theta) - W_p(\mu_t^{(A)},\delta_0) = \mathrm{m}_p^{1/p}(\mu_t^\theta) - \mathrm{m}_p^{1/p}(\mu_t^{(A)})$, we have

$$\begin{aligned}
&\mathrm{m}_p^{1/p}(\mu_t^\theta)\\
&\leq \mathrm{m}_p^{1/p}(\mu_t^{(A)}) + 2^{\frac{p-1}{p}}\left[e^{LT}W_p(\mu^{(A)},\mu^{(B)}) + LD\int_0^t \mathrm{m}_p^{1/p}(\mu_\tau^{(A)})\,d\tau + LD\int_0^t \mathrm{m}_p^{1/p}(\theta_\tau)\,d\tau\right]\\
&\leq (1+LTD)\sup_{t\in[0,T]}\mathrm{m}_p^{1/p}(\mu_t^{(A)}) + 2^{\frac{p-1}{p}}\left[e^{LT}W_p(\mu^{(A)},\mu^{(B)}) + LD\int_0^t \mathrm{m}_p^{1/p}(\theta_\tau)\,d\tau\right],
\end{aligned}$$

where we denoted with $D = Le^{LT}+1$. As in the proof of Corollary 1, without loss of generality we can assume that $0 \leq 2^{\frac{p-1}{p}}LDT < 1$. The general case will follow by concatenating finitely many pieces of admissible curves defined on time-subintervals of sufficiently small length. We take $R>0$ sufficiently large such that

$$R \geq \frac{(1+LTD)\sup\limits_{t\in[0,T]}\mathrm{m}_p^{1/p}(\mu_t^{(A)}) + 2^{\frac{p-1}{p}}e^{LT}W_p(\mu^{(A)},\mu^{(B)})}{1-2^{\frac{p-1}{p}}LDT} \geq \mathrm{m}_p^{1/p}(\mu^{(B)}),$$

and such that $\mathrm{m}_p^{1/p}(\theta_t) \leq R$ for all $t\in[0,T]$ and also $\mathrm{m}_p^{1/p}(\mu_t^\theta) \leq R$ for all $t\in[0,T]$. Define a sequence $\{\boldsymbol{\mu}^{(n)} = \{\mu_t^{(n)}\}_{t\in[0,T]}\}_{n\in\mathbb{N}} \subseteq C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ by setting $\boldsymbol{\mu}^{(0)}$ to be the constant $\mu^{(B)}$ and $\boldsymbol{\mu}^{(n)}$ to be equal to $\boldsymbol{\mu}^{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} = \boldsymbol{\mu}^{(n-1)}$. Notice that $\mu_0^{(n)} = \mu^{(B)}$ for all $n\in\mathbb{N}$. According to Lemma 3.5, the family $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ is relatively compact, thus up to passing to a subsequence, we may assume that it converges to $\boldsymbol{\mu}^\infty \in C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$. Since $\boldsymbol{\mu}^{(n)} \in \Upsilon_F(\mu^{(B)},\boldsymbol{\mu}^{(n-1)})$, by recalling the u.s.c. of $\Upsilon_F(\cdot,\cdot)$ proved in Proposition 1, we obtain that $\boldsymbol{\mu}^\infty \in \Upsilon_F(\mu^{(B)},\boldsymbol{\mu}^\infty)$, i.e., $\boldsymbol{\mu}^{\boldsymbol{\infty}}$ is an admissible trajectory, starting from $\mu^{(B)}$. Finally, by passing to the limit in

$$W_p(\mu_t^{(A)},\mu_t^{(n)}) \leq 2^{\frac{p-1}{p}}\left[e^{LT}W_p(\mu^{(A)},\mu^{(B)}) + LD\int_0^t W_p(\mu_\tau^{(A)},\mu_\tau^{(n-1)})\,d\tau\right],$$

we have

$$W_p(\mu_t^{(A)},\mu_t^\infty) \leq 2^{\frac{p-1}{p}}\left[e^{LT}W_p(\mu^{(A)},\mu^{(B)}) + LD\int_0^t W_p(\mu_\tau^{(A)},\mu_\tau^\infty)\,d\tau\right],$$

and, by Grönwall's Lemma,

$$W_p(\mu_t^{(A)},\mu_t^\infty) \leq 2^{\frac{p-1}{p}}e^{\hat{D}T}W_p(\mu^{(A)},\mu^{(B)}),$$

as desired, where $\hat{D} = L(2+Le^{LT})$. The proof is concluded by setting $\boldsymbol{\mu^{(B)}} = \boldsymbol{\mu^\infty}$. The last assertion trivially follows. □

Lemma 3.7 (Initial velocity set). *Assume Hypothesis 3.3 for F and let $\Xi(\cdot,\cdot)$ be as in Definition 3.4. Let $\mu\in\mathcal{P}_p(\mathbb{R}^d)$.*

1. *Given any Borel selection $v_\mu:\mathbb{R}^d\to\mathbb{R}^d$ of $F(\mu,\cdot)$, there exists $\boldsymbol{\eta}\in\mathcal{P}(\mathbb{R}^d\times\Gamma_T)$ such that, set $\mu_t = e_t\sharp\boldsymbol{\eta}$ for $t\in[0,T]$, we have $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$, $\boldsymbol{\eta}\in\Xi(\mu,\boldsymbol{\mu})$ and*

$$\left|\frac{\gamma(t)-\gamma(0)}{t} - v_\mu(x)\right| \leq Le^{Lt}\left[\frac{1}{t}\int_0^t W_p(\mu_\tau,\mu)\,d\tau + \frac{t}{2}|v_\mu(x)|\right]$$

for $\boldsymbol{\eta}$-a.e. $(x,\gamma)\in\mathbb{R}^d\times\Gamma_T$.

2. *Given any admissible trajectory* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$, *there exists* $\boldsymbol{\eta} \in \Xi(\mu, \boldsymbol{\mu})$ *such that* $\mu_t = e_t\sharp\boldsymbol{\eta}$ *and for* $\boldsymbol{\eta}$*-a.e.* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ *we have*

$$\lim_{t\to 0^+} \int_{\mathbb{R}^d\times\Gamma_T} d^p_{F(\mu,x)}\left(\frac{\gamma(t)-\gamma(0)}{t}\right) d\boldsymbol{\eta}(x,\gamma) = 0.$$

Proof.

1. Without loss of generality, we may assume $LT \le 1/2$, the general case will be obtained concatenating $\boldsymbol{\mu}$ with any other admissible trajectory starting from μ_T. Let $v_0 : \mathbb{R}^d \to \mathbb{R}^d$ be any Borel selection of $F(\mu,\cdot)$. Define $\gamma_x : [0,T] \to \mathbb{R}^d$ by $\gamma_x(t) = x + v_0(x)\cdot t$, and observe that $x \mapsto \gamma_x$ is a Borel map. Let $\boldsymbol{\theta} = \{\theta_t\}_{t\in[0,T]} \in C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ such that $\theta_0 = \mu$, and notice that

$$d_{F(\theta_t,\gamma_x(t))}(\dot\gamma_x(t)) \le L\left[W_p(\theta_t,\mu) + t|v_0(x)|\right].$$

Thus, by Filippov's Theorem (see [4, Theorem 10.4.1]) the set-valued map $R^{\boldsymbol{\theta}} : \mathbb{R}^d \rightrightarrows \Gamma_T$ defined as

$$R^{\boldsymbol{\theta}}(x) := \left\{\xi \in S^{\boldsymbol{\theta}}(x) : \begin{array}{l} |\gamma_x(t)-\xi(t)| \le Le^{Lt}\int_0^t \left[W_p(\theta_\tau,\mu) + \tau|v_0(x)|\right] d\tau \\ \text{for all } t\in[0,T] \end{array}\right\}$$

has nonempty images for every $x \in \mathbb{R}^d$. Notice that this set-valued map has closed images and it is Borel measurable by [4, Theorem 8.2.9], thus it admits a Borel selection $h_{\boldsymbol{\theta}} : \mathbb{R}^d \to \Gamma_T$. Set $\boldsymbol{\eta}^{\boldsymbol{\theta}} = \mu \otimes \delta_{h_{\boldsymbol{\theta}}(x)}$ and $\boldsymbol{\mu}^{\boldsymbol{\theta}} = \{\mu^\theta_t\}_{t\in[0,T]}$, $\mu^\theta_t = e_t\sharp\boldsymbol{\eta}^{\boldsymbol{\theta}}$. By construction we have $\boldsymbol{\mu}^{\boldsymbol{\theta}} \in \Upsilon_F(\mu,\boldsymbol{\theta})$, moreover for all $x \in \mathbb{R}^d$ we have

$$\begin{aligned}
\mathrm{m}_p^{1/p}(\mu^\theta_t) &= \left(\int_{\mathbb{R}^d} |h_{\boldsymbol{\theta}}(x)(t)|^p \, d\mu(x)\right)^{1/p} \\
&\le \left(\int_{\mathbb{R}^d} |h_{\boldsymbol{\theta}}(x)(t) - \gamma_x(t)|^p \, d\mu(x)\right)^{1/p} + \left(\int_{\mathbb{R}^d} |\gamma_x(t)|^p \, d\mu(x)\right)^{1/p} \\
&\le Le^{Lt}\left[\int_{\mathbb{R}^d}\left|\int_0^t (W_p(\theta_\tau,\mu) + \tau|v_0(x)|)\, d\tau\right|^p d\mu\right]^{1/p} + \mathrm{m}_p^{1/p}(\mu) + t\|v_0\|_{L^p_\mu} \\
&\le Le^{Lt}\int_0^t W_p(\theta_\tau,\mu)\, d\tau + \mathrm{m}_p^{1/p}(\mu) + (LTe^{Lt}+t)\|v_0\|_{L^p_\mu} \\
&\le Le^{Lt}\int_0^t \mathrm{m}_p^{1/p}(\theta_\tau)\, d\tau + (1+Le^{Lt}t)\mathrm{m}_p^{1/p}(\mu) + (LTe^{Lt}+t)\|v_0\|_{L^p_\mu}.
\end{aligned}$$

Furthermore,

$$\left|\frac{h_{\boldsymbol{\theta}}(x)(t) - h_{\boldsymbol{\theta}}(0)}{t} - v_0(x)\right| = \left|\frac{h_{\boldsymbol{\theta}}(x)(t) - \gamma_x(t)}{t}\right| \le Le^{Lt}\frac{1}{t}\int_0^t \left[W_p(\theta_\tau,\mu) + \tau|v_0(x)|\right] d\tau.$$

Choose

$$R \ge \frac{(1+LTe^{LT})\mathrm{m}_p^{1/p}(\mu) + (LTe^{LT}+T)\|v_0\|_{L^p_\mu}}{1-LTe^{LT}} \ge \mathrm{m}_p^{1/p}(\mu),$$

and notice that if $\mathrm{m}_p^{1/p}(\theta_t) \le R$ for all $t \in [0,T]$, then $\mathrm{m}_p^{1/p}(\mu^\theta_t) \le R$ for all $t \in [0,T]$. Define sequences $\{\boldsymbol{\mu}^{(n)} = \{\mu^{(n)}_t\}_{t\in[0,T]}\}_{n\in\mathbb{N}} \subseteq C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ and $\{\boldsymbol{\eta}^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathcal{P}(\mathbb{R}^d\times\Gamma_T)$ by setting $\boldsymbol{\mu}^{(0)}$ to be the constant μ, $\boldsymbol{\mu}^{(n)}$ and $\boldsymbol{\eta}^{(n)}$ to be equal to $\boldsymbol{\mu}^{\boldsymbol{\theta}}$ and $\boldsymbol{\eta}^{\boldsymbol{\theta}}$, respectively, with $\boldsymbol{\theta} = \boldsymbol{\mu}^{(n-1)}$ for all $n \in \mathbb{N}$. According to Lemma 3.5, the families $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ and $\{\boldsymbol{\eta}^{(n)}\}_{n\in\mathbb{N}}$ are relatively compact, thus up to passing to a subsequence, we may assume that the sequences converge to

$\boldsymbol{\mu}^\infty = \{\mu_t^\infty\}_{t\in[0,T]} \in C^0([0,T];\mathcal{P}_p(\mathbb{R}^d))$ and to $\boldsymbol{\eta^\infty} \in \mathcal{P}(\mathbb{R}^d \times \Gamma_T)$, with $\mu_t^\infty = e_t\sharp\boldsymbol{\eta^\infty}$ for all $t \in [0,T]$. Since $\boldsymbol{\mu}^{(n)} \in \Upsilon_F(\mu, \boldsymbol{\mu}^{(n-1)})$ ($\Upsilon_F(\cdot,\cdot)$ defined in Definition 3.4), by recalling the u.s.c. of $\Upsilon_F(\cdot,\cdot)$ proved in Proposition 1, we obtain that $\boldsymbol{\mu}^\infty \in \Upsilon_F(\mu, \boldsymbol{\mu}^\infty)$, i.e., $\boldsymbol{\mu^\infty}$ is an admissible trajectory, starting from μ. Recall that for $\boldsymbol{\eta^\infty}$-a.e. (x,γ) there exists a sequence $\{(x_n,\xi_n)\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d \times \Gamma_T$ converging to (x,γ) such that $(x_n,\xi_n) \in \operatorname{supp}\boldsymbol{\eta^{(n)}}$. Thus, without loss of generality, we may assume for all $t \in [0,T]$

$$\left|\frac{\xi_n(t)-\xi_n(0)}{t} - v_0(x)\right| \le Le^{Lt}\frac{1}{t}\int_0^t \left[W_p(\mu_\tau^{(n-1)},\mu) + \tau|v_0(x)|\right]\,d\tau,$$

and, by passing to the limit,

$$\left|\frac{\gamma(t)-\gamma(0)}{t} - v_0(x)\right| \le Le^{Lt}\frac{1}{t}\int_0^t \left[W_p(\mu_\tau^{\infty},\mu) + \tau|v_0(x)|\right]\,d\tau.$$

