# GENERALIZED CONTROL SYSTEMS IN THE SPACE OF PROBABILITY MEASURES 

GIULIA CAVAGNARI, ANTONIO MARIGONDA, KHAI T. NGUYEN, AND FABIO S. PRIULI


#### Abstract

In this paper we formulate a time-optimal control problem in the space of probability measures. The main motivation is to face situations in finitedimensional control systems evolving deterministically where the initial position of the controlled particle is not exactly known, but can be expressed by a probability measure on $\mathbb{R}^{d}$. We propose for this problem a generalized version of some concepts from classical control theory in finite dimensional systems (namely, target set, dynamic, minimum time function...) and formulate an Hamilton-JacobiBellman equation in the space of probability measure solved by the generalized minimum time function. We prove also some representation results linking the classical concept to the corresponding generalized ones. The main tool used is a superposition principle, proved by Ambrosio, Gigli and Savaré, which provides a probabilistic representation of the solution of the continuity equation as a weighted superposition of absolutely continuous solutions of the characteristic system.


## 1. Introduction

Classical minimum time problem in finite-dimension deals with the minimization of the time needed to steer a point $x_{0} \in \mathbb{R}^{d}$ to a given closed subset $S$ of $\mathbb{R}^{d}$, called the target set, along the trajectories of a controlled dynamic of the form

$$
\left\{\begin{array}{l}
\dot{x}(t) \in F(x(t)), \quad t>0  \tag{1.1}\\
x(0)=x_{0}
\end{array}\right.
$$

where $F$ is a set-valued map from $\mathbb{R}^{d}$ to $\mathbb{R}^{d}$ whose value at each point denote the set of admissible velocities at that point.

In this way it is possible to define the minimum time function $T$ : given $x \in \mathbb{R}^{d}$, we define $T(x)$ to be the minimum time needed to steer such point to the target $S$ along trajectories of (1.1). The study of regularity property of $T$ is a central topic in optimal control theory and it has been extensively treated in literature. In particular, we refer to $[11,12]$ and to references therein, for recent results on the regularity of $T$ in the framework of differential inclusions.

Our study moves from the natural consideration that in many real applications we do not know exactly the starting position $x_{0} \in \mathbb{R}^{d}$ of the particle, and we can

[^0]express it only with some uncertainty. This happens even if we assume to have a deterministic evolution of the system.

A natural choice to face this situation is to model the uncertainty on the initial position by a probability measure $\mu_{0} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ on $\mathbb{R}^{d}$, looking to a new macroscopic control system made by a suitable superposition of a continuum of weighted solutions of the classical differential inclusion (1.1) starting from each point of the support of $\mu_{0}$ (microscopic point of view).

The time evolution of the macroscopic system in the space of probability measures, under suitable assumptions, can be thought as ruled by the continuity equation

$$
\left\{\begin{array}{l}
\partial_{t} \mu(t, x)+\operatorname{div}\left(v_{t}(x) \mu(t, x)\right)=0, \quad \text { for } t>0, x \in \mathbb{R}^{d}  \tag{1.2}\\
\mu(0, \cdot)=\mu_{0}
\end{array}\right.
$$

which represents the conservation of the total mass $\mu_{0}\left(\mathbb{R}^{d}\right)$ during the evolution. Here $v_{t}(x)$ is a suitable time-depending Eulerian vector field, representing the velocity of the mass crossing position $x$ at time $t$.

In order to reflect the original control system (1.1) at a microscopic level, a natural requirement on the vector field $v_{t}(\cdot)$ is to be a selection of the set-valued map $F(\cdot)$ : this means that the microscopic particles still obey the nonholonomic constraints coming from (1.1). On the other hand, since the conservation of the mass gives us the property $\mu\left(t, \mathbb{R}^{d}\right)=\mu_{0}\left(\mathbb{R}^{d}\right)$ for all $t$, we are entitled - according to our motivation - to say that the measure $\mu(t, \cdot)$ actually represents the probability distribution in the space $\mathbb{R}^{d}$ of the evolving particles at time $t$.

The analysis of (1.2) by mean of the superposition of ODEs of the form $\dot{x}(t)=$ $v(x(t))$, or $\dot{x}(t)=v(t, x(t))$, has been extensively studied in the past years by many authors: for a general introduction, an overview of known results and open problems, and a comprehensive bibliography, we refer to the recent survey [1]. The main issue in these problems is to study existence, uniqueness and regularity of the solution of (1.2), for $\mu_{0}$ in a suitable class of measures, when the vector field $v$ has low regularity and, hence, it does not ensure that the corresponding ODEs have a (possibly not unique) solution among absolutely continuous functions, for every initial data $x_{0}$. In this case, the solution of (1.2) provides existence and uniqueness not in a pointwise sense, but rather generically. However we will not address this problem in this paper.

In order to face control problems involving measures, we need first of all a coherent generalization of the target set $S \subseteq \mathbb{R}^{d}$. To this aim, we consider an observer which measures the average of certain quantities $\phi(\cdot) \in \Phi$ on the system, and consider as target set $\tilde{S}^{\Phi} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ all the probability measures representing states which make the result of the measurements nonpositive. If we take for instance $\Phi=\left\{d_{S}(\cdot)\right\}$, the generalized target in $\mathscr{P}\left(\mathbb{R}^{d}\right)$ turns out to be the set of all probability measures supported on $S$.

This choice seems to be the simplest possible in this framework and it results in a quite natural definition of generalized minimum time: we aim to minimize the time needed to steer an initial measure towards a measure in the generalized target, along solutions of (1.2) with the additional constraint $v(x) \in F(x)$ a.e. in $\mathbb{R}^{d}$. This can be viewed as a controlled version of (1.2).

The links between continuity equation (1.2) and optimal transport theory have been investigated recently by many authors. One can prove that suitable subsets of $\mathscr{P}\left(\mathbb{R}^{d}\right)$ can be endowed with a metric structure - the Wasserstein metric - whose absolutely continuous curves turn out to be precisely the solutions of (1.2). This has been applied to solve many variational problems, among which we recall optimal transport problems, asymptotic limit for gradient flows of integral functionals, and calculus of variation in infinite dimensional spaces. We refer to [3] and [23] for an introduction to the subject, and for generalizations from $\mathbb{R}^{d}$ to infinite dimensional metric spaces.

Our main results can be summarized as follows:

- a theorem of existence of time-optimal curves in the space of probability measures (Theorem 4.14);
- a comparison result between classical and generalized minimum time functions in some cases (Proposition 4.8);
a sufficient condition for the generalized minimum time function to be finite, with an upper estimate based on the initial data (Theorem 4.19);
- the proof that the generalized minimum time function is a viscosity solution in a suitable sense of an Hamilton-Jacobi-Bellman equation analoguos to the classical one (Theorem 5.10).

Recent works (see e.g. [2,17]) have treated the problem of viscosity solutions of Hamilton-Jacobi equations in the space of probability measures endowed with Wasserstein metric. Since classical minimum time function can be characterized as unique viscosity solution of Hamilton-Jacobi-Bellman equation, it would be interesting to investigate if it is possible to characterize in similar way the generalized minimum time function in this setting. Indeed, in this paper we just proved that the generalized minimum time function solves in a suitable viscosity sense a natural Hamilton-Jacobi-Bellman, which presents strong analogies with the finitedimensional case.

Related to such a problem, a further application could be the theory of mean field games [19, 20]. According to this theory, in games with a continuum of agents, having the same dynamics and the same performance criteria, the value function for an average player can be retrieved by solving an infinite dimensional HamiltonJacobi equation, coupled with the continuity equation describing how the mass of players evolves in time.

Further applications of our approach, that we plan to investigate in the next future, are in the direction of the classical control problems. For instance, in the study of control-affine systems of the form $\dot{x}=\sum_{i=1}^{m} u_{i} f_{i}(x)$, where $u_{i} \in[-1,1]$ are the controls and $f_{i}(\cdot)$ are given vector fields. In these systems, controllability depends on the Lie algebra generated by vector fields $f_{i}(\cdot)$. When these vector fields are rough, classical Lie brackets may not be available at every point of $\mathbb{R}^{d}$, but just in some set of full measure. This problem was treated in [21], leading to a definition of nonsmooth Lie brackets. However, a valid alternative might be to extend the given system to the measure-valued context and to choose the initial data of such generalized system in a suitable subclass of measure, in the spirit of [1].

Another application might be in the context of discontinuous feedback controls for general nonlinear control systems $\dot{x}=f(x, u)$. Here, the construction of stabilizing or nearly optimal controls $x \mapsto u(x)$ cannot be performed, even for smooth dynamics, among continuous controls [22]. However, it is possible to construct discontinuous feedback controls which are stabilizing or nearly optimal, and whose discontinuities are sufficiently tame to ensure the existence of Carathéodory solutions for the closed loop system $\dot{x}=f(x, u(x))$, the so-called patchy feedback controls $[4,5,10]$, but uniqueness only holds for a set of full measure of initial data.

The paper is structured as follows: in Section 2 we review some notion from measure theory, optimal transport, continuity equation, differential inclusions, and control theory. In Section 3 we introduce the generalized target, studying its topological and metric properties in the space of probability measures. Finally, in Section 4 we give two definitions of generalized minimum time functions, providing some comparison results between them and with the classical minimum time function, and we prove the Existence Theorem 4.14 and the Attainability Theorem 4.19. In Section 5 we prove that the generalized minimum time function solves in a suitable viscosity sense an Hamilton-Jacobi-Bellman equation.

## 2. Preliminaries

In this section we review some concepts from measure theory, optimal transport, and control theory.
2.1. Probability measure. Our main references for this part are [3] and [23].

Definition 2.1 (Probability measures). Let $X$ be a complete separable metric space, $\mathscr{P}(X)$ be the set of Borel probability measures on $X$. Since $\mathscr{P}(X)$ can be identified with a convex subset of the unitary ball of $\left(C_{b}^{0}(X)\right)^{\prime}$ (the dual space of the space of bounded continuous functions on $X$ ), we can equip $\mathscr{P}(X)$ with the weak* topology induced by $\left(C_{b}^{0}(X)\right)^{\prime}$. In particular, we say that a sequence of probability measures $\left\{\mu_{n}\right\}_{n \in \mathbb{N}}$ is $w^{*}$-convergent (or narrowly converges) to a probability measure $\mu \in$ $\mathscr{P}(X)$, and write $\mu_{n} \rightharpoonup^{*} \mu$, if and only if for every $f \in C_{b}^{0}(X)$ it holds

$$
\lim _{n \rightarrow \infty} \int_{X} f(x) d \mu_{n}(x)=\int_{X} f(x) d \mu(x)
$$

We will consider on $\mathscr{P}(X)$ the $\sigma$-algebra of Borel sets generated by the $w^{*}$-open subsets of $\mathscr{P}(X)$.
Definition 2.2 (Tightness). Let $X$ be a metric space and $\mathscr{K} \subseteq \mathscr{P}(X)$. We say that $\mathscr{K}$ is tight if for every $\varepsilon>0$ there exists a compact subset $K_{\varepsilon}$ of $X$ such that $\mu\left(X \backslash K_{\varepsilon}\right) \leq \varepsilon$ for every $\mu \in \mathscr{K}$. Every tight subset of $\mathscr{P}(X)$ is relatively compact in $\mathscr{P}(X)$. The converse is true if there exists an equivalent complete metric on $X$.
Definition 2.3 (Push forward). If $X, Y$ are separable metric spaces, $\mu \in \mathscr{P}(X)$, and $r: X \rightarrow Y$ is a Borel (or, more generally, $\mu$-measurable) map, we denote by $r \sharp \mu \in \mathscr{P}(Y)$ the push-forward of $\mu$ through $r$, defined by

$$
r \sharp \mu(B):=\mu\left(r^{-1}(B)\right), \text { for all Borel sets } B \subseteq Y \text {. }
$$

Equivalently, we have

$$
\int_{X} f(r(x)) d \mu(x)=\int_{Y} f(y) d r \sharp \mu(y),
$$

for every bounded (or $r \sharp \mu$-integrable) Borel function $f: Y \rightarrow \mathbb{R}$.

Proposition 2.4 (Properties of push forward). Let $X, Y, Z$ be separable metric spaces, $\mu \in \mathscr{P}(X)$, and let $r: X \rightarrow Y$ be a Borel map.
(1) If $\nu \in \mathscr{P}(X)$ satisfies $\nu \ll \mu$, then $r \sharp \nu \ll r \sharp \mu$.
(2) Given a Borel map $s: Y \rightarrow Z$, the following composition rule holds

$$
(s \circ r) \sharp \mu=s \sharp(r \sharp \mu) .
$$

(3) If $r \in C^{0}(X ; Y)$ then $r \sharp: \mathscr{P}(X) \rightarrow \mathscr{P}(Y)$ is continuous with respect to the narrow convergence and

$$
r(\operatorname{supp} \mu) \subseteq \operatorname{supp}(r \sharp \mu)=\overline{r(\operatorname{supp} \mu)} .
$$

(4) Let $\left\{r_{n}: X \rightarrow Y\right\}_{n \in \mathbb{N}}$ be a sequence of Borel maps uniformly convergent to $r$ on compact subsets of $X$, and let $\left\{\mu_{n}\right\}_{n \in \mathbb{N}} \subseteq \mathscr{P}(X)$ be a tight sequence narrowly convergent to $\mu$. Then if $r$ is continuous, we have that $r_{n} \sharp \mu_{n} \rightharpoonup^{*}$ $r \sharp \mu$.

Proof. See [3], Chapter 5, Section 2.
Definition 2.5 ( $p$-moment). Let $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right), p \geq 1$. We say that $\mu$ has finite $p$-moment if

$$
\mathrm{m}_{p}(\mu):=\int_{\mathbb{R}^{d}}|x|^{p} d \mu(x)<+\infty .
$$

Equivalently, we have that $\mu$ has $p$-moment finite if and only if for every $x_{0} \in \mathbb{R}^{d}$ we have

$$
\int_{\mathbb{R}^{d}}\left|x-x_{0}\right|^{p} d \mu(x)<+\infty
$$

We denote by $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ the subset of $\mathscr{P}\left(\mathbb{R}^{d}\right)$ consisting of probability measures with finite $p$-moment.

Definition 2.6 (Uniform integrability). Let $\mathscr{K} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right), g: \mathbb{R}^{d} \rightarrow[0,+\infty]$ be a Borel function. We say that
(1) $g$ is uniformly integrable with respect to $\mathscr{K}$ if

$$
\lim _{k \rightarrow \infty} \sup _{\mu \in \mathscr{K}} \int_{\left\{x \in \mathbb{R}^{d}: g(x)>k\right\}} g(x) d \mu(x)=0 .
$$

(2) the set $\mathscr{K}$ has uniformly integrable p-moments, $p \geq 1$, if $|x|^{p}$ is uniformly integrable with respect to $\mathscr{K}$.

Lemma 2.7 (Uniform integrability criterion). Let $\mathscr{K}=\left\{\mu_{n}\right\}_{n \in \mathbb{N}} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right), p \geq 1$, $\mu_{n} \rightharpoonup^{*} \mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$. Then the set $\mathscr{K}$ has uniformly integrable $p$-moments if and only if

$$
\lim _{n \rightarrow \infty} \int_{\mathbb{R}^{d}} f(x) d \mu_{n}(x)=\int_{\mathbb{R}^{d}} f(x) d \mu(x)
$$

for every continuous function $f: \mathbb{R}^{d} \rightarrow \mathbb{R}$ such that there exist $a, b \geq 0$ and $x_{0} \in \mathbb{R}^{d}$ with $|f(x)| \leq a+b\left|x-x_{0}\right|^{p}$ for every $x \in \mathbb{R}^{d}$.

Proof. See Lemma 5.1.7 of [3].

### 2.2. Optimal transport and Wasserstein distances.

Definition 2.8 (Wasserstein distance). Given $\mu_{1}, \mu_{2} \in \mathscr{P}\left(\mathbb{R}^{d}\right), p \geq 1$, we define the p-Wasserstein distance between $\mu_{1}$ and $\mu_{2}$ by setting

$$
\begin{equation*}
W_{p}\left(\mu_{1}, \mu_{2}\right):=\left(\inf \left\{\iint_{\mathbb{R}^{d} \times \mathbb{R}^{d}}\left|x_{1}-x_{2}\right|^{p} d \pi\left(x_{1}, x_{2}\right): \pi \in \Pi\left(\mu_{1}, \mu_{2}\right)\right\}\right)^{1 / p} \tag{2.1}
\end{equation*}
$$

where the set of admissible transport plans $\Pi\left(\mu_{1}, \mu_{2}\right)$ is defined by

$$
\begin{aligned}
\Pi\left(\mu_{1}, \mu_{2}\right):=\left\{\pi \in \mathscr{P}\left(\mathbb{R}^{d} \times \mathbb{R}^{d}\right): \begin{array}{l}
\pi\left(A_{1} \times \mathbb{R}^{d}\right)=\mu_{1}\left(A_{1}\right) \\
\\
\pi\left(\mathbb{R}^{d} \times A_{2}\right)=\mu_{2}\left(A_{2}\right)
\end{array}\right. \\
\text { for all } \left.\mu_{i} \text {-measurable sets } A_{i}, i=1,2\right\} .
\end{aligned}
$$

We also denote with $\Pi_{o}^{p}\left(\mu_{1}, \mu_{2}\right)$ the subset of $\Pi\left(\mu_{1}, \mu_{2}\right)$ consisting of optimal transport plans, i.e. the set of all plans $\pi$ for which the infimum in (2.1) is attained. We will also use the notation $\Pi_{o}\left(\mu_{1}, \mu_{2}\right)$ when the context makes clear which distance $W_{p}$ is being considered.
Proposition 2.9. $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ endowed with the $p$-Wasserstein metric $W_{p}(\cdot, \cdot)$ is a complete separable metric space. Moreover, given a sequence $\left\{\mu_{n}\right\}_{n \in \mathbb{N}} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ and $\mu \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, we have that the following are equivalent
(1) $\lim _{n \rightarrow \infty} W_{p}\left(\mu_{n}, \mu\right)=0$,
(2) $\mu_{n} \rightharpoonup^{*} \mu$ and $\left\{\mu_{n}\right\}_{n \in \mathbb{N}}$ has uniformly integrable p-moments.