2. Recall the Superposition Principle in Proposition 2 and the definitions of Υ_F and Ξ in Definition 3.4. Then, from the assumption it follows that there exists $\boldsymbol{\eta} \in \Xi(\mu,\boldsymbol{\mu})$ such that $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0,T]$. By Jensen's inequality, for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$, we have

$$\begin{aligned}
d_{F(\mu,x)}\left(\frac{\gamma(t)-\gamma(0)}{t}\right) =& d_{F(\mu,x)}\left(\frac{1}{t}\int_0^t \dot\gamma(s)\,ds\right) \le \frac{1}{t}\int_0^t d_{F(\mu,x)}\left(\dot\gamma(s)\right)\,ds\\
\le& \frac{L}{t}\int_0^t \left(W_p(\mu_s,\mu) + |\gamma(s)-x|\right)\,ds\\
\le& \frac{L}{t}\int_0^t \left(W_p(\mu_s,\mu) + \|\dot\gamma\|_{L^\infty_{\boldsymbol{\eta}}} s\right)\,ds.
\end{aligned}$$

We conclude by taking the $L^p_{\boldsymbol{\eta}}$-norm and using Lemma 3.5. □

4. **Generalized targets.** In this section, we provide the generalized notion of target set in the space of probability measures, thus extending in a natural way the classical concept of target set in $\mathbb{R}^d$. A naive physical interpretation of the generalized target can be given as follows: to describe the state of the system, an observer chooses to measure some quantities ϕ. The results of the measurements are the *averages* of the quantities ϕ with respect to the measure μ_t, representing the state of the system at time t. Our aim is to steer the system to states where the result of such measurements is below a fixed threshold (that, without loss of generality, we assume to be 0). The following result provides a characterization of the class of such generalized target.

Lemma 4.1. *Let $\tilde S \subseteq \mathcal{P}(\mathbb{R}^d)$ be nonempty. Then, $\tilde S$ is w^*-closed and convex if and only if there exists a family $\Phi \subseteq C^0_b(\mathbb{R}^d)$ such that $\tilde S$ can be written as follows*

$$\tilde S = \left\{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} \varphi(x)\,d\mu(x) \le 0 \text{ for all } \varphi \in \Phi\right\}. \tag{3}$$

Proof. We first prove the necessity, so let $\tilde S$ be as in (3) for some fixed $\Phi \subseteq C^0_b(\mathbb{R}^d)$. Then, the convexity of $\tilde S$ comes by linearity of the integral w.r.t. the measure, while the closure in w^* topology follows immediately since Φ is a family of test functions for w^*-convergence.

We pass to the proof of the sufficiency. Recalling formula (5.1.7) in [3, Remark 5.1.2], we have that $\bar{\mu} \in \tilde{S}$ if and only if for all $\psi \in C_b^0(\mathbb{R}^d)$ it holds

$$\int_{\mathbb{R}^d} \psi(x)\, d\bar{\mu}(x) \leq \sup_{\mu \in \tilde{S}} \int_{\mathbb{R}^d} \psi(x)\, d\mu(x).$$

Given $\psi \in C_b^0(\mathbb{R}^d)$, set

$$C_\psi := \sup_{\mu \in \tilde{S}} \int_{\mathbb{R}^d} \psi(x)\, d\mu(x) \leq +\infty.$$

Then we have that $\bar{\mu} \in \tilde{S}$ if and only if for all $\psi \in C_b^0(\mathbb{R}^d)$ such that $C_\psi < +\infty$ it holds

$$\int_{\mathbb{R}^d} [\psi(x) - C_\psi]\, d\bar{\mu}(x) \leq 0.$$

Then, to get (3) it sufficies to take

$$\Phi := \left\{\varphi := \psi - C_\psi : \psi \in C_b^0(\mathbb{R}^d) \text{ and } C_\psi < +\infty\right\}.$$

□

Definition 4.2 (Generalized targets)**.** Let $\tilde{S} \subseteq \mathcal{P}(\mathbb{R}^d)$ be nonempty w^*-closed and convex, $\Phi \subseteq C_b^0(\mathbb{R}^d)$. We say that $\tilde{S}$ is a *generalized target generated by* Φ, and write $\tilde{S} = \tilde{S}^\Phi$ if

$$\tilde{S} := \left\{\mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} \varphi(x)\, d\mu(x) \leq 0 \text{ for all } \varphi \in \Phi\right\}. \tag{4}$$

Given $p \geq 1$ we set $\tilde{S}_p^\Phi = \tilde{S}^\Phi \cap \mathcal{P}_p(\mathbb{R}^d)$, and we define the *generalized distance* from $\tilde{S}_p^\Phi$ to be the 1-Lipschitz continuous map given by $\tilde{d}_{\tilde{S}_p^\Phi}(\cdot) := \inf_{\mu \in \tilde{S}_p^\Phi} W_p(\cdot, \mu)$.

Remark 4.

- In Definition 4.2 we can equivalently assume that Φ is a set of continuous bounded functions, or bounded Lipschitz functions, or even just l.s.c. functions bounded from below. Moreover, without loss of generality, we can always assume that Φ is convex. Indeed, assume that Ψ is a set of l.s.c. functions bounded from below. For all $\psi \in \Psi$ and $k \in \mathbb{N} \setminus \{0\}$ we define a Lipschitz continuous bounded map $\varphi_k^\psi : \mathbb{R}^d \to \mathbb{R}$ by setting

  $$\varphi_k^\psi(x) := \min\left\{\inf_{y \in \mathbb{R}^d} \{\psi(y) + k|x-y|\}, k\right\}.$$

  We recall that $\{\varphi_k^\psi\}_{k \in \mathbb{N}}$ is an increasing sequence of bounded Lipschitz functions bounded from below and pointwise converging to ψ. Hence, by Monotone Convergence Theorem, we have

  $$\begin{aligned}\sup_{\psi \in \Psi} \int_{\mathbb{R}^d} \psi(x)\, d\mu(x) &= \sup_{\psi \in \Psi} \int_{\mathbb{R}^d} \sup_{k \in \mathbb{N}} \varphi_k^\psi(x)\, d\mu(x) = \sup_{k \in \mathbb{N}} \sup_{\psi \in \Psi} \int_{\mathbb{R}^d} \varphi_k^\psi(x)\, d\mu(x) \\ &= \sup_{\varphi \in \Phi} \int_{\mathbb{R}^d} \varphi(x)\, d\mu(x),\end{aligned}$$

  where $\Phi = \{\varphi_k^\psi : k \in \mathbb{N} \setminus \{0\},\, \psi \in \Psi\}$. Replacing Φ with its convex hull does not change anything due to the linearity of the integral operator.

- Since convergence in $W_p(\cdot,\cdot)$ implies w^*-convergence, if $\tilde{S}^\Phi$ is a generalized target, then $\tilde{S}^\Phi_p$ is closed and convex in $\mathcal{P}_p(\mathbb{R}^d)$ endowed with the p-Wasserstein metric $W_p(\cdot,\cdot)$.
- We notice that if there exists $\bar{x} \in \mathbb{R}^d$ such that $\varphi(\bar{x}) \le 0$ for all $\varphi \in \Phi$ then the set $\tilde{S}$ given by (4) is nonempty, since $\delta_{\bar{x}} \in \tilde{S}$.

The last condition of Remark 4 is indeed not necessary to have the nontriviality of $\tilde{S}$.

Example 4.3. For every $y \in \mathbb{R}$, $\varepsilon > 0$, define

$$\varphi^\varepsilon_y(x) = \begin{cases} -(x+y)^2 + \varepsilon, & \text{if } |x+y| \le 1, \\ -1+\varepsilon, & \text{if } |x+y| \ge 1. \end{cases}$$

and set $\Phi_\varepsilon := \{\varphi^\varepsilon_y : y \in \mathbb{R}\}$. Clearly, we have that φ^ε_y attains its maximum at $x = -y$ and the value of the maximum is $\varepsilon > 0$. Thus the sufficient condition of the last assertion in Remark 4 is violated. For $0 < \varepsilon \le \dfrac{1}{12}$ sufficiently small we have

$$\int_{-1/2}^{1/2} \varphi^\varepsilon_y(x)\,dx \le \int_{-1/2}^{1/2} \varphi^\varepsilon_0(x)\,dx = \varepsilon - \frac{1}{12} \le 0,$$

thus the measure $\chi_{[-1/2,1/2]}\mathcal{L}^1 \in \tilde{S}$.

Indeed, by the translation invariance of the problem, we have that $\mu_a := \chi_{[a,a+1]}\mathcal{L}^1 \in \tilde{S}$ for all $a \in \mathbb{R}$, in particular, we have that $\tilde{S}$ is not tight, hence not w^*-compact, since for any $K \subseteq \mathbb{R}$ it is possible to find $a \in \mathbb{R}$ such that $\mu_a(\mathbb{R} \setminus K) = 1$.

Lemma 4.4 (Compactness)**.** *Let $\tilde{S}$ be a nonempty generalized target generated by the family $\Phi \subseteq C^0(\mathbb{R}^d)$. If there exists $\bar{\phi} \in \Phi$, $A, C > 0$ and $p \ge 1$ such that $\bar{\phi}(x) \ge A|x|^p - C$, then $\tilde{S}^\Phi = \tilde{S}^\Phi_p$ is compact in the w^*-topology and in the W_p-topology.*

Proof. Trivially we have that $\tilde{S}^\Phi_p \subseteq \tilde{S}^\Phi$ for any $p \ge 1$. Conversely, given $\mu \in \tilde{S}^\Phi$, we have

$$A \cdot \mathrm{m}_p(\mu) - C \le \int_{\mathbb{R}^d} \bar{\phi}(x)\,d\mu \le 0,$$

hence $\mu \in \tilde{S}^\Phi_p$ and all the measures in $\tilde{S}^\Phi_p = \tilde{S}^\Phi$ have p-moments uniformly bounded by C/A. This means that the w^*-topology and W_p-topology coincide on $\tilde{S}^\Phi = \tilde{S}^\Phi_p$, which turns out to be tight, according to [3, Remark 5.1.5], and w^*-closed, hence w^*-compact and W_p-compact. □

We mention the following example, which may be relevant for the applications.

Example 4.5. Given a nonempty and closed set $S \subseteq \mathbb{R}^d$ and $\alpha \ge 0$, a natural choice for Φ can be for example $\Phi_\alpha = \{d_S(\cdot) - \alpha\}$. If $\alpha = 0$ we have that $\tilde{S}^{\Phi_0} = \{\mu \in \mathcal{P}(\mathbb{R}^d) : \mu(\mathbb{R}^d \setminus S) = 0\}$. More generally, for all $r > 0$ let $B_r(S) := \{z \in \mathbb{R}^d : d_S(z) \le r\}$. Then, if $\mu \in \tilde{S}^{\Phi_\alpha}$,

$$r\mu(\mathbb{R}^d \setminus B_r(S)) = \int_{\mathbb{R}^d \setminus B_r(S)} r\,d\mu \le \int_{\mathbb{R}^d \setminus B_r(S)} d_S(x)\,d\mu(x) \le \alpha,$$

thus, in particular, we must have $\mu(\mathbb{R}^d \setminus B_r(S)) \le \min\left\{1, \dfrac{\alpha}{r}\right\}$ for all $r > 0$, which, if α is sufficiently small can be interpreted as a relaxed version of the case $\alpha = 0$.

Given a generalized target $\tilde{S} \subseteq \mathcal{P}(\mathbb{R}^d)$, a natural question is wheter it is possible to *localize* it, i.e., to describe it as the set of all the measures supported a certain (closed) subset of $\mathbb{R}^d$. Equivalently, we want to find a nonempty closed set $S \subseteq \mathbb{R}^d$, such that, set $\Phi = \{d_S(\cdot)\}$, we have $\tilde{S} = \tilde{S}^\Phi$. To this aim, we give the following definition.

Definition 4.6 (Classical counterpart of generalized target)**.** Let $\tilde{S} \subseteq \mathcal{P}(\mathbb{R}^d)$ be a generalized target. Given a set $S \subseteq \mathbb{R}^d$, we say that S is a *classical counterpart of the generalized target* $\tilde{S}$ if

$$\tilde{S} = \{\mu \in \mathcal{P}(\mathbb{R}^d) : \operatorname{supp} \mu \subseteq S\}.$$

An analogous definition is given for the classical counterpart of $\tilde{S} \cap \mathcal{P}_p(\mathbb{R}^d)$, $p \geq 1$ by taking intersection of the right hand side with $\mathcal{P}_p(\mathbb{R}^d)$.

Remark 5.