Proof. See Proposition 7.1.5 in [3].
In the following Proposition we summarize some properties of the Wasserstein distance (see chapter 6 in [23] or section 7.1 in [3]).
Proposition 2.10. The Wasserstein distances defined above satisfies the following properties:

- Metric character. $W_{p}$ is a pseudo-distance on $\mathscr{P}\left(\mathbb{R}^{d}\right)$, i.e. it satisfies the axioms of the distance, but it can assume the value $+\infty$. Namely, for all $\mu_{0}, \mu_{1}, \mu_{2} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ we have
(i) $W_{p}\left(\mu_{0}, \mu_{1}\right) \geq 0$, and $W_{p}\left(\mu_{0}, \mu_{1}\right)=0$ if and only if $\mu_{0}=\mu_{1}$ (positive definiteness);
(ii) $W_{p}\left(\mu_{0}, \mu_{1}\right)=W_{p}\left(\mu_{1}, \mu_{0}\right)$ (symmetry);
(iii) $W_{p}\left(\mu_{0}, \mu_{2}\right) \leq W_{p}\left(\mu_{0}, \mu_{1}\right)+W_{p}\left(\mu_{1}, \mu_{2}\right)$ (triangle inequality).

When restricted to $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right), W_{p}$ is actually finite, so it is a metric.

- Topological properties. The topology induced by $W_{p}$ on $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ is finer (equivalently stronger) than or equal to the narrow one.
- Lower semicontinuity. If $\mu_{n}^{0} \rightharpoonup^{*} \mu^{0}, \mu_{n}^{1} \rightharpoonup^{*} \mu^{1}$ in $\mathscr{P}\left(\mathbb{R}^{d}\right)$ when $n \rightarrow+\infty$, then

$$
W_{p}\left(\mu^{0}, \mu^{1}\right) \leq \liminf _{n \rightarrow+\infty} W_{p}\left(\mu_{n}^{0}, \mu_{n}^{1}\right)
$$

Proposition 2.11 (Monge-Kantorovich duality). Given $\mu_{1}, \mu_{2} \in \mathscr{P}\left(\mathbb{R}^{d}\right), p \geq 1$, the following dual representation holds

$$
\begin{align*}
& W_{p}^{p}\left(\mu_{1}, \mu_{2}\right)=  \tag{2.2}\\
& =\sup \left\{\int_{\mathbb{R}^{d}} \varphi\left(x_{1}\right) d \mu_{1}\left(x_{1}\right)+\int_{\mathbb{R}^{d}} \psi\left(x_{2}\right) d \mu_{2}\left(x_{2}\right): \begin{array}{l}
\varphi, \psi \in C_{b}^{0}\left(\mathbb{R}^{d}\right) \\
\quad \begin{array}{l}
\varphi\left(x_{1}\right)+\psi\left(x_{2}\right) \leq\left|x_{1}-x_{2}\right|^{p} \\
\text { for } \mu_{i}-\text { a.e. } x_{i} \in \mathbb{R}^{d}
\end{array}
\end{array}\right\} .
\end{align*}
$$

Proof. See Theorem 6.1.1 in [3].
2.3. Continuity equation. For this part the main reference is [3].

Definition 2.12 (Continuity equation). Given $\tau>0$, a Borel family of probability measures $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, \tau]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ and a Borel map $v:[0, \tau] \times \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ (we will write also $\left.v_{t}(x)=v(t, x)\right)$, we say that $\boldsymbol{\mu}$ solves the continuity equation

$$
\begin{equation*}
\partial_{t} \mu_{t}+\operatorname{div}\left(v_{t} \mu_{t}\right)=0, \tag{2.3}
\end{equation*}
$$

if for every $\varphi \in C_{c}^{\infty}\left(\mathbb{R}^{d}\right)$ there holds

$$
\frac{d}{d t} \int_{\mathbb{R}^{d}} \varphi(x) d \mu_{t}(x)=\int_{\mathbb{R}^{d}}\left\langle v_{t}(x), \nabla \varphi(x)\right\rangle d \mu_{t}(x),
$$

in the sense of distributions on $] 0, \tau[$.
According to Lemma 8.1.2 in [3], if the above $v$ satisfies

$$
\begin{equation*}
\int_{0}^{\tau} \int_{\mathbb{R}^{d}}\left|v_{t}(x)\right| d \mu_{t}(x) d t<+\infty \tag{2.4}
\end{equation*}
$$

then there exists a curve $\tilde{\mu}:[0, \tau] \rightarrow \mathscr{P}\left(\mathbb{R}^{d}\right)$ which is continuous with respect to narrow convergence and such that $\tilde{\mu}(t)=\mu_{t}$ for $\mathscr{L}^{1}$-a.e. $t \in(0, \tau)$, i.e. each solution of the continuity equation admits a unique narrowly continuous representative.

The following gluing lemma will be also used.
Lemma 2.13. Let $T_{1}, T_{2}>0$ be given. For $i=1,2$, assume that $\boldsymbol{\mu}^{i}=\left\{\mu_{t}^{i}\right\}_{t \in\left[0, T_{i}\right]}$ are narrowly continuous families of probability measures on $\mathbb{R}^{d}$, and $v^{i}:\left[0, T_{i}\right] \times \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ are Borel maps such that $\mu_{\mid t=T_{1}}^{1}=\mu_{\mid t=0}^{2}$ and

$$
\left\{\begin{array}{l}
\partial_{t} \mu_{t}^{i}+\operatorname{div}\left(v_{t}^{i} \mu_{t}^{i}\right)=0 \\
\int_{0}^{T_{i}} \int_{\mathbb{R}^{d}}\left|v_{t}^{i}(x)\right| d \mu_{t}^{i}(x) d t<+\infty
\end{array} \quad i=1,2\right.
$$

Then if we set

$$
\left(\mu_{t}, v_{t}\right)= \begin{cases}\left(\mu_{t}^{1}, v_{t}^{1}\right), & \text { for } 0 \leq t \leq T_{1} \\ \left(\mu_{t-T_{1}}^{2}, v_{t-T_{1}}^{2}\right), & \text { for } T_{1} \leq t \leq T_{1}+T_{2}\end{cases}
$$

we have that $\boldsymbol{\mu}:=\left\{\mu_{t}\right\}_{t \in\left[0, T_{1}+T_{2}\right]}$ solves the continuity equation $\partial_{t} \mu_{t}+\operatorname{div}\left(v_{t} \mu_{t}\right)=0$.
Proof. See Lemma 4.4 in [16].
Theorem 2.14 (Superposition principle). Let $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ be a solution of the continuity equation $\partial_{t} \mu_{t}+\operatorname{div}\left(v_{t} \mu_{t}\right)=0$ for a suitable Borel vector field $v:[0, T] \times$ $\mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ satisfying

$$
\int_{0}^{T} \int_{\mathbb{R}^{d}} \frac{\left|v_{t}(x)\right|}{1+|x|} d \mu_{t}(x) d t<+\infty
$$

Then there exists a probability measure $\boldsymbol{\eta} \in \mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{T}\right)$, with $\Gamma_{T}=C^{0}\left([0, T] ; \mathbb{R}^{d}\right)$ endowed with the sup norm, such that
(i) $\boldsymbol{\eta}$ is concentrated on the pairs $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$ such that $\gamma$ is an absolutely continuous solution of

$$
\left\{\begin{array}{l}
\dot{\gamma}(t)=v_{t}(\gamma(t)), \quad \text { for } \mathscr{L}^{1} \text {-a.e } t \in(0, T) \\
\gamma(0)=x,
\end{array}\right.
$$

(ii) for all $t \in[0, T]$ and all $\varphi \in C_{b}^{0}\left(\mathbb{R}^{d}\right)$ we have

$$
\int_{\mathbb{R}^{d}} \varphi(x) d \mu_{t}(x)=\iint_{\mathbb{R}^{d} \times \Gamma_{T}} \varphi(\gamma(t)) d \boldsymbol{\eta}(x, \gamma)
$$

Conversely, given any $\boldsymbol{\eta}$ satisfying (i) above and defined $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ as in (ii) above, we have that $\partial_{t} \mu_{t}+\operatorname{div}\left(v_{t} \mu_{t}\right)=0$ and $\mu_{\mid t=0}=\gamma(0) \sharp \boldsymbol{\eta}$.
Proof. See Theorem 5.8 in [9] and Theorem 8.2.1 in [3].
2.4. Differential inclusions and classical minimum time. For this part, our main references are [7] and [8].

Definition 2.15 (Standing Assumption). We will say that a set-valued function $F: \mathbb{R}^{d} \rightrightarrows \mathbb{R}^{d}$ satisfies the assumption $\left(F_{j}\right), j=0,1,2,3$ if the following hold true
$\left(F_{0}\right) F(x) \neq \emptyset$ is compact and convex for every $x \in \mathbb{R}^{d}$, moreover $F(\cdot)$ is continuous with respect to the Hausdorff metric, i.e. given $x \in X$, for every $\varepsilon>0$ there exists $\delta>0$ such that $|y-x| \leq \delta$ implies $F(y) \subseteq F(x)+B(0, \varepsilon)$ and $F(x) \subseteq F(y)+B(0, \varepsilon)$.
( $F_{1}$ ) $F(\cdot)$ has linear growth, i.e. there exist nonnegative constants $L_{1}$ and $L_{2}$ such that $F(x) \subseteq \overline{B\left(0, L_{1}|x|+L_{2}\right)}$ for every $x \in \mathbb{R}^{d}$,
$\left(F_{2}\right) F(\cdot)$ is Lipschitz continuous with respect to the Hausdorff metric, i.e., there exists $L>0, L \in \mathbb{R}$, such that for all $x, y \in \mathbb{R}^{d}$ it holds

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

$\left(F_{3}\right) F(\cdot)$ is bounded, i.e. there exist $M>0$ such that $\|y\| \leq M$ for all $x \in \mathbb{R}^{d}$, $y \in F(x)$.
Theorem 2.16. Under assumption $\left(F_{0}\right)$ and $\left(F_{1}\right)$, the differential inclusion

$$
\begin{equation*}
\dot{x}(t) \in F(x(t)) \tag{2.5}
\end{equation*}
$$

has at least one Carathéodory solution defined in $[0,+\infty[$ for every initial data $x(0)$ in $\mathbb{R}^{d}$, i.e., an absolutely continuous function $x(\cdot)$ satisfying (2.5) for a.e. $t \geq 0$.
Moreover, the set of trajectories of the differential inclusions (2.5) is closed in the topology of uniform convergence.
Proof. See e.g. Theorem 2 p. 97 in [7] and Theorem 1.11 p. 186 in Chapter 4 of [14].

The following simple classical lemma will be used.
Lemma 2.17 (A priori estimate on differential inclusions). Assume $\left(F_{0}\right)$ and $\left(F_{1}\right)$. Let $K \subset \mathbb{R}^{d}$ be compact and $T>0$ and set $\|K\|=\max _{y \in K}\|y\|$. Then, for all Carathéodory solutions $\gamma:[0, T] \rightarrow \mathbb{R}^{d}$ of (2.5) we have
(i) forward estimate: if $\gamma(0) \in K$ then $|\gamma(t)| \leq\left(\|K\|+L_{2} T\right) e^{L_{1} T}$ for all $t \in$ $[0, T]$;
(ii) backward estimate: if $\gamma(T) \in K$ then $|\gamma(t)| \leq\left(\|K\|+L_{2} T\right) e^{L_{1} T}$ for all $t \in[0, T]$.
where $L_{1}, L_{2}$ are the constants in $\left(F_{1}\right)$.
Proof. Recalling that $\dot{\gamma}(s) \in F(\gamma(s))$ for a.e. $s \in[0, T]$ and that $F(\gamma(s)) \subseteq$ $\overline{B\left(0, L_{1}|\gamma(s)|+L_{2}\right)}$, we have

$$
|\gamma(t)| \leq|\gamma(0)|+\int_{0}^{t}|\dot{\gamma}(s)| d s \leq\|K\|+L_{2} T+L_{1} \int_{0}^{t}|\gamma(s)| d s .
$$

According to Gronwall's inequality, we then have $|\gamma(t)| \leq\left(\|K\|+L_{2} T\right) e^{L_{1} t}$, whence (i) follows.

Next, we define $w(t)=\gamma(T-t)$ and observe that $w$ is a solution of $\dot{w}(t) \in-F(w(t))$. Since $-F(\cdot)$ still satisfies $\left(F_{0}\right)$ and $\left(F_{1}\right)$ and $w(0) \in K$, the previous analysis implies

$$
|\gamma(t)|=|w(T-t)| \leq\left(\|K\|+L_{2} T\right) e^{L_{1}(T-t)},
$$

whence (ii) follows.
Definition 2.18 (Weak invariance). We say that $S \subseteq \mathbb{R}^{d}$ is weakly invariant for $F(\cdot)$ if for every $x \in S$ there exists a Carathéodory solution $x(\cdot)$ of $(2.5)$, defined in $[0,+\infty[$, such that $x(0)=x$ and $x(t) \in S$ for every $t \geq 0$.

For conditions on $S$ and $F$ ensuring weak invariance, we refer to Theorem 2.10 in Chapter 4 of [14].

Definition 2.19 (Minimum time function). Let $F(\cdot)$ be a set-valued function satisfying $\left(F_{0}\right), S$ be a nonempty closed subset of $\mathbb{R}^{d}$. We define the minimum time function $T: \mathbb{R}^{d} \rightarrow[0,+\infty]$ by setting
$T\left(x_{0}\right)=\inf \left\{\tau>0: \exists x(\cdot) \in A C\left([0, \tau] ; \mathbb{R}^{d}\right)\right.$ s.t. $\left.\begin{array}{l}\dot{x}(t) \in F(x(t)) \text { for a.e. } t \in[0, \tau], \\ x(0)=x_{0}, x(\tau) \in S\end{array}\right\}$,
where by convention $\inf \emptyset=+\infty$.
Definition 2.20. Let $X$ be a set, $A \subseteq X$. The indicator function of $A$ is the function $I_{A}: X \rightarrow\{0,+\infty\}$ defined as $I_{A}(x)=0$ for all $x \in A$ and $I_{A}(x)=+\infty$ for all $x \notin A$. The characteristic function of $A$ is the function $\chi_{A}: X \rightarrow\{0,1\}$ defined as $\chi_{A}(x)=1$ for all $x \in A$ and $\chi_{A}(x)=0$ for all $x \notin A$.

Definition 2.21 (Support function). Let $X$ be a Banach space, $X^{\prime}$ be its topological dual, $A \subseteq X$ be nonempty. We define the support function to $A$ at $x^{*} \in X^{\prime}$ by setting

$$
\begin{equation*}
\sigma_{A}\left(x^{*}\right):=\sup _{x \in A}\left\langle x^{*}, x\right\rangle_{X^{\prime}, X} . \tag{2.6}
\end{equation*}
$$

It turns out that $\sigma_{A}\left(x^{*}\right)=\sigma_{\overline{c o}(A)}\left(x^{*}\right)$ for every $x^{*} \in X^{\prime}$ and that $\sigma_{A}: X^{\prime} \rightarrow \mathbb{R} \cup\{+\infty\}$ is convex and lower semicontinuous.

## 3. Generalized targets

In this section we propose some suitable generalizations of the classical target set that can be used in our framework.

Definition 3.1 (Generalized targets). Let $p \geq 1, \Phi \subseteq C^{0}\left(\mathbb{R}^{d}, \mathbb{R}\right)$ such that the following property holds
$\left(T_{E}\right)$ there exists $x_{0} \in \mathbb{R}^{d}$ with $\phi\left(x_{0}\right) \leq 0$ for all $\phi \in \Phi$.

We define the generalized targets $\tilde{S}^{\Phi}$ and $\tilde{S}_{p}^{\Phi}$ as follows

$$
\begin{aligned}
& \tilde{S}^{\Phi}:=\left\{\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right): \int_{\mathbb{R}^{d}} \phi^{+}(x) d \mu(x) \leq \int_{\mathbb{R}^{d}} \phi^{-}(x) d \mu(x) \text { for all } \phi \in \Phi\right\}, \\
& \tilde{S}_{p}^{\Phi}:=\tilde{S}^{\Phi} \cap \mathscr{P}_{p}\left(\mathbb{R}^{d}\right),
\end{aligned}
$$

where $\phi^{+}(x):=\max \{0, \phi(x)\}, \phi^{-}(x):=\max \{0,-\phi(x)\}$ for all $x \in \mathbb{R}^{d}$, thus $\phi=$ $\phi^{+}-\phi^{-}$and $\phi^{+}, \phi^{-} \geq 0$.

We define also the generalized distance from $\tilde{S}_{p}^{\Phi}$ as

$$
\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\cdot):=\inf _{\mu \in \tilde{S}_{p}^{\Phi}} W_{p}(\cdot, \mu) .
$$

Notice that $\tilde{S}_{p}^{\Phi} \neq \emptyset$ because $\delta_{x_{0}} \in \tilde{S}_{p}^{\Phi}$, hence $\tilde{S}^{\Phi} \neq \emptyset$. The 1-Lipschitz continuity of $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\cdot)$ follows from the structure of metric space: indeed let $\mu, \nu \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, and fix $\varepsilon>0$. Choose $\sigma_{\nu} \in \tilde{S}_{p}^{\Phi}$ such that $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\nu) \geq W_{p}\left(\nu, \sigma_{\nu}\right)-\varepsilon$. Then we have by triangular inequality

$$
\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\mu)-\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\nu) \leq W_{p}\left(\mu, \sigma_{\nu}\right)-W_{p}\left(\nu, \sigma_{\nu}\right)+\varepsilon \leq W_{p}(\mu, \nu)+\varepsilon .
$$

By switching the role of $\mu, \nu$ and letting $\varepsilon \rightarrow 0^{+}$, we obtain the desired Lipschitz continuity property.