- From the very definition of classical counterpart, if $\tilde{S}$ admits S and S' as classical counterparts, then $S = S'$.
- In general a classical counterpart may not exists: in $\mathbb{R}$, take $\Phi = \{\phi\}$ where $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(y) := |y| - 1$. Defined $\mu_0 := \dfrac{1}{2}(\delta_0 + \delta_2)$, we have $\mu_0 \in \tilde{S}_p^\Phi$ for every $p \geq 1$. If a classical counterpart S of $\tilde{S}^\Phi$ would exists, by definition it should contain the support of μ_0, i.e. $0, 2 \in S$. However, $\delta_2 \notin \tilde{S}^\Phi$ even if $\operatorname{supp}(\delta_2) \subseteq S$. So neither $\tilde{S}^\Phi$ nor $\tilde{S}_p^\Phi$ admit a classical counterpart.
- If S is the classical counterpart of $\tilde{S}^\Phi$ (or $\tilde{S}_p^\Phi$), there exists a representation of $\tilde{S}^\Phi$ as $\tilde{S}^{\hat{\Phi}}$, where $\hat{\Phi} = \{\hat{\phi}\}$ and $\hat{\phi}(x) \geq 0$ for every $x \in \mathbb{R}^d$ where the inequality is strict at every $x \notin S$. In particular we can take $\hat{\Phi} = \{\arctan \circ d_S\}$ (resp. $\hat{\Phi} = \{d_S\}$), i.e., we can replace Φ with the set $\{\arctan \circ d_S\}$ (resp. $\{d_S\}$).

Our aim is now to characterize the generalized target possessing a classical counterpart.

Proposition 3 (Existence, uniqueness and properties of the classical counterpart)**.** *Let $\tilde{S} \subseteq \mathcal{P}(\mathbb{R}^d)$ be a generalized target, $S \subseteq \mathbb{R}^d$.*

1. *if $\tilde{S}$ admits S as classical counterpart then S is closed;*
2. *$\tilde{S}$ admits S as classical counterpart if and only if*
$$\int_{\mathbb{R}^d} \left[\varphi(x) - \sup_{y \in S} \varphi(y)\right] d\mu(x) \leq 0,$$
for all $\varphi \in C_b^0(\mathbb{R}^d)$ and $\mu \in \tilde{S}$;
3. *if $\tilde{S}$ admits S as classical counterpart, then $\tilde{S}_p$ admits S as classical counterpart for all $p \geq 1$.*
4. *If $\tilde{S} = \tilde{S}^\Phi$ (resp. $\tilde{S} \cap \mathcal{P}_p(\mathbb{R}^d) = \tilde{S}_p^\Phi$), for a suitable $\Phi \subseteq C_b^0(\mathbb{R}^d)$, admits a classical counterpart S, then*
$$S = \bigcap_{\phi \in \Phi} \{x \in \mathbb{R}^d : \phi(x) \leq 0\}.$$

Proof.

1. Assume that $\tilde{S}$ admits S as a classical counterpart and $\tilde{S} = \tilde{S}^\Phi$ for a suitable $\Phi \in C_b^0(\mathbb{R}^d)$. In particular, we have $\delta_x \in \tilde{S}$ for all $x \in S$, i.e. $\phi(x) \leq 0$ for all $x \in S$. Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in S converging to $x \in \mathbb{R}^d$. Then for all $\varphi \in \Phi$ we have $\varphi(x_n) \leq 0$ for all $n \in \mathbb{N}$, which implies $\varphi(x) \leq 0$, and so

$\delta_x \in \tilde{S}$. Since S is a classical counterpart of $\tilde{S}$ and supp$\delta_x = \{x\}$, we have that thus $x \in S$, so S is closed.

2. $\tilde{S}$ admits S as classical counterpart if and only if $\tilde{S} = \overline{\text{co}}\{\delta_x : x \in S\}$, where the closure is the weak* closure in $\mathcal{P}(\mathbb{R}^d)$. Indeed, every measure supported in S is w^*-limit of convex combinations of Dirac deltas concentrated in points of S, and conversely all such deltas belong to $\tilde{S}$ by definition of classical counterpart, and $\tilde{S}$ is convex and w^*-closed. Recalling formula (5.1.7) in [3, Remark 5.1.2], we have that $\mu \in \tilde{S}$ if and only if
$$\int_{\mathbb{R}^d} \varphi(x)\, d\mu(x) \le \sup_{y\in S} \varphi(y),$$
for all $\varphi \in C_b^0$, as desired.
3. It is sufficient to use the same argument as in (2) but taking the intersection with $\mathcal{P}_p(\mathbb{R}^d)$ and the closure w.r.t. W_p distance.
4. Trivially, if there exist $\bar{x} \in \mathbb{R}^d$ and $\varphi \in \Phi$ such that $\varphi(\bar{x}) > 0$, then $\delta_{\bar{x}} \notin \tilde{S}$, thus $\bar{x}$ does not belong to the classical counterpart of $\tilde{S}$. Conversely, if $\varphi(\bar{x}) \le 0$ for all $\varphi \in \Phi$, then $\delta_{\bar{x}} \in \tilde{S}$, and so $\bar{x} \in S$ by definition of classical counterpart. □

A useful sufficient condition can be expressed as follows.

Corollary 2. *Assume that for every $\phi \in \Phi$ we have either $\phi(x) \ge 0$ or $\phi(x) \le 0$ for all $x \in \mathbb{R}^d$. Then $\tilde{S}^\Phi$ (and so $\tilde{S}^\Phi_p$) admits classical counterpart.*

Proof. Denote by
$$S = \bigcap_{\phi\in\Phi} \{x \in \mathbb{R}^d : \phi(x) \le 0\}.$$
If for all $\phi \in \Phi$ and $x \in \mathbb{R}^d$ we had $\phi(x) \le 0$, then we would trivially have $S = \mathbb{R}^d$ and $\tilde{S}^\Phi = \mathcal{P}(\mathbb{R}^d)$ as desired since $\delta_x \in \tilde{S}^\Phi$ for all $x \in \mathbb{R}^d$, thus concluding with the thesis.

Otherwise, let $\mu \in \tilde{S}^\Phi$ and suppose by contradiction that $\mu(\mathbb{R}^d \setminus S) > 0$. Thus there exists $y \in \mathbb{R}^d \setminus S$ of density 1 w.r.t. μ. In particular, there exists a neighborhood A_y of y contained in $\mathbb{R}^d \setminus S$ such that $\mu(A_y) > 0$. If for all $\varphi \in \Phi$ we had $\varphi(y) \le 0$, we would have $y \in S$, contradicting the fact that $y \notin S$. So, according to the assumptions, there exists $\hat{\phi} \in \Phi$ such that $\hat{\phi}(x) \ge 0$ for all $x \in \mathbb{R}^d$ and such that $\hat{\phi}(y) > 0$. Thus we have
$$\sup_{\phi\in\Phi} \int_{\mathbb{R}^d} \phi(x)\, d\mu(x) \ge \int_{\mathbb{R}^d} \hat{\phi}(x)\, d\mu(x) \ge \int_{A_y} \hat{\phi}(x)\, d\mu(x) > 0,$$
hence $\mu \notin \tilde{S}^\Phi$, leading to a contradiction. Thus $\tilde{S}^\Phi \subseteq \{\mu \in \mathcal{P}(\mathbb{R}^d) : \operatorname{supp}\mu \subseteq S\}$. Since the converse inclusion is always true, equality holds. □

Remark 6. The condition of Corollary 2 is not necessary in general. In $\mathbb{R}$, take $\Phi = \{\phi_1, \phi_2, \phi_3\}$ where $\phi_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$ are defined to be $\phi_1(x) = \min\{\max\{x, 0\}, 1\}$, $\phi_2(x) = \min\{\max\{-x, -1\}, 0\}$, $\phi_3(x) = \min\{\max\{x, -1\}, 1\}$. Then both $\tilde{S}^\Phi_p$ and $\tilde{S}^\Phi$ admits S as their classical counterpart, with $S =]-\infty, 0]$, but ϕ_3 changes its sign.

We are now ready to state some comparison results between the generalized distance and the classical one.

Proposition 4 (Comparison with classical distance)**.** *Let $p \geq 1$, $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$, $\Phi \subseteq C_b^0(\mathbb{R}^d;\mathbb{R})$ be such that $\tilde{S}_p^\Phi \neq \emptyset$, and define*

$$S := \bigcap_{\phi\in\Phi} \{x \in \mathbb{R}^d : \phi(x) \leq 0\}. \tag{5}$$

Then $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \leq \|d_S\|_{L^p_{\mu_0}}$, and equality holds if and only if the generalized target $\tilde{S}_p^\Phi$ admits classical counterpart. In this last case, the classical counterpart of $\tilde{S}_p^\Phi$ is S, moreover $\tilde{d}^p_{\tilde{S}_p^\Phi} : \mathcal{P}_p(\mathbb{R}^d) \to [0, +\infty[$ is convex.

Proof. If $S = \emptyset$ we have $d_S(x) \equiv +\infty$ at all $x \in \mathbb{R}^d$ so the statement is trivially true, thus suppose $S \neq \emptyset$. Since S is closed and nonempty, [4, Corollary 8.2.13] implies the existence of a Borel map $g : \mathbb{R}^d \to S$ such that $|x - g(x)| = d_S(x)$. We have

$$\mathrm{m}_p^{1/p}(g\sharp\mu_0) = \|g\|_{L^p_{\mu_0}} \leq \|\mathrm{Id}_{\mathbb{R}^d} - g\|_{L^p_{\mu_0}} + \|\mathrm{Id}_{\mathbb{R}^d}\|_{L^p_{\mu_0}} \leq \|d_S\|_{L^p} + \mathrm{m}_p^{1/p}(\mu_0) < +\infty,$$

moreover, for all $\phi \in \Phi$, we have

$$\int_{\mathbb{R}^d} \phi(x)\, dg\sharp\mu_0(x) = \int_{\mathbb{R}^d} \phi(g(y))\, d\mu_0(y) \leq 0,$$

since $g(y) \in S$ for all $y \in \mathbb{R}^d$ and so $\phi \circ g(y) \leq 0$ for all $y \in \mathbb{R}^d$. Therefore, $g\sharp\mu_0 \in \tilde{S}_p^\Phi$, and so

$$\tilde{d}^p_{\tilde{S}^\Phi}(\mu_0) \leq W_p^p(\mu_0, g\sharp\mu_0) \leq \|\mathrm{Id}_{\mathbb{R}^d} - g\|^p_{L^p_{\mu_0}} = \|d_S\|^p_{L^p_{\mu_0}}.$$

Assume now that $\tilde{S}_p^\Phi$ admits classical counterpart. As noticed in Proposition 3, S must be the classical counterpart of $\tilde{S}_p^\Phi$. For every $\nu_0 \in \tilde{S}_p^\Phi$ we have thus $\operatorname{supp}\nu_0 \subseteq S$ and hence $|x - y| \geq d_S(x)$ for all $\pi \in \Pi(\mu_0, \nu_0)$ and π-a.e. $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. This leads to

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} |x - y|^p\, d\pi(x, y) \geq \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_0(x).$$

By taking the infimum on $\pi \in \Pi(\mu_0, \nu_0)$ and then on $\nu_0 \in \tilde{S}_p^\Phi$, we obtain $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \geq \|d_S\|_{L^p_{\mu_0}}$, thus equality holds.

Without the assumption of existence of a classical counterpart for $\tilde{S}_p^\Phi$, the inequality $\tilde{d}_{\tilde{S}_p^\Phi}(\mu_0) \leq \|d_S\|_{L^p_{\mu_0}}$ is strict. Indeed, since $\tilde{S}_p^\Phi$ does not admit S as a classical counterpart, there exist a measure $\mu \in \tilde{S}_p^\Phi$ and $n \in \mathbb{N}$ such that

$$\mu\left(\left\{z \in \mathbb{R}^d : d_S(z) > \frac{1}{n}\right\}\right) > 0,$$

and so there exists a Borel set $A \subseteq \mathbb{R}^d$ and $\varepsilon > 0$ such that $d_S^p(z) \geq \varepsilon$ for μ-a.e. $z \in A$, $\mu(A) > 0$. This implies

$$0 = \tilde{d}^p_{\tilde{S}_p^\Phi}(\mu) < \varepsilon\mu(A) \leq \int_A d_S^p(z)\, d\mu(z) \leq \int_{\mathbb{R}^d} d_S^p(x)\, d\mu(x).$$

Finally, the last statement is trivial, and it follows from the fact that

$$\tilde{d}^p_{\tilde{S}_p^\Phi}(\mu) = \int_{\mathbb{R}^d} d_S^p(x)\, d\mu,$$

is linear in μ. $\square$

Without the p-th power, the generalized distance in the case of the Proposition 4 above may fail to be convex.

Example 4.7. Let $p > 1$. In $\mathbb{R}^2$, consider $P = (0,0)$, $Q_1 = (1,0)$, $Q_2 = \big(0, 2^{1/p}\big)$. Set $S = \{P\}$, $\Phi = \{d_S(\cdot)\}$, hence $\tilde{S}_p^\Phi := \{\delta_P\}$, and define $\nu_\lambda = \lambda\delta_{Q_1} + (1-\lambda)\delta_{Q_2}$, $\lambda \in [0,1]$. By Proposition 4, we have

$$\tilde{d}^p_{\tilde{S}_p^\Phi}(\nu_\lambda) = W_p^p(\delta_P, \nu_\lambda) = \lambda + 2(1-\lambda) = 2 - \lambda,$$

whence $\tilde{d}_{\tilde{S}_p^\Phi}(\nu_\lambda) = \sqrt[p]{2-\lambda}$, which is not convex.