For further use, we will say that $\Phi$ satisfies property $\left(T_{p}\right)$ with $p \geq 0$ if the following holds true
$\left(T_{p}\right)$ for all $\phi \in \Phi$ there exist $A_{\phi}, C_{\phi}>0$ such that $\phi(x) \geq A_{\phi}|x|^{p}-C_{\phi}$.
Remark 3.2. Given $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$, observe that if $\phi^{+} \in L_{\mu}^{1}$ we have

$$
\int_{\mathbb{R}^{d}} \phi(x) d \mu(x) \leq 0 \text { if and only if } \int_{\mathbb{R}^{d}} \phi^{+}(x) d \mu(x) \leq \int_{\mathbb{R}^{d}} \phi^{-}(x) d \mu(x),
$$

but in general only the last inequality make sense. Thus, defined $\tilde{S}^{\Phi}\left(\right.$ and $\left.\tilde{S}_{p}^{\Phi}\right)$ as above, if $\phi^{+} \in L_{\mu}^{1}$ for all $\mu \in \tilde{S}^{\Phi}$ (resp. $\mu \in \tilde{S}_{p}^{\Phi}$ ), then

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

In this case, roughly speaking, a 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 $\phi$. The results of the measurements are the average of the quantities $\phi$ with respect to the measure $\mu_{t}$ 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 ).
Remark 3.3. Given a nonempty and closed $S \subseteq \mathbb{R}^{d}$ and $\left.\left.\alpha \in\right] 0,1\right]$, a natural choice for $\Phi$ can be for example $\Phi=\left\{d_{S}(\cdot)-\alpha\right\}$. In this case, a measure belonging to $\tilde{S}^{\Phi}$ corresponds to the state of a particle which is on $S$ with probability $1-\alpha$. If $\alpha=0$, i.e. $\Phi=\left\{d_{S}(\cdot)\right\}$, then $\tilde{S}^{\Phi}$ reduces to the set of all probability measures supported on $S$.

The following proposition establishes some straightforward properties of the generalized targets.

Proposition 3.4 (Properties of the generalized targets). Let $p \geq 1$ and $\Phi \subseteq$ $C^{0}\left(\mathbb{R}^{d}, \mathbb{R}\right)$ be such that $\left(T_{E}\right)$ holds. Then:
(1) $\tilde{S}^{\Phi}$ and $\tilde{S}_{p}^{\Phi}$ are convex;
(2) $\tilde{S}^{\Phi}$ is $w^{*}$-closed in $\mathscr{P}\left(\mathbb{R}^{d}\right)$;
(3) $\tilde{S}_{p}^{\Phi}$ is closed in $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ endowed with the $p$-Wasserstein metric $W_{p}(\cdot, \cdot)$;
(4) for every $\mu \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ we have $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\mu)=0$ if and only if $\mu \in \tilde{S}_{p}^{\Phi}$;
(5) if there exists $\bar{\phi} \in \Phi, A, C \in \mathbb{R}$ and $p \geq 1$ such that $\bar{\phi}(x) \geq A|x|^{p}-C$, then $\tilde{S}^{\Phi}=\tilde{S}_{p}^{\Phi}$ is compact in the $w^{*}$-topology and in the $W_{p}$-topology. In particular, this holds if $\Phi$ satisfies property $\left(T_{p}\right)$.

Proof.
(1) The convexity property is trivial from the definition.
(2) Let $\left\{\mu_{n}\right\}_{n \in \mathbb{N}}$ be a sequence in $\tilde{S}^{\Phi}$, and $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ be such that $\mu_{n} \rightharpoonup^{*} \mu$. Take a sequence $\left\{\varphi_{k}\right\}_{k \in \mathbb{N}} \subseteq C^{0}\left(\mathbb{R}^{d} ;[0,1]\right)$ of maps such that $\varphi_{k}(x)=1$ for $x \in \overline{B(0, k)}$ and $\varphi_{k}(x)=0$ for $x \in B(0, k+1)$.

Since $\mu_{n} \in \tilde{S}^{\Phi}$, we have

$$
\int_{\mathbb{R}^{d}} \phi^{+}(x) d \mu_{n} \leq \int_{\mathbb{R}^{d}} \phi^{-}(x) d \mu_{n}, \quad \forall \phi \in \Phi .
$$

Fix $\varepsilon>0$. By Monotone Convergence Theorem, we have that there exists $k_{\varepsilon} \in \mathbb{N}$ such that for all $k \geq k_{\varepsilon}$

$$
\int_{\mathbb{R}^{d}} \phi^{-}(x) d \mu_{n} \leq \int_{\mathbb{R}^{d}} \varphi_{k}(x) \phi^{-}(x) d \mu_{n}+\varepsilon,
$$

thus

$$
\int_{\mathbb{R}^{d}} \varphi_{k}(x) \phi^{+}(x) d \mu_{n} \leq \int_{\mathbb{R}^{d}} \varphi_{k}(x) \phi^{-}(x) d \mu_{n}+\varepsilon,
$$

and by letting $n \rightarrow+\infty$, recalling the weak* convergence of $\mu_{n}$ to $\mu$, we have

$$
\int_{\mathbb{R}^{d}} \varphi_{k}(x) \phi^{+}(x) d \mu \leq \int_{\mathbb{R}^{d}} \varphi_{k}(x) \phi^{-}(x) d \mu+\varepsilon,
$$

hence, by letting $\varepsilon \rightarrow 0^{+}$we have

$$
\int_{\mathbb{R}^{d}} \phi^{+}(x) d \mu \leq \int_{\mathbb{R}^{d}} \phi^{-}(x) d \mu
$$

and so $\mu \in \tilde{S}^{\Phi}$.
(3) It follows from the fact that convergence in $W_{p}(\cdot, \cdot)$ implies $w^{*}$-convergence, and that $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ endowed with the $p$-Wasserstein metric $W_{p}(\cdot, \cdot)$ is a complete separable metric space according to Proposition 2.9.
(4) It is obvious that if $\mu \in \tilde{S}_{p}^{\Phi}$ then $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\mu)=0$. Conversely, if $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\mu)=0$ there exists a sequence $\left\{\mu_{n}\right\}_{n \in \mathbb{N}} \subseteq \tilde{S}_{p}^{\Phi}$ such that $\lim _{n \rightarrow \infty} W_{p}\left(\mu_{n}, \mu\right)=0$, and, by the closedness of $\tilde{S}_{p}^{\Phi}$, we conclude that $\mu \in \tilde{S}_{p}^{\Phi}$.
(5) Given $p \geq 1$, trivially we have that $\tilde{S}_{p}^{\Phi} \subseteq \tilde{S}^{\Phi}$. Conversely, given $\mu \in \tilde{S}^{\Phi}$, we have that $\bar{\phi}^{-}$is bounded, thus the integral of $\bar{\phi}$ w.r.t. $\mu$ is well defined. So

$$
\int_{\mathbb{R}^{d}}\left(A|x|^{p}-C\right) d \mu \leq \int_{\mathbb{R}^{d}} \bar{\phi}(x) d \mu \leq 0
$$

where $\bar{\phi}, A, C, p$, are as in the assumptions. Thus for all $\mu \in \tilde{S}^{\Phi}$ we have

$$
\int_{\mathbb{R}^{d}}|x|^{p} d \mu \leq \frac{C}{A}<+\infty,
$$

hence $\mu \in \tilde{S}_{p}^{\Phi}$. So all the measures in $\tilde{S}_{p}^{\Phi}=\tilde{S}^{\Phi}$ have uniformly bounded $p$-moments. Hence, if $\left\{\mu_{n}\right\}_{n \in \mathbb{N}} \subseteq \tilde{S}^{\Phi}$ and $\mu_{n} \rightharpoonup^{*} \mu$, by the $w^{*}$-closure of $\tilde{S}^{\Phi}$ we have that $\mu \in \tilde{S}^{\Phi}=\tilde{S}_{p}^{\Phi}$ and it has finite $p$-moment. Thus, the family $\left\{\mu_{n}\right\}_{n \in \mathbb{N}}$ has equiuniformly integrable $p$-moments, and $W_{p}\left(\mu_{n}, \mu\right) \rightarrow 0$ by Proposition 2.9. This means that the $w^{*}$-topology and $W_{p}$-topology coincide on $\tilde{S}^{\Phi}=\tilde{S}_{p}^{\Phi}$, which turns out to be tight, according to Remark 5.1.5 in [3], and $w^{*}$-closed, hence $w^{*}$-compact and $W_{p}$-compact.

Given a nonempty closed set $S \subseteq \mathbb{R}^{d}$, and set $\Phi=\left\{d_{S}(\cdot)\right\}$, a natural problem is to express the generalized distance $\tilde{d}_{\tilde{S}_{p}^{\Phi}}(\cdot)$ in terms of $d_{S}(\cdot)$. More generally, we give the following definition.

Definition 3.5 (Classical counterpart of generalized target). Let $p \geq 1$ and $\Phi \subseteq$ $C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ satisfying assumptions of Definition 3.1. Given a set $S \subseteq \mathbb{R}^{d}$, we say that
(1) $S$ is a classical counterpart of the generalized target $\tilde{S}^{\Phi}$ if the following equality holds

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

(2) $S$ is a classical counterpart of the generalized target $\tilde{S}_{p}^{\Phi}$ if the following equality holds

$$
\tilde{S}_{p}^{\Phi}=\left\{\mu \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right): \operatorname{supp} \mu \subseteq S\right\}
$$

Proposition 3.6 (Existence, uniqueness and properties of the classical counterpart). Let $p \geq 1$ and $\Phi \subseteq C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ satisfying assumptions of Definition 3.1. Then
(1) if $\tilde{S}^{\Phi}$ admits a classical counterpart $S$, then $\tilde{S}_{p}^{\Phi}$ admits $S$ as a classical counterpart for all $p \geq 1$.
(2) if $S, S^{\prime}$, are classical counterparts of the generalized targets $\tilde{S}^{\Phi}, \tilde{S}_{p}^{\Phi}$, respectively, then $S=S^{\prime}$;
(3) if $S$ is a classical counterpart of $\tilde{S}^{\Phi}$ or of $\tilde{S}_{p}^{\Phi}$, then $S$ is closed;
(4) if $S$ is the classical counterpart of $\tilde{S}^{\Phi}$ then $\phi(x) \leq 0$ for all $\phi \in \Phi, x \in S$;
(5) if $\phi(x) \geq 0$ for all $\phi \in \Phi$ and $x \in \mathbb{R}^{d}$ then the set

$$
S:=\left\{x \in \mathbb{R}^{d}: \phi(x)=0 \text { for all } \phi \in \Phi\right\}
$$

is the classical counterpart of $\tilde{S}^{\Phi}$ and of $\tilde{S}_{p}^{\Phi}$ (uniqueness follows from item (2) above);
(6) if $S$ is the classical counterpart of $\tilde{S}^{\Phi}$, then there exists a representation of $\tilde{S}^{\Phi}$ as $\tilde{S}^{\Phi^{\prime}}$, where $\phi^{\prime}(x) \geq 0 \forall x \in \mathbb{R}^{d}$, $\phi^{\prime} \in \Phi^{\prime}$. In particular we can take $\Phi^{\prime}=\left\{d_{S}\right\}$ and we have $\tilde{S}^{\Phi}=\tilde{S}^{\left\{d_{S}\right\}}$ and $\tilde{S}_{p}^{\Phi}=\tilde{S}_{p}^{\left\{d_{S}\right\}}$, i.e., we can replace $\Phi$ with the set $\left\{d_{S}\right\}$;
(7) if for every $\phi \in \Phi$ we have either $\phi(x) \geq 0$ or $\phi(x) \leq 0$ for all $x \in \mathbb{R}^{d}$, then $\tilde{S}^{\Phi}$ and $\tilde{S}_{p}^{\Phi}$ admit as classical counterpart the set

$$
S=\bigcap_{\phi \in \Phi}\left\{x \in \mathbb{R}^{d}: \phi(x) \leq 0\right\}=\bigcap_{\phi \in \Phi^{+}}\left\{x \in \mathbb{R}^{d}: \phi(x)=0\right\},
$$

where $\Phi^{+}=\left\{\phi \in \Phi: \phi(x) \geq 0\right.$ for all $\left.x \in \mathbb{R}^{d}\right\}$, and if $\Phi^{+}=\emptyset$ we set $S=\mathbb{R}^{d}$.

Proof.
(1) By definition, for all $p \geq 1$ we have

$$
\begin{aligned}
\tilde{S}_{p}^{\Phi}:=\tilde{S}^{\Phi} \cap \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) & =\left\{\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right): \operatorname{supp} \mu \subseteq S\right\} \cap \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) \\
& =\left\{\mu \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right): \operatorname{supp} \mu \subseteq S\right\} .
\end{aligned}
$$

(2) Let $S$ and $S^{\prime}$ be two classical counterparts of $\tilde{S}^{\Phi}$ and of $\tilde{S}_{p}^{\Phi}$, respectively. For every $x \in S$ we have that $\delta_{x} \in \tilde{S}_{p}^{\Phi} \subseteq \tilde{S}^{\Phi}$ for all $p \geq 1$, hence we must have also $x \in S^{\prime}$ since $S^{\prime}$ is a classical counterpart of the generalized target $\tilde{S}_{p}^{\Phi}$. So $S \subseteq S^{\prime}$. By reversing the roles of $S$ and $S^{\prime}$ we obtain $S=S^{\prime}$.
(3) Let $S$ be the classical counterpart of $\tilde{S}^{\Phi}$ (the proof is analoguos for $\tilde{S}_{p}^{\Phi}$ ). Let $\left\{x_{n}\right\}_{n \in \mathbb{N}} \subseteq S$ be s.t. $x_{n} \rightarrow \bar{x}$ for some $\bar{x} \in \partial S$. By contradiction, let us suppose $\bar{x} \notin S$, thus $\delta_{\bar{x}} \notin \tilde{S}^{\Phi}$. Then there exists $\bar{\phi} \in \Phi$ s.t. $\bar{\phi}(\bar{x})>0$, and thus for $n$ sufficiently large we have $\bar{\phi}\left(x_{n}\right)>0$ by continuity of $\bar{\phi}$. It follows that $\delta_{x_{n}} \notin \tilde{S}^{\Phi}$ for $n$ sufficiently large, thus $x_{n} \notin S$ by definition of classical counterpart and we get a contradiction.
(4) Immediate by definition of generalized target and of classical counterpart, in fact we have $\delta_{\bar{x}} \in \tilde{S}^{\Phi}$ for all $\bar{x} \in S$.
(5) Obviously we have

$$
\left\{\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right): \operatorname{supp} \mu \subseteq S\right\} \subseteq \tilde{S}^{\Phi}
$$

Let us prove the other inclusion. Note that by hypothesis $\phi=\phi^{+}$and $\phi^{-}=0$ for every $\phi \in \Phi$, hence

$$
\tilde{S}^{\Phi}=\left\{\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right): \int_{\mathbb{R}^{d}} \phi(x) d \mu(x)=0 \text { for all } \phi \in \Phi\right\}
$$

Let $\mu \in \tilde{S}^{\Phi}$, then

$$
\int_{\mathbb{R}^{d}} \phi(x) d \mu(x)=0 \quad \forall \phi \in \Phi
$$

i.e. $\phi(x)=0$ for $\mu$-a.e. $x \in \mathbb{R}^{d}, \forall \phi \in \Phi$, i.e. $\phi(x)=0$ for all $x \in \operatorname{supp} \mu$, $\forall \phi \in \Phi$. Thus $\operatorname{supp} \mu \subseteq S$. By item (1), $S$ is the classical counterpart also of $\tilde{S}_{p}^{\Phi}$.
(6) Let us prove that $\tilde{S}^{\left\{d_{S}\right\}}=\tilde{S}^{\Phi}$. First $\tilde{S}^{\left\{d_{S}\right\}} \subseteq \tilde{S}^{\Phi}$, in fact if $\mu \in \tilde{S}^{\left\{d_{S}\right\}}$ then $\mu\left(\mathbb{R}^{d} \backslash S\right)=0$, and so $\mu \in \tilde{S}^{\Phi}$ by definition of classical counterpart. Moreover, $\tilde{S}^{\{d S\}} \supseteq \tilde{S}^{\Phi}$, in fact if $\mu \in \tilde{S}^{\Phi}$, then $\operatorname{supp} \mu \subseteq S$ and it follows that $\int_{\mathbb{R}^{d}} d_{S}(x) d \mu(x)=0$, thus $\mu \in \tilde{S}^{\left\{d_{S}\right\}}$.
(7) By item (1), it is sufficient to prove that $S$ is the classical counterpart of $\tilde{S}^{\Phi}$. Assume that $\Phi^{+}=\emptyset$. This means that $\phi(x) \leq 0$ for all $x \in \mathbb{R}^{d}$ and for all $\phi \in \Phi$, thus $\phi=-\phi^{-}$and $\phi^{+}=0$ for all $\phi \in \Phi$. In this case we have that $\tilde{S}^{\Phi}=\mathscr{P}\left(\mathbb{R}^{d}\right)$ since for every $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ we have

$$
\int_{\mathbb{R}^{d}} \phi(x) d \mu(x) \leq 0
$$

Thus we have trivially $S=\mathbb{R}^{d}$.