In the metric space $\mathcal{P}_p(\mathbb{R}^d)$ endowed with the W_p-distance, another concept of convexity can be given, related more to the metric structure rather than to the linear one inherited by the set of all Borel signed measures.

Given any product space X^N ($N \geq 1$), in the following we denote with $\mathrm{pr}^i \colon X^N \to X$ the projection on the i–th component, i.e., $\mathrm{pr}^i(x_1, \dots, x_N) = x_i$.

Definition 4.8 (Geodesics)**.** Given a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]} \subseteq \mathcal{P}_p(\mathbb{R}^d)$, we say that it is a *(constant speed) geodesic* if for all $0 \leq s \leq t \leq 1$ we have

$$W_p(\mu_s, \mu_t) = (t-s)W_p(\mu_0, \mu_1).$$

In this case, we will also say that the curve $\boldsymbol{\mu}$ is a *geodesic connecting μ_0 and μ_1.*

Theorem 4.9 (Characterization of geodesics)**.** *Let $\mu_0, \mu_1 \in \mathcal{P}_p(\mathbb{R}^d)$ and let $\pi \in \Pi_o^p(\mu_0, \mu_1)$ be an optimal transport plan between μ_0 and μ_1, i.e.*

$$W_p^p(\mu_0, \mu_1) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x_1 - x_2|^p \, d\pi(x_1, x_2)\,.$$

Then the curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ defined by

$$\mu_t := \big((1-t)\,\mathrm{pr}^1 + t\,\mathrm{pr}^2\big)\sharp\pi \;\in\; \mathcal{P}_p(\mathbb{R}^d), \tag{6}$$

is a (constant speed) geodesic connecting μ_0 and μ_1.

Conversely, any (constant speed) geodesic $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ connecting μ_0 and μ_1 admits the representation (6) for a suitable plan $\pi \in \Pi_o^p(\mu_0, \mu_1)$.

Proof. See [3, Theorem 7.2.2]. □

Definition 4.10 (Geodesically and strongly geodesically convex sets)**.** A subset $A \subseteq \mathcal{P}_p(\mathbb{R}^d)$ is said to be

1. *geodesically convex* if for every pair of measures μ_0, μ_1 in A, there exists a geodesic connecting μ_0 and μ_1 which is contained in A.
2. *strongly geodesically convex* if for every pair of measures μ_0, μ_1 in A and for every admissible transport plan $\pi \in \Pi(\mu_0, \mu_1)$, the curve $t \mapsto \mu_t$ defined by (6) is contained in A.

The interest in this alternative concept of convexity comes from the fact that, in many problems, functionals defined on probability measures are convex along geodesics (a notion related to geodesically convex sets) and not convex with respect to the linear structure in the usual sense. We refer to [3, Section 9.1] for further details.

Remark 7. Notice that, even if the notations do not highlight this fact, the notions of *geodesic* and *geodesical convexity* depend on the exponent p which has been fixed.

Proposition 5 (Strong geodesic convexity of $\tilde{S}_p^\Phi$)**.** *Let $p \geq 1$, Φ as in Definition 4.2 and assume that all the elements of Φ are also convex. Then the generalized target $\tilde{S}_p^\Phi$ is strongly geodesically convex.*

Proof. Let $\mu_0, \mu_1 \in \tilde{S}_p^\Phi$ and let $\pi \in \Pi(\mu_0, \mu_1)$ be an admissible transport plan between μ_0 and μ_1. Consider the corresponding curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ defined by (6), and fix $t \in [0,1]$. We have for every $\phi(\cdot) \in \Phi$

$$\begin{aligned}
&\int_{\mathbb{R}^d} \phi(x)\, d\mu_t(x) \leq \\
&\quad \leq (1-t) \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi\left(\mathrm{pr}^1(\xi,\eta)\right)\, d\pi(\xi,\eta) + t \iint_{\mathbb{R}^d\times\mathbb{R}^d} \phi\left(\mathrm{pr}^2(\xi,\eta)\right)\, d\pi(\xi,\eta) \\
&\quad = (1-t) \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) + t \int_{\mathbb{R}^d} \phi(y)\, d\mu_1(y) \leq 0\,,
\end{aligned}$$

since $\mathrm{pr}^i \sharp \pi$ are the marginal measures of π, which belong to $\tilde{S}_p^\Phi$. The conclusion follows from the arbitrariness of $\phi(\cdot) \in \Phi$. □

Remark 8. In particular, considering also the first item in Remark 4, the above result holds for $\Phi := \{d_S(\cdot) - \alpha\}$ when S is nonempty, closed and convex, and $\alpha \in [0,1]$. In this case, since in the above proof we use only the convexity property of $d_S(\cdot)$, the statement holds also if we equip $\mathbb{R}^d$ with a different norm than the Euclidean one.

We conclude this section by investigating the semiconcavity properties of the generalized distance along geodesics. The case $p = 2$ is particularly easy thanks to the geometric structure of the metric space $\mathcal{P}_2(\mathbb{R}^d)$.

Proposition 6 (Semiconcavity of $\tilde{d}^2_{\tilde{S}_2^\Phi}$)**.** *Let $\tilde{S}_2^\Phi$ be a generalized target in $\mathcal{P}_2(\mathbb{R}^d)$. Then the square of the generalized distance satisfies the following* global semiconcavity inequality: *for every $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and every $t \in [0,1]$*

$$\tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_t) \geq (1-t)\, \tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_0) + t\, \tilde{d}^2_{\tilde{S}_2^\Phi}(\mu_1) - t(1-t)\, W_2^2(\mu_0,\mu_1),$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ is any constant speed geodesic for W_2 joining μ_0 and μ_1.

Proof. Owing to [3, Theorem 7.3.2], we have that for any measure $\sigma \in \mathcal{P}_2(\mathbb{R}^d)$ the function $\mu \mapsto W_2^2(\mu,\sigma)$ is semiconcave along geodesics, with semiconcavity constant independent by σ, i.e. it satisfies for every $t \in [0,1]$

$$W_2^2(\mu_t,\sigma) + t(1-t)\, W_2^2(\mu_0,\mu_1) \geq (1-t)\, W_2^2(\mu_0,\sigma) + t\, W_2^2(\mu_1,\sigma).$$

The conclusion follows by passing to the infimum on $\sigma \in \tilde{S}_2^\Phi$. □

In the case $p \neq 2$ we need additional requirements on Φ.

Proposition 7 (Semiconcavity of $\tilde{d}^p_{\tilde{S}_p^\Phi}$)**.** *Let $p \geq 1$, and $\tilde{S}_p^\Phi$ be a generalized target. Assume that $\tilde{S}_p^\Phi$ admits a classical counterpart $S \subseteq \mathbb{R}^d$. Let $K \subseteq \mathbb{R}^d \setminus S$ be compact and convex. Then the p-th power of the generalized distance $\tilde{d}_{\tilde{S}_p^\Phi}(\cdot)$ satisfies the following* local semiconcavity inequality: *there exists a constant $C = C(p,K) > 0$ such that for every $\mu_0, \mu_1 \in \mathcal{P}_p(K)$ we have*

$$\tilde{d}^p_{\tilde{S}_p^\Phi}(\mu_t) \geq (1-t)\, \tilde{d}^p_{\tilde{S}_p^\Phi}(\mu_0) + t\, \tilde{d}^p_{\tilde{S}_p^\Phi}(\mu_1) - Ct(1-t)\, W_p^{\min\{p,2\}}(\mu_0,\mu_1), \tag{7}$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]}$ is any constant speed geodesic for W_p joining μ_0 and μ_1.

Proof. In this proof to make clearer the notation we will omit the superscript Φ, since Φ is fixed. Under the above assumptions, and recalling Proposition 4, we have $\tilde{d}_{\tilde{S}_p}(\mu_0) = \|d_S\|_{L^p_{\mu_0}}$.

Given $x_0, x_1 \in K$ and $t \in [0,1]$ we set

$$x_t := (1-t)x_0 + tx_1, \qquad\qquad d_t := (1-t)d_S(x_0) + td_S(x_1).$$

According to [7, Proposition 2.2.2], there exists $c = c(K) > 0$ such that d_S satisfies the following inequality for all $x_0, x_1 \in K$:

$$d_S(x_t) \geq d_t - ct(1-t)|x_0 - x_1|^2,$$

By using [7, Proposition 2.1.12 (i)], we obtain that

$$d_S^p(x_t) \geq (1-t)\, d_S^p(x_0) + t\, d_S^p(x_1) - C' t(1-t)|x_0 - x_1|^{\min\{p,2\}}, \tag{8}$$

with $C' = C'(p, K)$.

For any Borel sets $A, B \subseteq \mathbb{R}^d$ and $\pi \in \Pi(\mu_0, \mu_1)$, we now have $\mathrm{supp}(\pi) \subseteq K \times K$. Therefore, we choose a transport plan $\pi \in \Pi_o^p(\mu_0, \mu_1)$ realizing the p-Wasserstein distance between μ_0 and μ_1, so that the representation in formula (6) holds, and we integrate the estimate (8) to find that

$$\begin{aligned}\int_{\mathbb{R}^d} d_S^p(x)\, d\mu_t = \iint_{\mathbb{R}^d \times \mathbb{R}^d} d_S^p(x_t)\, d\pi \geq (1-t) \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_0 + t \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_1 \\ - C'\, t\, (1-t) \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^{\min\{p,2\}}\, d\pi,\end{aligned}$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,1]} \subseteq \mathcal{P}_p(\mathbb{R}^d)$ is the constant speed geodesic corresponding to π. But according to Proposition 4, there holds

$$\tilde{d}^p_{\tilde{S}_p}(\mu_t) = \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_t(x), \qquad \text{and} \qquad \tilde{d}^p_{\tilde{S}_p}(\mu_i) = \int_{\mathbb{R}^d} d_S^p(x)\, d\mu_i(x), \quad i = 0, 1,$$

and applying Jensen's inequality to the concave map $\xi \mapsto \xi^{\gamma/p}$ on $\mathbb{R}^+$, with $\gamma = \min\{p, 2\}$, we obtain that

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^{\min\{p,2\}}\, d\pi \leq \begin{cases} \displaystyle\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^p\, d\pi, & \text{for } 1 \leq p < 2, \\[2ex] \displaystyle\left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x_0 - x_1|^p\, d\pi\right)^{2/p}, & \text{for } p \geq 2. \end{cases}$$

We thus conclude that

$$\tilde{d}^p_{\tilde{S}_p}(\mu_t) \geq (1-t)\, \tilde{d}^p_{\tilde{S}_p}(\mu_0) + t\, \tilde{d}^p_{\tilde{S}_p}(\mu_1) - C'\, t\, (1-t)\, W_p^{\min\{p,2\}}(\mu_0, \mu_1),$$

and the proof is completed. □

Remark 9. Notice that inequality (7) implies that, for $p \geq 2$ and under the assumption of Proposition 7, the functional $-\tilde{d}^p_{\tilde{S}_p}(\cdot) \colon \mathcal{P}_p(K) \to]-\infty, 0]$ is λ-geodesically convex, in the sense of [3, Definition 9.1.1], with $\lambda = -2C$.

5. **The generalized minimum time function.**

Definition 5.1 (Generalized minimum time function)**.** Given a generalized target $\tilde{S}_p = \tilde{S}_p^\Phi$ defined in Definition 4.2, we define the generalized minimum time function $\tilde{T}_p : \mathcal{P}_p(\mathbb{R}^d) \to [0, +\infty]$ by

$$\tilde{T}_p(\mu) := \inf\{T \geq 0 : \text{ there exists } \boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu) \text{ s.t. } \mu_T \in \tilde{S}_p\},$$

where we set $\inf \emptyset = +\infty$ by convention. We say that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$ is time optimal from μ if $\tilde{T}_p(\mu) \leq T < +\infty$ and $\mu_{\tilde{T}_p(\mu)} \in \tilde{S}_p$.

Proposition 8 (Properties of $\tilde{T}_p$)**.** *Assume Hypothesis 3.3 for F. Then*

1. *for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ with $\tilde{T}_p(\mu) < +\infty$ there exists a time optimal admissible trajectory from μ;*
2. *the function $\tilde{T}_p(\cdot)$ is lower semicontinuous;*
3. *the following Dynamic Programming Principle holds*

$$\tilde{T}_p(\mu) = \inf\{t + \tilde{T}_p(\mu_t) : \boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu),\, T > 0\}. \tag{9}$$

In particular, $t \mapsto t+\tilde{T}_p(\mu_t)$ is nondecreasing along every admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$, and it is constant if and only if $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$ is the restriction to $[0,T] \cap [0, \tilde{T}_p(\mu)]$ of an optimal trajectory.

Proof.