Suppose now $\Phi^{+} \neq \emptyset$. Clearly, every measure supported in $S$ belongs to $\tilde{S}^{\Phi}$, since all the elements of $\Phi$ are nonpositive on $S$, i.e. $\tilde{S}^{\left\{d_{S}\right\}} \subseteq \tilde{S}^{\Phi}$. Conversely, let $\mu \in \tilde{S}^{\Phi}$ and by contradiction assume that there exists $\bar{x} \in$ $\operatorname{supp} \mu \backslash S$. This implies that there exists an open neighborhood $A$ of $\bar{x}$ such that $\mu(A)>0$, and an element $\phi \in \Phi^{+}$such that $\phi(\bar{x}) \neq 0$. By continuity of $\phi$, we can assume that $\phi>0$ on the whole of $A$, thus, recalling that $\phi(x) \geq 0$ for all $x \in \mathbb{R}^{d}$, we obtain

$$
\int_{\mathbb{R}^{d}} \phi(x) d \mu(x) \geq \int_{A} \phi(x) d \mu(x)>0
$$

contradicting the fact that $\mu \in \tilde{S}^{\Phi}$.

## Example 3.7.

(1) In general $\tilde{S}^{\Phi}$ may fail to possess a classical counterpart: in $\mathbb{R}$, take $\Phi=\{\phi\}$ where $\phi: \mathbb{R} \rightarrow \mathbb{R}, \phi(x):=|x+1|-1$ (notice that $\phi^{-}$is bounded). Then if $\tilde{S}^{\Phi}$ or $\tilde{S}_{p}^{\Phi}$ admitted a classical counterpart $S$, we should have $S \subseteq[-2,0]$ by item (4) of the Proposition above. Define $\mu_{0}:=\frac{1}{2}\left(\delta_{-1}+\delta_{1}\right)$. Thus we have $\mu_{0} \in \tilde{S}_{p}^{\Phi}$, in fact $\int_{\mathbb{R}^{2}} \phi(x) d \mu_{0}(x)=0$, but supp $\mu_{0}=\{-1,1\} \nsubseteq S$ for any possible $S$. So neither $\widetilde{S}^{\Phi}$ nor $\tilde{S}_{p}^{\Phi}$ admit a classical counterpart.
(2) The converse of item (7) of Proposition 3.6 is not true: in $\mathbb{R}$, take $\Phi=$ $\left\{\phi_{1}, \phi_{2}, \phi_{3}\right\}$ where $\phi_{i}: \mathbb{R} \rightarrow \mathbb{R}, i=1,2,3$ are defined to be $\phi_{1}(x)=$ $\max \{x, 0\}, \phi_{2}(x)=\min \{-x, 0\}, \phi_{3}(x)=x$. Then both $\tilde{S}_{p}^{\Phi}$ and $\tilde{S}^{\Phi}$ admits $S$ as their classical counterpart, with $S=]-\infty, 0]$, but $\phi_{3}$ can change its sign.

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

Proposition 3.8 (Comparison with classical distance). Let $p \geq 1, \mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, $\Phi \subseteq C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ satisfying assumptions of Definition 3.1, and set

$$
C:=\left\{x \in \mathbb{R}^{d}: \phi(x) \leq 0 \text { for all } \phi \in \Phi\right\} .
$$

Then
(1) $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \leq\left\|d_{C}\right\|_{L_{\mu_{0}}^{p}}$,
(2) if there exists $\tilde{\phi}(\cdot) \in \Phi$ such that $\tilde{\phi}(x) \geq 0$ for all $x \in \mathbb{R}^{d}$, then $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \geq$ $\left\|d_{D}\right\|_{L_{\mu_{0}}^{p}}$, where

$$
D:=\left\{x \in \mathbb{R}^{d}: \tilde{\phi}(x)=0\right\} .
$$

(3) if $\tilde{S}_{p}^{\Phi}$ admits a classical counterpart $S$, then $C=S$ and $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right)=\left\|d_{S}\right\|_{L_{\mu_{0}}^{p}}$, moreover $\tilde{d}_{\tilde{S}_{p}^{\text {d }}}^{p}: \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) \rightarrow[0,+\infty[$ is convex.
Proof. Clearly, according to assumption of Definition 3.1 we have $C \neq \emptyset$.
(1) If $\left\|d_{C}\right\|_{L^{p}\left(\mu_{0}\right)}=+\infty$ then there is nothing to prove. So let us assume that $\left\|d_{C}\right\|_{L^{p}\left(\mu_{0}\right)}<+\infty$.

Define the multifunction

$$
G(x):=\left\{y \in \mathbb{R}^{d}:|x-y|=d_{C}(x)\right\} \cap C=\overline{B\left(x, d_{C}(x)\right)} \cap C
$$

Since the map $f: \mathbb{R}^{d} \times \mathbb{R}^{d} \rightarrow \mathbb{R}$ defined by setting $f(x, y):=|x-y|-d_{C}(x)$ is continuous, we have that $G(\cdot)$ has closed graph in $\mathbb{R}^{d} \times \mathbb{R}^{d}$, and in particular $G(\cdot)$ is measurable. According to Theorem 8.1.3 in [8], there exists a Borel map $g: \mathbb{R}^{d} \rightarrow C$ such that $|x-g(x)|=d_{C}(x)$ for all $x \in \mathbb{R}^{d}$ (that is $g(x) \in G(x)$ for all $\left.x \in \mathbb{R}^{d}\right)$.

We define $\nu_{0}:=g \sharp \mu_{0}$ and prove now that $\nu_{0} \in \tilde{S}_{p}^{\Phi}$. Indeed, since $g(x) \in C$ for all $x \in \mathbb{R}^{d}$, we have $(\phi \circ g)^{+}=0$ for all $\phi(\cdot) \in \Phi$, thus

$$
\int_{\mathbb{R}^{d}} \phi(x) d g \sharp \mu_{0}(x)=\int_{\mathbb{R}^{d}} \phi(g(x)) d \mu_{0}(x) \leq 0, \text { for all } \phi(\cdot) \in \Phi,
$$

whence $\nu_{0} \in \tilde{S}^{\Phi}$.
It remains to prove that the $p$-moment of $\nu_{0}$ is finite. Owing to

$$
\begin{aligned}
\left(\int_{\mathbb{R}^{d}}|x|^{p} d \nu_{0}\right)^{1 / p} & =\left(\int_{\mathbb{R}^{d}}|g(x)|^{p} d \mu_{0}\right)^{1 / p} \\
& =\|g\|_{L^{p}\left(\mu_{0}\right)} \leq\|g-\operatorname{Id}\|_{L^{p}\left(\mu_{0}\right)}+\|\operatorname{Id}\|_{L^{p}\left(\mu_{0}\right)}
\end{aligned}
$$

we have to prove that the sum in the right hand side is finite. But $\mu_{0} \in$ $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ implies $\|\mathrm{Id}\|_{L^{p}\left(\mu_{0}\right)}<+\infty$ and $|g(x)-x|=d_{C}(x)$ holds by construction, so that $\|g-\mathrm{Id}\|_{L^{p}\left(\mu_{0}\right)}=\left\|d_{C}\right\|_{L^{p}\left(\mu_{0}\right)}<+\infty$. Therefore, we conclude $\nu_{0} \in \tilde{S}_{p}^{\Phi}$ and we have

$$
\begin{aligned}
\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) & \leq W_{p}\left(\mu_{0}, \nu_{0}\right) \leq\left(\iint_{\mathbb{R}^{d} \times \mathbb{R}^{d}}|x-y|^{p} d(\operatorname{Id} \times g) \sharp \mu_{0}\right)^{1 / p} \\
& =\left(\int_{\mathbb{R}^{d}}|x-g(x)|^{p} d \mu_{0}\right)^{1 / p}=\left(\int_{\mathbb{R}^{d}} d_{C}^{p}(x) d \mu_{0}\right)^{1 / p}
\end{aligned}
$$

as desired.
(2) Let us now assume that there exists $\tilde{\phi}(\cdot) \in \Phi$ such that $\tilde{\phi}(x) \geq 0$ for all $x \in \mathbb{R}^{d}$ (i.e. $\tilde{\phi}^{-}=0$ ) and prove that $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \geq\left\|d_{D}\right\|_{L_{\mu_{0}}^{p}}$. Let $\left\{\varphi_{n}\right\}_{n \in \mathbb{N}} \subset$ $C_{c}^{0}\left(\mathbb{R}^{d} ;[0,1]\right)$ be such that

$$
\varphi_{n}(x)=\left\{\begin{array}{l}
1, \text { if } x \in \overline{B(0, n)} \\
0, \text { if } x \notin B(0, n+1)
\end{array}\right.
$$

Set $\psi_{2}^{n}(y)=\varphi_{n}(y) \tilde{\phi}(y)$ and $\psi_{1}^{n}(x)=\varphi_{n}(x) d_{D}^{p}(x)$, hence we have $\psi_{1}^{n}, \psi_{2}^{n} \in$ $C_{b}^{0}\left(\mathbb{R}^{d}\right)$. Given $\theta \in \tilde{S}_{p}^{\Phi}$, we notice that for $\theta$-a.e. $y \in \mathbb{R}^{d}$ we must have $\tilde{\phi}(y)=0$, and so $y \in D$ thus for $\theta$-a.e. $y \in \mathbb{R}^{d}$ and $\mu_{0}$-a.e. $x \in \mathbb{R}^{d}$ it holds

$$
\psi_{1}^{n}(x)+\psi_{2}^{n}(y)=\varphi_{n}(x) d_{D}^{p}(x) \leq d_{D}^{p}(x) \leq|x-y|^{p}
$$