1. Fix $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ with $\tilde{T}_p(\mu) < +\infty$. For any $n \in \mathbb{N}\setminus\{0\}$ there exists $T_n > 0$ and $\boldsymbol{\mu^{(n)}} = \{\mu_t^{(n)}\}_{t\in[0,T_n]} \in \mathcal{A}^p_{[0,T_n]}(\mu)$ such that $\mu^{(n)}_{T_n} \in \tilde{S}_p$ and $T_n \leq \tilde{T}_p(\mu) + 1/n$. We can extend each $\boldsymbol{\mu^{(n)}}$ to an admissible curve defined on $\tilde{T}_p(\mu)+1$ (possibly concatenating it with an element of $\mathcal{A}_{[T_n,\tilde{T}_p(\mu)+1]}(\mu^{(n)}_{T_n})$, which is nonempty for all $n \in \mathbb{N}\setminus\{0\}$). Thus, without loss of generality, we may assume that we have a sequence $\boldsymbol{\hat{\mu}^{(n)}} = \{\hat{\mu}_t^{(n)}\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$ of admissible trajectories, which are all defined in $[0,T]$ with $T = \tilde{T}_p(\mu) + 1$, and satisfying $\hat{\mu}^{(n)}_{T_n} \in \tilde{S}_p$ where $T_n \leq \tilde{T}_p(\mu) + 1/n$. Recalling the compactness of $\mathcal{A}^p_{[0,T]}(\mu)$ (see Corollary 1), up to passing to a subsequence, the sequence of curves $\{\boldsymbol{\hat{\mu}^{(n)}}\}_{n\in\mathbb{N}}$ uniformly converges to $\boldsymbol{\hat{\mu}^{\infty}} = \{\hat{\mu}^\infty_t\}_{t\in[0,T]}$ and, moreover, we have $T_n \to \ell$. In particular, $\hat{\mu}^{(n)}_{T_n} \to \hat{\mu}^\infty_\ell$ which, by the closedness of $\tilde{S}_p$, implies $\hat{\mu}^\infty_\ell \in \tilde{S}_p$, and so $\tilde{T}_p(\mu) \leq \ell$. But passing to the limit in $T_n \leq \tilde{T}_p(\mu)+1/n$ yields the reverse inequality, thus $\ell = \tilde{T}_p(\mu)$, hence $\boldsymbol{\hat{\mu}^{\infty}}$ is optimal.
2. Let $\{\mu^{(n)}\}_{n\in\mathbb{N}} \subseteq \mathcal{P}_p(\mathbb{R}^d)$ be a W_p-converging sequence satisfying $\mu^{(n)} \to \mu^\infty$ and $\liminf\limits_{n\to+\infty} \tilde{T}_p(\mu^{(n)}) =: \ell \in \mathbb{R}$. If $\ell = +\infty$ there is nothing to prove, so let us assume $\ell < +\infty$. As before, up to concatenation and restriction and by taking n sufficiently large, this implies that there exists a sequence $\{\boldsymbol{\mu^{(n)}}\}_{n\in\mathbb{N}}$ such that $\boldsymbol{\mu^{(n)}} = \{\mu^{(n)}_t\}_{t\in[0,\ell+1]} \in \mathcal{A}_{[0,\ell+1]}(\mu^{(n)})$ and $\mu^{(n)}_{\tilde{T}_p(\mu^{(n)})} \in \tilde{S}_p$ for all $n \in \mathbb{N}$.

   By Theorem 3.6, there exists a sequence $\{\boldsymbol{\hat{\mu}^{(n)}} = \{\hat{\mu}^{(n)}_t\}_{t\in[0,\ell+1]}\}_{n\in\mathbb{N}} \subseteq \mathcal{A}_{[0,\ell+1]}(\mu^\infty)$ such that

$$W_p(\hat{\mu}^{(n)}_t, \mu^{(n)}_t) \leq D \cdot W_p(\mu^{(n)}, \mu^\infty),$$

for all $t \in [0, \ell+1]$, where $D := 2^{\frac{p-1}{p}} e^{L(2+Le^{L(\ell+1)})(\ell+1)}$. Recalling the compactness of $\mathcal{A}^p_{[0,\ell+1]}(\mu^\infty)$ (see Corollary 1), up to a passing to a subsequence, the sequence of curves $\{\hat{\boldsymbol{\mu}}^{(\boldsymbol{n})}\}_{n\in\mathbb{N}}$ uniformly converges to $\hat{\boldsymbol{\mu}}^{\boldsymbol{\infty}} = \{\hat{\mu}^\infty_t\}_{t\in[0,\ell+1]}$, in particular, we have that

$$\begin{aligned}
W_p\left(\mu^{(n)}_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^\infty_\ell\right) \leq & W_p\left(\mu^{(n)}_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^{(n)}_{\tilde{T}_p(\mu^{(n)})}\right) + W_p\left(\hat{\mu}^{(n)}_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^\infty_{\tilde{T}_p(\mu^{(n)})}\right) + \\
& + W_p\left(\hat{\mu}^\infty_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^\infty_\ell\right) \\
\leq & D \cdot W_p\left(\mu^{(n)}, \mu^\infty\right) + \sup_{t\in[0,\ell+1]} W_p\left(\hat{\mu}^{(n)}_t, \hat{\mu}^\infty_t\right) + \\
& + W_p\left(\hat{\mu}^\infty_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^\infty_\ell\right).
\end{aligned}$$

By taking the limit for $n \to +\infty$, we have that $W_p\left(\mu^{(n)}_{\tilde{T}_p(\mu^{(n)})}, \hat{\mu}^\infty_\ell\right) \to 0$, hence, by the closedness of $\tilde{S}_p$, we obtain $\hat{\mu}^\infty_\ell \in \tilde{S}_p$, and so $\tilde{T}_p(\mu^\infty) \leq \ell$.

3. Let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$ and $\hat{\boldsymbol{\mu}}^{(\boldsymbol{t})} \in \mathcal{A}^p_{[0,\tilde{T}_p(\mu_t)]}(\mu_t)$ such that $\hat{\boldsymbol{\mu}}^{(\boldsymbol{t})}$ is optimal for μ_t (such an optimal trajectory exists by item (1)). For any $t \in [0,T]$, the concatenation $\boldsymbol{\mu}_{|[0,t]} \odot \hat{\boldsymbol{\mu}}^{(\boldsymbol{t})} \in \mathcal{A}^p_{[0,t+\tilde{T}_p(\mu_t)]}(\mu)$, and so $\tilde{T}_p(\mu) \leq t + \tilde{T}_p(\mu_t)$ for every $t \in [0,T]$, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathcal{A}^p_{[0,T]}(\mu)$, $t > 0$, giving the first inequality in (9). In particular, for $0 \leq t \leq s \leq T$, we have
$$\tilde{T}_p(\mu) \leq t + \tilde{T}_p(\mu_t) \leq t + (s-t) + \tilde{T}_p(\mu_s) = s + \tilde{T}_p(\mu_s),$$
since the restriction of $\boldsymbol{\mu}$ to $[t,T]$ is an admissible trajectory from μ_t. Thus $t \mapsto t + \tilde{T}_p(\mu_t)$ is nondecreasing along all the admissible trajectories. If $\boldsymbol{\mu}$ is an optimal trajectory, by taking $s = \tilde{T}_p(\mu)$ we have $\tilde{T}_p(\mu_s) = 0$ and so $\tilde{T}_p(\mu) = t + \tilde{T}_p(\mu_t)$ for all $t \in [0, \tilde{T}_p(\mu)]$, which gives equality in (9). Finally, assume that $t \mapsto t + \tilde{T}_p(\mu_t)$ is constant along an admissible trajectory $\boldsymbol{\mu} \in \mathcal{A}^p_{[0,T]}(\mu)$. By (9) we have that $\tilde{T}_p(\mu) = t + \tilde{T}_p(\mu_t)$ for all $t \in [0,T]$. If $T \geq \tilde{T}_p(\mu)$, this implies that $\boldsymbol{\mu}$ is optimal, since by taking $t = \tilde{T}_p(\mu)$ we obtain $\tilde{T}_p(\mu_{\tilde{T}_p(\mu)}) = 0$ and so $\mu_{\tilde{T}_p(\mu)} \in \tilde{S}_p$. If $T < \tilde{T}_p(\mu)$ we concatenate $\boldsymbol{\mu}$ with an optimal trajectory $\hat{\boldsymbol{\mu}} = \{\hat{\mu}_s\}_{s\in[0,\tilde{T}_p(\mu_T)]} \in \mathcal{A}_{[0,\tilde{T}_p(\mu_T)]}(\mu_T)$ for μ_T. Set $\boldsymbol{\mu} \odot \hat{\boldsymbol{\mu}} = \{\tilde{\mu}_s\}_{s\in[0,T+\tilde{T}_p(\mu_T)]}$. In particular, we have $\tilde{T}_p(\mu_T) = s + \tilde{T}_p(\hat{\mu}_s)$ for all $s \in [0, \tilde{T}_p(\mu_T)]$, thus $\tilde{T}_p(\mu) = \tau + \tilde{T}_p(\tilde{\mu}_\tau)$ for all $\tau \in [0, \tilde{T}_p(\mu)]$. By taking $\tau = \tilde{T}_p(\mu)$ we obtain $\tilde{T}_p(\mu_{\tilde{T}_p(\mu)}) = 0$ and so $\mu_{\tilde{T}_p(\mu)} \in \tilde{S}_p$ thus the concatenation $\boldsymbol{\mu} \odot \hat{\boldsymbol{\mu}}$ is an optimal trajectory, whose restriction to $[0,T]$ is $\boldsymbol{\mu}$.

□

The following definition of Small-Time Local Attainability (STLA) has been introduced in [27] for finite-dimensional control systems, but can be easily generalized in our framework.

Definition 5.2 (STLA for Wasserstein spaces)**.** We say that the system with generalized target $\tilde{S}_p$ satisfies the STLA property if

Property **(STLA)**: for any $\varepsilon > 0$ and $\hat{\mu} \in \tilde{S}_p$ there exists $\delta > 0$ such that $\tilde{T}_p(\mu) \leq \varepsilon$ for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ satisfying $W_p(\mu, \hat{\mu}) \leq \delta$.

The link between STLA and continuity of the generalized minimum time is provided by the following result.

Proposition 9 (STLA and continuity of $\tilde{T}_p$)**.** *Let $\tilde{S}_p$ be a generalized target. Assume Hypothesis 3.3 for F, and that* **(STLA)** *holds for the system. Then $\tilde{T}_p : \mathcal{P}_p(\mathbb{R}^d) \to [0, +\infty]$ is continuous at every point where it is finite.*

Proof. Recalling the l.s.c. of $\tilde{T}_p(\cdot)$, given $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ with $\tilde{T}_p(\mu) = +\infty$, we have $\lim_{n\to+\infty} \tilde{T}_p(\mu^{(n)}) = +\infty$ for every sequence $\{\mu^{(n)}\}_{n\in\mathbb{N}}$ converging to μ in W_p.

Therefore, we assume $T := \tilde{T}_p(\mu) < +\infty$. Since $\tilde{T}_p(\cdot)$ is l.s.c., it is enough to prove that for all $\{\bar{\mu}^{(n)}\}_n \subseteq \mathcal{P}_p(\mathbb{R}^d)$ such that $W_p(\bar{\mu}^{(n)}, \mu) \to 0$ as $n \to +\infty$, we have

$$\limsup_{n\to+\infty} \tilde{T}_p(\bar{\mu}^{(n)}) \le T.$$

Fix an optimal trajectory $\boldsymbol{\mu}^\infty := \{\mu_t^\infty\}_{t\in[0,T]}$ starting from $\mu^\infty_{|t=0} = \mu$. Let $\{\mu^{(n)}\}_{n\in\mathbb{N}}$ be a sequence converging to μ in W_p and such that $\lim_{n\to+\infty} \tilde{T}_p(\mu^{(n)})$ exists. By Theorem 3.6, there exists a sequence of admissible trajectories $\{\boldsymbol{\mu}^{(n)}\}_{n\in\mathbb{N}}$ such that

- $\boldsymbol{\mu}^{(n)} = \{\mu_t^{(n)}\}_{t\in[0,T]}$, $\mu_0^{(n)} = \mu^{(n)}$ for all $n \in \mathbb{N}$ and
- $\tilde{d}_{\tilde{S}_p}(\mu_T^{(n)}) \le W_p(\mu_T^{(n)}, \mu_T^\infty) \le D \cdot W_p(\mu^{(n)}, \mu)$, recalling that $\mu_T^\infty \in \tilde{S}_p$,

where $D := 2^{\frac{p-1}{p}} e^{L(2+Le^{LT})T}$. In particular, by **(STLA)**, given $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that for all $n > n_\varepsilon$ we have $\tilde{T}_p(\mu_T^{(n)}) \le \varepsilon$. By Dynamic Programming principle, we have

$$\tilde{T}_p(\mu^{(n)}) \le T + \tilde{T}_p(\mu_T^{(n)}) \le T + \varepsilon,$$

By letting $n \to +\infty$ and $\varepsilon \to 0$, we have

$$\lim_{n\to+\infty} \tilde{T}_p(\mu^{(n)}) \le T.$$

We conclude by the arbitrariness of the sequence $\{\mu^{(n)}\}_{n\in\mathbb{N}}$. $\square$

Definition 5.3. Given $\Phi \subset C_b^0(\mathbb{R}^d)$, $\phi \in \Phi$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$ we define

$$L_\phi(\mu) := \int_{\mathbb{R}^d} \phi(x)\, d\mu(x), \qquad \sigma_\Phi(\mu) := \sup_{\phi\in\Phi} L_\phi(\mu).$$

Our aim is to provide a sufficient condition for **(STLA)**, following the line of [28] and [29] for finite-dimensional systems. We recall that the l.s.c. of $\tilde{T}_p(\cdot)$ was already showed in [13, Theorem 4] in a simplified setting, while a stronger sufficient condition was provided in [11, Theorem 4.1] to prove the Lipschitz continuity regularity. The continuity of $\tilde{T}_p(\cdot)$ was a crucial assumption also in [13, Theorem 8] to prove that it solves an Hamilton-Jacobi-Bellman equation in Wasserstein space.

The following definition establishes a quantitative estimate of the maximal infinitesimal decreasing of the functions $\phi \in \Phi$ defining the generalized target, along the admissible trajectories of the system.

Definition 5.4. We say that the generalized target $\tilde{S}_p^\Phi$ is (r, Q)-*attainable* if there exist continuous maps

$$r : [0, +\infty[\to \left[0, \min\left\{1, \frac{1}{2L}\right\}\right], \quad Q : \left[0, \min\left\{1, \frac{1}{2L}\right\}\right] \times [0, +\infty[\to \mathbb{R}$$

such that

1. $r(q) = 0$ if and only if $q = 0$;
2. $Q(r(q), q) < 0$ for all $q \in]0, +\infty[$;

3. the function $q \mapsto \dfrac{r(q)}{|Q(r(q),q)|}$ is decreasing and integrable on $[0,+\infty[$.
4. for any $\mu \in \mathcal{P}_p(\mathbb{R}^d) \setminus \tilde{S}_p^\Phi$ there exists $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,r(\sigma_\Phi(\mu))]} \in \mathcal{A}^p_{[0,r(\sigma_\Phi(\mu))]}(\mu)$ such that
$$\inf_{t\in[0,r(\sigma_\Phi(\mu))]} \{\sigma_\Phi(\mu_t) - \sigma_\Phi(\mu)\} \le 2\,Q(r(\sigma_\Phi(\mu)),\sigma_\Phi(\mu)).$$

Remark 10. Roughly speaking, (r,Q)-attainability expresses a relation between the variation of the distance (or a related positive function vanishing only on the target) along a particular admissible trajectory in a time interval, and the size of the time interval itself. The integrability condition asks that the approaching speed, which can be seen as the quotient between the variation of the distance and the time needed to realize it, is sufficiently high to ensure that the target will be reached in finite time. Finite-dimensional examples of similar constructions can be found e.g. in [28], while an example in this setting with no interactions can be found in [11].

With the notations of Definitions 5.3, 5.4 we state the following results of this section.

Proposition 10. *Assume Hypothesis 3.3 for F and that the generalized target $\tilde{S}_p^\Phi$ is (r,Q)-attainable. Then*
$$\boldsymbol{T}(\mu) := \int_0^{\sigma_\Phi(\mu)} \frac{r(q)\,dq}{|Q(r(q),q)|} \ge \tilde{T}_p(\mu).$$

Proof. Define sequences $\{\mu^{(i)}\}_{i\in\mathbb{N}} \subseteq \mathcal{P}(\mathbb{R}^d)$, $\{\sigma_i\}_{i\in\mathbb{N}}$, $\{t_i\}_{i\in\mathbb{N}} \subseteq [0,1]$ as follows. Set $\mu^{(0)} = \mu$. Suppose to have defined $\mu^{(i)}$, then define $\sigma_i = \sigma_\Phi(\mu^{(i)})$. We notice that, by assumption, if $\mu^{(i)} \notin \tilde{S}_p^\Phi$ we have $\sigma_i > 0$, and so $Q(r(\sigma_i),\sigma_i) < 0$.

By property (4) in Definition 5.4, if $\mu^{(i)} \notin \tilde{S}_p^\Phi$ there exists $\boldsymbol{\mu}^{(i)} = \{\mu_t^{(i)}\}_{t\in[0,r(\sigma_i)]} \in \mathcal{A}^p_{[0,r(\sigma_i)]}(\mu^{(i)})$ such that
$$\inf_{t\in[0,r(\sigma_i)]} \{\sigma_\Phi(\mu_t^{(i)}) - \sigma_i\} \le 2\,Q(r(\sigma_i),\sigma_i).$$
Thus, for any $\varepsilon > 0$ there exists $t_i^\varepsilon \in [0,r(\sigma_i)]$ such that
$$\sigma_\Phi(\mu^{(i)}_{t_i^\varepsilon}) - \sigma_i \le 2\,Q(r(\sigma_i),\sigma_i) + \varepsilon.$$
Notice that, if we choose ε sufficiently small, in particular $0 < \varepsilon < -2\,Q(r(\sigma_i),\sigma_i)$, then $t_i^\varepsilon \neq 0$. We thus fix $\hat{\varepsilon}(i) = -Q(r(\sigma_i),\sigma_i)$ and set $t_i = t_i^{\hat{\varepsilon}(i)} > 0$ and $\mu^{(i+1)} = \mu^{(i)}_{t_i}$.

While, if $\mu^{(i)} \in \tilde{S}_p^\Phi$, then we set $t_i = 0$ and $\mu^{(i+1)} = \mu^{(i)}_{t_i} = \mu^{(i)}$.

Thus, together with property (1) in Definition 5.4, this implies that $\mu^{(i)} \notin \tilde{S}_p^\Phi$ if and only if $\sigma_i,\, t_i > 0$.

Notice that $\sigma_i \ge 0$ for all $i \in \mathbb{N}$, moreover, if $\sigma_i = 0$ then $\sigma_m = t_m = 0$ for all $m \ge i$.

For every $i \in \mathbb{N}$ such that $\sigma_i \neq 0$ we have
$$\sigma_{i+1} - \sigma_i \le 2\,Q(r(\sigma_i),\sigma_i) + \hat{\varepsilon}(i) = Q(r(\sigma_i),\sigma_i) < 0, \tag{10}$$
by property (2) in Definition 5.4. Thus the sequence $\{\sigma_i\}_{i\in\mathbb{N}}$ is decreasing and bounded from below, and so it has a limit $\sigma_\infty \ge 0$. If $\sigma_i = 0$ for some $i \in \mathbb{N}$ then $\sigma_\infty = 0$. If $\sigma_i \neq 0$ for all $i \in \mathbb{N}$, by passing to the limit in (10) we get $Q(r(\sigma_\infty),\sigma_\infty) = 0$, by continuity of $Q(\cdot,\cdot)$ and $r(\cdot)$, which implies $\sigma_\infty = 0$.

We have
$$\boldsymbol{T}(\mu) \geq \sum_{\substack{i\in\mathbb{N}\\ \sigma_i\neq 0}} \frac{r(\sigma_i)(\sigma_i-\sigma_{i+1})}{|Q(r(\sigma_i),\sigma_i)|} \geq \sum_{i\in\mathbb{N}} t_i.$$
To conclude the proof, we consider two cases

- assume that $\sigma_i \neq 0$ for all $i\in\mathbb{N}$. Then for any $i\in\mathbb{N}$ there exists an admissible trajectory $\boldsymbol{\mu}^{(\infty)} = \{\mu_t^{(\infty)}\}_{[0,\boldsymbol{T}]}$ starting from μ and coinciding with $\boldsymbol{\mu}^{(i)}$ on $[t_{i-1},t_i]$. In particular, $\mu^{(\infty)}_{\sum t_i} \in \tilde{S}^{\Phi}_p$ since $\sigma_\infty = 0$, and so $\boldsymbol{T}(\mu) \geq \tilde{T}_p(\mu)$.
- let $\hat{\imath}$ the minimum of the set $\{i\in\mathbb{N} :\ t_i = 0\}$. Then there exists an admissible trajectory $\hat{\boldsymbol{\mu}}^{(\hat{\imath})} = \{\hat{\mu}_t^{(\hat{\imath})}\}_{[0,\boldsymbol{T}]}$ starting from μ and coinciding with $\boldsymbol{\mu}^{(i)}$ on $[t_{i-1},t_i]$, for all $i\geq 1$. In particular, $\hat{\mu}^{(\hat{\imath})}_{\sum_{i=1}^{\hat{\imath}-1} t_i} \in \tilde{S}^{\Phi}_p$, and so
$$\boldsymbol{T}(\mu) \geq \sum_{i=1}^{\infty} t_i = \sum_{i=1}^{\hat{\imath}-1} t_i \geq \tilde{T}_p(\mu).$$

Thus in both cases we have $\boldsymbol{T}(\mu) \geq \tilde{T}_p(\mu)$, which concludes the proof. □

Theorem 5.5 (Sufficient condition for **(STLA)**). *Assume Hypothesis 3.3 for F and that the generalized target $\tilde{S}^{\Phi}_p$ is (r,Q)-attainable. Assume that there exists $C>0$ and an open set $U\subseteq \mathcal{P}_p(\mathbb{R}^d)$ such that $U\supseteq \tilde{S}^{\Phi}_p$ and $\sigma_\Phi(\mu)\leq C$ for all $\mu\in U$. Then* **(STLA)** *holds.*

Proof. Fix $\varepsilon>0$. Since $\max\{\sigma_\Phi(\mu),0\}\leq C$ in a neighborhood of $\tilde{S}^{\Phi}_p$, we have that the convex function $\mu\mapsto \max\{\sigma_\Phi(\mu),0\}$ is continuous in a neighborhood of $\tilde{S}^{\Phi}_p$ and vanishes exactly on $\tilde{S}^{\Phi}_p$. Thus for any $\varepsilon>0$ there exists $\rho,\delta>0$ such that if $d_{\tilde{S}^{\Phi}_p}(\mu)\leq\delta$ we have $\sigma_\Phi(\mu)\leq\rho$ and
$$\varepsilon > \int_0^\rho \frac{r(q)\,dq}{|Q(r(q),q)|} \geq \tilde{T}^{\Phi}_p(\mu),$$
recalling that by the integrability assumption in item (3) in Definition 5.4, the map
$$\rho\mapsto \int_0^\rho \frac{r(q)\,dq}{|Q(r(q),q)|}$$
is continuous. □

In conclusion, in order to check the (r,Q)-attainability of a set from the data of the problem, the following result may serve the purpose.

Corollary 3. *Given $\alpha\geq 0$, an interval $I\subseteq\mathbb{R}$, $\gamma\in AC(I;\mathbb{R}^d)$ and $v:\mathbb{R}^d\to\mathbb{R}^d$, we define*
$$\Delta^v_{\alpha,\gamma}(t) := \left|\frac{\gamma(t)-\gamma(0)}{t^{1+\alpha}} - v(\gamma(0))\right|.$$
Let $\mathcal{D}\subseteq\mathcal{P}_2(\mathbb{R}^d)$ and assume that there exist constants $C_\phi\geq 0$, $\alpha,\beta,K>0$ such that, by defining for any $\mu\in\mathcal{D}$
$$t_\mu := \min\left\{1,\frac{1}{2L},\sigma_\Phi^{1/\beta}(\mu)\right\} \text{ and } I_\mu := [0,t_\mu],$$
we have

a.) $\Phi := \{\phi\}$, where ϕ is semiconcave with constant C_ϕ;

b.) for all $\mu\in\mathcal{D}\setminus\tilde S_2^\Phi$ there exist functions $v_\mu,\xi_\mu\in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$, $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_{I_\mu})$, and constants $C_{2,\mu},C_{3,\mu},C_{4,\mu}>0$ satisfying

- $0\le\alpha<\beta-1$;
- $\boldsymbol{\mu}=\{e_t\sharp\boldsymbol{\eta}\}_{t\in I_\mu}\in\mathcal{A}_{I_\mu}(\mu)$, *with* $e_{t_\mu}\sharp\boldsymbol{\eta}\in\mathcal{D}$;
- $\xi_\mu(x)\in\partial^P\phi(x)$ *for μ-a.e.* $x\in\mathbb{R}^d$;
- $\displaystyle\int_{\mathbb{R}^d}\langle\xi_\mu(x),v_\mu(x)\rangle\,d\mu(x)\le -C_{2,\mu}<0$;
- $\displaystyle\left(\int_{\mathbb{R}^d\times\Gamma_{I_\mu}}|\Delta^{v_\mu}_{\alpha,\gamma}(t_\mu)|^2\,d\boldsymbol{\eta}(x,\gamma)\right)^{1/2}\le C_{3,\mu}t_\mu$;
- $\|v_\mu\|_{L^2_\mu}\le C_{4,\mu}$;
- $\left(-C_{2,\mu}+C_{3,\mu}\|\xi_\mu\|_{L^2_\mu}t_\mu+2C_\phi(C^2_{3,\mu}t^2_\mu+C^2_{4,\mu})t_\mu^{\alpha+1}\right)\le -2K\cdot t_\mu$.

Then **(STLA)** *holds in $\mathcal{D}$ and for all $\mu\in\mathcal{D}$ we have*

$$\tilde T_2^\Phi(\mu)\le\begin{cases}\dfrac{\beta\sigma_\Phi^{\frac{\beta-\alpha-1}{\beta}}(\mu)}{K(\beta-\alpha-1)}, & \text{if } \sigma_\Phi(\mu)\le\min\{1,(2L)^{-\beta}\}\\[2ex] \dfrac{\beta\,(2L)^{-\beta+\alpha+1}}{K(\beta-\alpha-1)}+\dfrac{1}{K}(2L)^{\alpha+1}\,(\sigma_\Phi(\mu)-(2L)^{-\beta}), & \text{if } \sigma_\Phi(\mu)\ge(2L)^{-\beta}=\min\{1,(2L)^{-\beta}\}\\[2ex] \dfrac{\beta}{K(\beta-\alpha-1)}+\dfrac{1}{K}(\sigma_\Phi(\mu)-1), & \text{if } \sigma_\Phi(\mu)\ge 1=\min\{1,(2L)^{-\beta}\}.\end{cases}$$

Proof. Indeed, for $\boldsymbol{\eta}$-a.e. $(x,\gamma)\in\mathbb{R}^d\times\Gamma_{I_\mu}$ we have

$$\begin{aligned}\phi(\gamma(t_\mu))-\phi(\gamma(0))\le&\langle\xi_\mu(\gamma(0)),\gamma(t_\mu)-\gamma(0)\rangle+C_\phi|\gamma(t_\mu)-\gamma(0)|^2\\ \le&t_\mu^{\alpha+1}\langle\xi_\mu(\gamma(0)),v_\mu(\gamma(0))\rangle+t_\mu^{\alpha+1}|\xi_\mu(\gamma(0))|\Delta^{v_\mu}_{\alpha,\gamma}(t_\mu)+\\ &+C_\phi t_\mu^{2(\alpha+1)}\left(\Delta^{v_\mu}_{\alpha,\gamma}(t_\mu)+|v_\mu(\gamma(0))|\right)^2.\end{aligned}$$

Integrating w.r.t. $\boldsymbol{\eta}$ and using Hölder's inequality yields

$$\begin{aligned}&\sigma_\Phi(\mu_{t_\mu})-\sigma_\Phi(\mu)\le\\ &\le-C_{2,\mu}t_\mu^{\alpha+1}+\int_{\mathbb{R}^d\times\Gamma_{I_\mu}}|\xi_\mu(x)|\cdot t_\mu^{\alpha+1}\Delta^{v_\mu}_{\alpha,\gamma}(t_\mu)\,d\boldsymbol{\eta}(x,\gamma)+2C_\phi(C^2_{3,\mu}t^2_\mu+C^2_{4,\mu})t_\mu^{2(\alpha+1)}\\ &\le-C_{2,\mu}t_\mu^{\alpha+1}+t_\mu^{\alpha+2}\|\xi_\mu\|_{L^2_\mu}\cdot C_{3,\mu}+2C_\phi(C^2_{3,\mu}t^2_\mu+C^2_{4,\mu})t_\mu^{2(\alpha+1)}\\ &\le t_\mu^{\alpha+1}\left(-C_{2,\mu}+C_{3,\mu}\|\xi_\mu\|_{L^2_\mu}t_\mu+2C_\phi(C^2_{3,\mu}t^2_\mu+C^2_{4,\mu})t_\mu^{\alpha+1}\right)\\ &\le-2K\cdot t_\mu^{\alpha+1}\cdot t_\mu=-2K\,t_\mu^{\alpha+2}.\end{aligned}\tag{11}$$

Choose

$$r(q):=\min\left\{1,\frac{1}{2L},q^{1/\beta}\right\},\qquad Q(t,q):=-K\cdot t^{\alpha+1}\cdot r(q).$$

In particular, we have $r(q)=0$ if and only if $q=0$, $Q(r(q),q)<0$ if $q\ne0$,

$$\frac{r(q)}{|Q(r(q),q)|}=\frac{1}{K\min\left\{1,\frac{1}{(2L)^{\alpha+1}},q^{\frac{\alpha+1}{\beta}}\right\}}=\frac{1}{K}\max\left\{1,(2L)^{\alpha+1},q^{-\frac{\alpha+1}{\beta}}\right\},$$

which is a decreasing integrable function of q. Furthermore, we notice that by definition $t_\mu=r(\sigma_\Phi(\mu))$ and $Q(r(\sigma_\Phi(\mu)),\sigma_\Phi(\mu))=-K\,t_\mu^{\alpha+2}$. Thus, from (11) we get

$$\sigma_\Phi(\mu_{r(\sigma_\Phi(\mu))})-\sigma_\Phi(\mu)\le 2Q(r(\sigma_\Phi(\mu)),\sigma_\Phi(\mu)),$$

and so we showed that $\tilde{S}_2^{\Phi}$ is (r, Q)-attainable. The result now follows from Theorem 5.5 and Proposition 10. □

6. **A brief comparison with classical attainability.** As reported in p. 352 [10] and at the beginning of Sec. 6.1 in [9], we recall that if $(\Omega, \mathcal{B}, \mathbb{P})$ is 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, given any $\mu_1, \mu_2 \in \mathcal{P}_p(\mathbb{R}^d)$ there exist $X_1, X_2 \in L^p_{\mathbb{P}}(\Omega)$ such that $\mu_i = X_i \sharp \mathbb{P}$, $i = 1, 2$, and $W_p(\mu_1, \mu_2) = \|X_1 - X_2\|_{L^p_{\mathbb{P}}}$. For instance, we can take $\Omega = [0, 1]$ endowed with the restriction of the Lebesgue measure to $[0, 1]$. This allows to use the well-known differential structure on $L^p_{\mathbb{P}}(\Omega)$ (the case $p = 2$ is the most common, in particular in the context of *mean field games*) in order to formulate the problem and possible derive finer properties of regularity for the solutions, by relying for instance on the theory of viscosity solution in infinite-dimensional Banach spaces. For instance, in this setting (more oriented to a stochastic process interpretation) the representation of Remark 2 can be expressed as follows.

Corollary 4 (Stochastic SP)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a reference probability space, where Ω is a Polish space and $\mathbb{P} \in \mathcal{P}(\Omega)$ an atomless measure. Then, $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ is an admissible trajectory if and only if there exists a stochastic process $X = X(\cdot)$ with*

$$\Omega \ni \omega \mapsto X(\cdot) := X(\cdot, \omega) \in AC(0, T) \cap C([0, T]; \mathbb{R}^d)$$

such that

- *$\mu_t = X_t \sharp \mathbb{P}$ for all $t \in [0, T]$;*
- *$\dot{X}(t) \in F(\mu_t, X(t))$ for a.e. $t \in [0, T]$.*

Proof. The proof is a consequence of Proposition 2, it sufficies to give the relation between the measure $\boldsymbol{\eta}$ of Proposition 2 and the stochastic process X. By e.g. Lemma 5.29 in [10], there exists a Borel map $\mathcal{V} : \Omega \to \mathbb{R}^d \times \Gamma_T$ such that $\boldsymbol{\eta} = \mathcal{V}\sharp\mathbb{P}$. We can thus conclude by setting $X_t := e_t \circ \mathcal{V}$ for all $t \in [0, T]$, indeed we have $\mu_t = e_t \sharp \boldsymbol{\eta} = X_t \sharp \mathbb{P}$. □

Here we will not follow this approach, since it is not in the purposes of the present paper to enter into this theory. However we actually implemented this more stochastic approach in a similar context in the preprint [15]. Indeed, there our interest is the study of a viability problem in the probability measure space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ by means of a suitable lifted Hamiltonian in $L^2(\Omega)$. For completeness, we mention that a theory of well-posedness for Hamilton-Jacobi equations in metric spaces has been introduced and developed for instance by [1, 2, 20, 21, 24, 25].

In this section, we provide viscosity results related to our study with the purpose to compare the concept of (r, Q)-attainability given in Definition 5.4 with the classical one provided by [27] in the finite dimensional framework. For this sake, we first study an Hamilton-Jacobi-Bellman equation associated with our time-optimal control problem in a suitable viscosity sense. In particular, we prove that the minimum time function is a viscosity supersolution of an HJB equation similarly to what occurs in the finite dimensional case.

Definition 6.1 (Superdifferential)**.** Let $1 < p < +\infty$, $U : \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}$, $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ and let p' be the conjugate exponent of p. We say that $q \in L^{p'}(\mathbb{R}^d)$ belongs to the viscosity *superdifferential* of U at μ, and we write $q \in D^+U(\mu)$, if for all $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ and all $\pi \in \Pi(\mu, \nu)$ we have

$$U(\nu) \le U(\mu) + \int_{\mathbb{R}^d \times \mathbb{R}^d} \langle q(x), y - x \rangle \, d\pi(x, y) + o\left(\left[\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \, d\pi(x, y)\right]^{1/p}\right).$$

Similarly, the set of viscosity *subdifferentials* of U at μ is defined by $D^-U(\mu) = -D^+(-U)(\mu)$.

Definition 6.2 (Viscosity solution)**.** Let $1 < p < +\infty$, and $U : \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}$. Let $\mathcal{H}(\mu, q)$ be defined for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ and $q \in L^{p'}_\mu(\mathbb{R}^d)$. We say that

- U is a *viscosity subsolution* of $\mathcal{H}(\mu, DU(\mu)) = 0$ if U is u.s.c. and $\mathcal{H}(\mu, q_\mu) \le 0$ for all $q_\mu \in D^+U(\mu)$ and $\mu \in \mathcal{P}_p(\mathbb{R}^d)$;
- U is a *viscosity supersolution* of $\mathcal{H}(\mu, DU(\mu)) = 0$ if U is l.s.c. and $\mathcal{H}(\mu, p_\mu) \ge 0$ for all $p_\mu \in D^-U(\mu)$ and $\mu \in \mathcal{P}_p(\mathbb{R}^d)$;
- U is a *viscosity solution* of $\mathcal{H}(\mu, DU(\mu)) = 0$ if it is both a super and a subsolution.

In the following, we prove that the minimum time function is a viscosity solution of an Hamilton-Jacobi-Bellman equation with the Hamiltonian $\mathcal{H}$ defined as follows

$$\mathcal{H}(\mu, q(\cdot)) := -1 - \inf_{\substack{v \in L^p_\mu \\ v(x) \in F(\mu, x)}} \langle q(\cdot), v(\cdot) \rangle_{L^{p'}, L^p} \tag{12}$$

for $\mu \in \mathcal{P}_p(\mathbb{R}^d)$, $q(\cdot) \in L^{p'}_\mu(\mathbb{R}^d)$ where p' is the conjugate exponent of $1 < p < +\infty$.

Proposition 11. *Let $1 < p < +\infty$ and $\tilde{S}_p$ be a generalized target. Assume Hypothesis 3.3 for F, and that* **(STLA)** *holds for the system. Then the minimum time function $\tilde{T}_p$ is a viscosity solution of the HJB equation $\mathcal{H}(\mu, D\tilde{T}_p(\mu)) = 0$, with Hamiltonian $\mathcal{H}$ defined in (12).*

Proof. By **(STLA)** assumption and Proposition 9, we get the continuity of $\tilde{T}_p$.

Let $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Given a function $v_\mu \in L^p(\mu)$ with $v_\mu(x) \in F(\mu, x)$ for μ-a.e. x, there exists an admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t \in [0,T]}$ represented by $\boldsymbol{\eta}$ satisfying Lemma 3.7(1). According to the Dynamic Programming Principle (Proposition 8(3)), for all $q \in D^+\tilde{T}_p(\mu)$ and for all $\pi_t \in \Pi(\mu, \mu_t)$

$$\begin{aligned}
0 \le & \frac{\tilde{T}_p(\mu_t) - \tilde{T}_p(\mu) + t}{t} \\
\le & 1 + \frac{1}{t} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \langle q(x), y - x \rangle d\pi_t(x, y) + \frac{1}{t} o\left(\left[\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \, d\pi_t(x, y)\right]^{1/p}\right) \\
\le & 1 + \iint_{\mathbb{R}^d \times \Gamma_T} \langle q(x), \frac{\gamma(t) - \gamma(0)}{t} \rangle d\boldsymbol{\eta}(x, \gamma) + \frac{1}{t} o\left(\left[\iint_{\mathbb{R}^d \times \Gamma_T} |\gamma(t) - \gamma(0)|^p \, d\boldsymbol{\eta}(x, \gamma)\right]^{1/p}\right)
\end{aligned}$$

where we chose $\pi_t = (e_t, e_0)\sharp\boldsymbol{\eta}$ in the last line. By letting $t \to 0^+$, Lemma 3.7(1) yields

$$0 \le 1 + \int_{\mathbb{R}^d} \langle q(x), v_\mu(x) \rangle \, d\mu(x).$$

By taking the infimum on $v_\mu \in L^p(\mu)$ s.t. $v_\mu(x) \in F(\mu, x)$ for μ-a.e. $x \in \mathbb{R}^d$, we have for all $\pi_t^{(\varepsilon)} \in \Pi(\mu, \mu_t)$

$$\mathcal{H}(\mu, q(\mu)) \le 0,$$

thus $\tilde{T}_p(\cdot)$ is a viscosity subsolution of $\mathcal{H}(\mu, D\tilde{T}_p(\mu)) = 0$.

On the other hand, from the Dynamic Programming Principle, for any $\varepsilon > 0$ we also get the existence of an admissible trajectory $\boldsymbol{\mu}^{(\varepsilon)} = \{\mu_t^{(\varepsilon)}\}_{t\in[0,T]}$, represented by $\boldsymbol{\eta}^{(\varepsilon)}$ satisfying Lemma 3.7(2), such that for all $p \in D^-\tilde{T}_p(\mu)$ and for all $\pi_t^{(\varepsilon)} \in \Pi(\mu, \mu_t^{(\varepsilon)})$

$$\begin{aligned}\varepsilon \geq & \frac{\tilde{T}_p(\mu_t^{(\varepsilon)}) - \tilde{T}_p(\mu) + t}{t} \\ \geq & 1 + \frac{1}{t}\int_{\mathbb{R}^d} \langle p(x), y - x\rangle d\pi_t^{(\varepsilon)}(x, y) + \frac{1}{t} o\left(\left[\iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p \, d\pi_t^{(\varepsilon)}(x,y)\right]^{1/p}\right) \\ \geq & 1 + \int_{\mathbb{R}^d\times\Gamma_T} \langle p(x), \frac{\gamma(t)-\gamma(0)}{t}\rangle d\boldsymbol{\eta}^{(\varepsilon)}(x,\gamma) + \frac{1}{t} o\left(\left[\iint_{\mathbb{R}^d\times\mathbb{R}^d} |\gamma(t)-\gamma(0)|^p \, d\boldsymbol{\eta}^{(\varepsilon)}(x,\gamma)\right]^{1/p}\right)\end{aligned}$$

where we chose $\pi_t^{(\varepsilon)} = (e_t, e_0)\sharp\boldsymbol{\eta}^{(\varepsilon)}$ in the last line. Consider the disintegration of $\boldsymbol{\eta}^{(\varepsilon)}$ with respect to the evaluation operator at time $t = 0$, i.e. $\boldsymbol{\eta}^{(\varepsilon)} = \mu \otimes \eta_x^{(\varepsilon)}$. By Filippov's measurable selection Theorem (see [4, Theorem 8.2.10, Corollary 8.2.13]) there exists $w_t \in L^p_\mu(\mathbb{R}^d)$ such that $w_t(x) \in F(\mu, x)$ for μ-a.e. $x \in \mathbb{R}^d$ and

$$\begin{aligned}\int_{\mathbb{R}^d}\left|w_t(x) - \int_{e_0^{-1}(x)} \frac{\gamma(t)-\gamma(0)}{t}\, d\eta_x^{(\varepsilon)}\right|^p d\mu(x) &= \int_{\mathbb{R}^d} d^p_{F(\mu,x)}\left(\int_{e_0^{-1}(x)} \frac{\gamma(t)-\gamma(0)}{t}\, d\eta_x^{(\varepsilon)}\right) d\mu(x) \\ &\leq \int_{\mathbb{R}^d\times\Gamma_T} d^p_{F(\mu,x)}\left(\frac{\gamma(t)-\gamma(0)}{t}\right) d\boldsymbol{\eta}^{(\varepsilon)}(x,\gamma),\end{aligned}$$

where we used Jensen's inequality in the last step. The last term vanishes as $t \to 0^+$ by Lemma 3.7(2), and therefore for t sufficiently small

$$\begin{aligned}\varepsilon \geq & 1 - \varepsilon\|p\|_{L^{p'}_\mu} + \int_{\mathbb{R}^d} \langle p(x), w_t(x)\rangle d\mu(x) + \frac{1}{t} o\left(t\cdot(\varepsilon + \|w_t\|_{L^p_\mu})\right) \\ \geq & -\mathcal{H}(\mu, p(\cdot)) - \varepsilon\|p\|_{L^{p'}_\mu} + \frac{1}{t} o\left(t\cdot(\varepsilon + \|w_t\|_{L^p_\mu})\right)\end{aligned}$$

Recalling that for μ-a.e. $x \in \mathbb{R}^d$ we have

$$w_t(x) \in F(\mu, x) \subseteq F(\delta_0, 0) + \mathrm{Lip}(F)(W_p(\mu, \delta_0) + |x|),$$

we have that $\|w_t\|_{L^p_\mu}$ is bounded uniformly w.r.t. t, ε, and so by letting $t \to 0^+$ and $\varepsilon \to 0^+$ we get

$$\mathcal{H}(\mu, p(\cdot)) \geq 0.$$

Thus $\tilde{T}_p(\cdot)$ is also a viscosity supersolution of $\mathcal{H}(\mu, D\tilde{T}_p(\mu)) = 0$. □

Remark 11. It seems natural that the combined use of **(STLA)** condition, of Grönwall estimate (see Theorem 3.6) together with a semiconcavity assumption for the distance to the target can be used to derive a semiconcavity estimate for the generalized minimum time function, pretty much as in the finite-dimensional case (see e.g. [6]). This topic will be subject of future investigation.

To conclude, we want to perform a comparison between the concept of (r, Q)-attainability of Definition 5.4 and the results obtained in [27].

Proposition 12. *Let $1 < p < +\infty$. Assume Hypothesis 3.3 for F and that the generalized target $\tilde{S}^\Phi_p$ is (r, Q)-attainable, with $Q = Q(t, s)$ differentiable and such* *that $\partial_t Q(0,\cdot) < 0$. Then $\sigma_\Phi(\cdot)$, defined in Definition 5.3, is a viscosity supersolution of $1+\mathcal{H}(\mu, D\sigma_\Phi(\mu))=0$, in particular*

$$\inf_{\substack{v\in L^p_\mu\\ v(x)\in F(\mu,x)}} \langle p(x), v_\mu(x)\rangle < 0, \tag{13}$$

for all $p(\cdot)\in D^-\sigma_\Phi(\mu)$, $\mu\in\mathcal{P}_p(\mathbb{R}^2)\setminus \tilde{S}^\Phi_p$.

Proof. Let $\mu\in\mathcal{P}_p(\mathbb{R}^d)\setminus\tilde{S}^\Phi_p$ and $\boldsymbol{\mu}\in\mathcal{A}^p_{[0,t]}(\mu)$ be an admissible trajectory represented by $\boldsymbol{\eta}$ according to Proposition 2 and Remark 2, with $0<t<r(\sigma_\Phi(q))$. We notice that, for all $p(\cdot)\in D^-\sigma_\Phi(\mu)$ we have

$$\sigma_\Phi(\mu_t)-\sigma_\Phi(\mu)\geq \int_{\mathbb{R}^d\times\Gamma_{[0,t]}} \langle p(x), \gamma(t)-\gamma(0)\rangle\, d\boldsymbol{\eta}(x,\gamma) + o\left(\|e_t-e_0\|_{L^p_{\boldsymbol{\eta}}}\right).$$

In particular, according to (r,Q)-attainability condition, for every $\varepsilon>0$ and $0<t\leq\varepsilon$ we can find $\boldsymbol{\mu}^{(\varepsilon)}\in\mathcal{A}^p_{[0,t]}(\mu)$ represented by $\boldsymbol{\eta}^{(\varepsilon)}$ such that

$$\frac{2Q(t,\sigma_\Phi(\mu))}{t}+\varepsilon\geq\frac{1}{t}\iint_{\mathbb{R}^d\times\Gamma_{[0,t]}}\langle p(x),\gamma(t)-\gamma(0)\rangle\, d\boldsymbol{\eta}^{(\varepsilon)}(x,\gamma)+\frac{1}{t}o\left(\|e_t-e_0\|_{L^p_{\boldsymbol{\eta}^{(\varepsilon)}}}\right),$$

where we used the previous estimate and divided by $t>0$. Recalling the hypothesis on Q and that without loss of generality we can consider $Q(0,\cdot)=0$, then by letting $t\to 0^+$ and $\varepsilon\to 0^+$ we obtain

$$0>2\partial_t Q(0,\sigma_\Phi(\mu))\geq -1-\mathcal{H}(\mu,p(\cdot)).$$

Here, we used the same arguments used in Proposition 11 when proving that $\tilde{T}_p$ is a viscosity supersolution of an HJB equation with Hamiltonian $\mathcal{H}$ defined in (12). We remind that in order to be a supersolution, the continuity of $\tilde{T}_p$ is not needed, and the l.s.c. is provided by Proposition 8.

So, we got that $\sigma_\Phi(\cdot)$ is a viscosity supersolution of $1+\mathcal{H}(\mu,D\sigma_\Phi(\mu))=0$, thus the conclusion noting that the expression above can be equivalently written as

$$\inf_{\substack{v\in L^p_\mu\\ v(x)\in F(\mu,x)}} \langle p(x), v_\mu(x)\rangle \leq 2\partial_t Q(0,\sigma_\Phi(\mu))<0.$$

□

Remark 12. Notice that the expression obtained above in (13) represents a natural counterpart of formula (27) in [27] (and subsequent extensions, see Corollary 3.5 and Remark 3.3 of [27]). Here, the Lipschitz continuity requested in [27] is replaced by an integrability assumption of the derivative $\partial_t Q$. Therefore, under these hypothesis, (r,Q)-attainability can be seen as a *sampled* form of the assumption in [27], indeed in our framework the decreasing condition (27) is checked along admissible trajectories only at time steps of size given by $r(\cdot)$.

Acknowledgments. G.C. has been partially 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.

The authors finally thank both the referees for their suggestions which definetely helped us to improve the presentation of the paper and to better frame it in the existing literature.

REFERENCES

[1] L. Ambrosio and J. Feng, On a class of first order Hamilton-Jacobi equations in metric spaces, *J. Differential Equations*, **256** (2014), 2194–2245.

[2] L. Ambrosio and W. Gangbo, Hamiltonian ODEs in the Wasserstein space of probability measures, *Comm. Pure Appl. Math.*, **61** (2008), 18–53.

[3] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2^{nd} edition, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[4] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Reprint of the 1990 edition [MR1048347], Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2009.

[5] Y. Averboukh, Viability theorem for deterministic mean field type control systems, *Set-Valued and Variational Analysis*, **26** (2018), 993–1008.

[6] P. Cannarsa and C. Sinestrari, Convexity properties of the minimum time function, *Calc. Var. Partial Differential Equations*, **3** (1995), 273–298.

[7] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston Inc., Boston, MA, 2004.

[8] M. Caponigro, M. Fornasier, B. Piccoli and E. Trélat, Sparse stabilization and control of alignment models, *Mathematical Models and Methods in Applied Sciences*, **25** (2015), 521–564.

[9] P. Cardaliaguet, *Notes on Mean Field Games*, 2013.

[10] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I*, Probability Theory and Stochastic Modeling, **83**, Springer International Publishing, 2018.

[11] G. Cavagnari, Regularity results for a time-optimal control problem in the space of probability measures, *Math. Control Relat. Fields*, **7** (2017), 213–233.

[12] G. Cavagnari, S. Lisini, C. Orrieri and G. Savaré, Lagrangian, Eulerian and Kantorovich formulations of multi-agent optimal control problems: equivalence and Gamma-convergence, preprint.

[13] G. Cavagnari, A. Marigonda, K. T. Nguyen and F. S. Priuli, Generalized control systems in the space of probability measures, *Set-Valued Var. Anal.*, **26** (2018), 663–691.

[14] G. Cavagnari, A. Marigonda and B. Piccoli, Superposition principle for differential inclusions, in *Large-Scale Scientific Computing. LSSC 2017. Lecture Notes in Computer Science* (eds. I. Lirkov and S. Margenov), **10665**, Springer, Cham, (2018), 201–209.

[15] G. Cavagnari, A. Marigonda and M. Quincampoix, Compatibility of state constraints and dynamics for multiagent control systems, preprint.

[16] J. Dolbeault, B. Nazaret and G. Savaré, A new class of transport distances between measures, *Calc. Var. Partial Differential Equations*, **34** (2009), 193–231.

[17] M. Duprez, M. Morancey and F. Rossi, Approximate and exact controllability of the continuity equation with a localized vector field, *SIAM J. Control Optim.*, **57** (2019), 1284–1311.

[18] M. Duprez, M. Morancey and F. Rossi, Minimal time for the continuity equation controlled by a localized perturbation of the velocity vector field, *Journal of Differential Equations*, **269** (2020), 82–124.

[19] K. Fan, Fixed-point and minimax theorems in locally convex topological linear spaces, *Proc. Nat. Acad. Sci. U. S. A.*, **38** (1952), 121–126.

[20] J. Feng and T. G. Kurtz, *Large Deviations for Stochastic Processes*, Mathematical Surveys and Monographs, **131**, American Mathematical Society, Providence, RI, 2006.

[21] J. Feng and T. Nguyen, Hamilton-Jacobi equations in space of measures associated with a system of conservation laws, *J. Math. Pures Appl. (9)*, **97** (2012), 318–390.

[22] M. Fornasier, S. Lisini, C. Orrieri and G. Savaré, Mean-field optimal control as gamma-limit of finite agent controls, *European J. Appl. Math.*, **30** (2019), 1153–1186.

[23] M. Fornasier, B. Piccoli and F. Rossi, Mean-field sparse optimal control, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **372** (2014), 20130400, 21 pp.

[24] W. Gangbo and A. Tudorascu, On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations, *Journal de Mathématiques Pures et Appliquées*, **125** (2019), 119–174.

[25] W. Gangbo and A. Święch, Optimal transport and large number of particles, *Discrete Contin. Dyn. Syst.*, **34** (2014), 1397–1441.

[26] C. Jimenez, A. Marigonda and M. Quincampoix, Optimal control of multiagent systems in the Wasserstein space, *Calc. Var. Partial Differential Equations*, **59** (2020), Paper No. 58, 45pp.

[27] M. I. Krastanov and M. Quincampoix, Local small time controllability and attainability of a set for nonlinear control system, *ESAIM Control Optim. Calc. Var.*, **6** (2001), 499–516.

[28] T. T. T. Le and A. Marigonda, Small-time local attainability for a class of control systems with state constraints, *ESAIM Control Optim. Calc. Var.*, **23** (2017), 1003–1021.

[29] A. Marigonda and S. Rigo, Controllability of some nonlinear systems with drift via generalized curvature properties, *SIAM J. Control and Optimization*, **53** (2015), 434–474.

[30] A. Marigonda and M. Quincampoix, Mayer control problem with probabilistic uncertainty on initial positions, *J. Differential Equations*, **264** (2018), 3212–3252.

[31] C. Orrieri, Large deviations for interacting particle systems: joint mean-field and small-noise limit, `arXiv:1810.12636`.

[32] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, **58**, American Mathematical Society, Providence, RI, 2003.

E-mail address: `giulia.cavagnari@polimi.it`
E-mail address: `antonio.marigonda@univr.it`