So, according to Kantorovich duality (2.2), we have

$$
\begin{aligned}
W_{p}^{p}\left(\mu_{0}, \theta\right) & =\sup _{\substack{\psi_{1}, \psi_{2} \in C_{b}^{0}\left(\mathbb{R}^{d}\right) \\
\psi_{1}(x)+\psi_{2}(y) \leq|x-y|^{p}}}\left\{\int_{\mathbb{R}^{d}} \psi_{1}(x) d \mu_{0}(x)+\int_{\mathbb{R}^{d}} \psi_{2}(y) d \theta(y)\right\} \\
& \geq \int_{\mathbb{R}^{d}} \varphi_{n}(x) d_{D}^{p}(x) d \mu_{0}(x)
\end{aligned}
$$

Since $\left\{\psi_{1}^{n}(\cdot)\right\}_{n \in \mathbb{N}} \subseteq C_{b}^{0}\left(\mathbb{R}^{d}\right)$ is an increasing sequence of nonnegative functions pointwise convergent to $d_{D}^{p}(\cdot)$, by letting $n \rightarrow+\infty$ and applying the Monotone Convergence Theorem we obtain

$$
W_{p}^{p}\left(\mu_{0}, \theta\right) \geq \int_{\mathbb{R}^{d}} d_{D}^{p}(x) d \mu_{0}(x)
$$

for all $\theta \in \tilde{S}_{p}^{\Phi}$.
(3) The equality $C=S$ is trivial: from item (4) in Proposition 3.6 we have $S \subseteq C$, moreover if $\mu$ is a measure supported in $C$ we have that $\mu \in \tilde{S}_{p}^{\Phi}$, since all the functions of $\Phi$ are nonpositive on $C$, thus $C \subseteq S$, and so equality holds. By item (1) above we have already $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \leq\left\|\bar{d}_{S}\right\|_{L_{\mu_{0}}^{p}}$. By item (6) in Proposition 3.6, we have $\tilde{S}_{p}^{\Phi}=\tilde{S}_{p}^{\left\{d_{C}\right\}}$, hence by applying item (2) above with $D=C=S$ and $\tilde{\phi}=d_{S}$ we obtain $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \geq\left\|d_{S}\right\|_{L_{\mu_{0}}^{p}}$, thus equality holds. Finally, the last statement is trivial, and it follows from the fact that

$$
\tilde{d}_{\tilde{S}_{p}^{\Phi}}^{p}(\mu)=\int_{\mathbb{R}^{d}} d_{C}^{p}(x) d \mu,
$$

is linear in $\mu$.

Without the assumption of existence of a classical counterpart for $\tilde{S}_{p}^{\Phi}$, the inequality $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right) \leq\left\|d_{C}\right\|_{L_{\mu_{0}}^{p}}$ may be strict.
Example 3.9. In $\mathbb{R}$, take $\Phi=\left\{\phi_{1}, \phi_{2}, \phi_{3}\right\}$ where $\phi_{i}: \mathbb{R} \rightarrow \mathbb{R}$ are defined by

$$
\phi_{1}(x)=|x-1|-1, \quad \phi_{2}(x)=|x+1|-1, \quad \phi_{3}(x)=\left|x\left(x^{2}-1\right)\right| .
$$

Define also $\mu_{0}=\frac{1}{2}\left(\delta_{-1}+\delta_{1}\right)$. For any $x \in \mathbb{R}$, we have $\phi_{i}(x) \geq-1$ for $i=1,2$ and $\phi_{3}(x) \geq 0$ (thus $\phi^{-}$is uniformly bounded for $i=1,2,3$ ), moreover

$$
\begin{aligned}
C: & \left\{x \in \mathbb{R}: \phi_{i} \leq 0, \text { for } i=1,2,3\right\}=\{0\} \\
= & \left\{x \in \mathbb{R}: \phi_{i}=0, \text { for } i=1,2,3\right\}, \\
& \int_{\mathbb{R}} \phi_{i}(x) d \mu_{0}(x)=0, \quad i=1,2,3,
\end{aligned}
$$

hence, $\mu_{0} \in \tilde{S}_{p}^{\Phi}$ for all $p \geq 1$, thus $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\mu_{0}\right)=0$. However, since $d_{C}^{p}(x)=|x|^{p}$, we have

$$
\int_{\mathbb{R}} d_{C}^{p}(x) d \mu_{0}(x)=1>0
$$

We notice that $\tilde{S}_{p}^{\Phi}$ does not admit a classical counterpart: indeed if a classical counterpart would exist, it would be reduced to $C=\{0\}$, however $\mu_{0} \in \tilde{S}_{p}^{\Phi} \subseteq \tilde{S}^{\Phi}$ and supp $\mu_{0} \nsubseteq C$, thus no classical counterpart may exist.

Without the $p$-th power, the generalized distance in the case of the Proposition 3.8 above may fail to be convex.

Example 3.10. Let $p>1$. In $\mathbb{R}^{2}$, consider $P=(0,0), Q_{1}=(1,0), Q_{2}=\left(0,2^{1 / p}\right)$. Set $S=\{P\}, \Phi=\left\{d_{S}(\cdot)\right\}$, hence $\tilde{S}_{p}^{\Phi}:=\left\{\delta_{P}\right\}$, and define $\nu_{\lambda}=\lambda \delta_{Q_{1}}+(1-\lambda) \delta_{Q_{2}}$, $\lambda \in[0,1]$. By Proposition 3.8, we have

$$
\tilde{d}_{\tilde{S}_{p}^{\Phi}}^{p}\left(\nu_{\lambda}\right)=W_{p}^{p}\left(\delta_{P}, \nu_{\lambda}\right)=\lambda+2(1-\lambda)=2-\lambda,
$$

whence $\tilde{d}_{\tilde{S}_{p}^{\Phi}}\left(\nu_{\lambda}\right)=\sqrt[p]{2-\lambda}$, which is not convex.
In the metric space $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ endowed with $W_{p}$-distance, another concept of convexity can be given, related more to the metric structure rather than to the linear one.

Given any product space $X^{N}(N \geq 1)$, in the following we denote with pri${ }^{i}: X^{N} \rightarrow$ $X$ the projection on the $i$-th component, i.e., $\operatorname{pr}^{i}\left(x_{1}, \ldots, x_{N}\right)=x_{i}$.

Definition 3.11 (Geodesics). Given a curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0,1]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, we say that it is a (constant speed) geodesic if for all $0 \leq s \leq t \leq 1$ we have

$$
W_{p}\left(\mu_{s}, \mu_{t}\right)=(t-s) W_{p}\left(\mu_{0}, \mu_{1}\right)
$$

In this case, we will also say that the curve $\boldsymbol{\mu}$ is a geodesic connecting $\mu_{0}$ and $\mu_{1}$.
Theorem 3.12 (Characterization of geodesics). Let $\mu_{0}, \mu_{1} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ and let $\pi \in$ $\Pi_{o}^{p}\left(\mu_{0}, \mu_{1}\right)$ be an optimal transport plan between $\mu_{0}$ and $\mu_{1}$, i.e.

$$
W_{p}^{p}\left(\mu_{0}, \mu_{1}\right)=\iint_{\mathbb{R}^{d} \times \mathbb{R}^{d}}\left|x_{1}-x_{2}\right|^{p} d \pi\left(x_{1}, x_{2}\right) .
$$

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

$$
\begin{equation*}
\mu_{t}:=\left((1-t) \mathrm{pr}^{1}+t \mathrm{pr}^{2}\right) \sharp \pi \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) \tag{3.1}
\end{equation*}
$$

is a (constant speed) geodesic connecting $\mu_{0}$ and $\mu_{1}$.
Conversely, any (constant speed) geodesic $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0,1]}$ connecting $\mu_{0}$ and $\mu_{1}$ admits the representation (3.1) for a suitable plan $\pi \in \Pi_{o}^{p}\left(\mu_{0}, \mu_{1}\right)$.

Proof. See Theorem 7.2.2 in [3].
Definition 3.13 (Geodesically and strongly geodesically convex sets). A subset $A \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ is said to be
(1) geodesically convex if for every pair of measures $\mu_{0}, \mu_{1}$ in $A$, there exists a geodesic connecting $\mu_{0}$ and $\mu_{1}$ which is contained in $A$.
(2) strongly geodesically convex if for every pair of measures $\mu_{0}, \mu_{1}$ in $A$ and for every admissible transport plan $\pi \in \Pi\left(\mu_{0}, \mu_{1}\right)$, the curve $t \mapsto \mu_{t}$ defined by (3.1) 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 Section 9.1 in [3] for further details.

Remark 3.14. Notice that, even if the notations does not highlight this fact, the notions of geodesic and geodesical convexity depend on the exponent $p$ which has been fixed.

Proposition 3.15 (Strong geodesic convexity of $\left.\tilde{S}_{p}^{\Phi}\right)$. Let $p \geq 1$. Assume that all the elements of $\Phi$ are continuous and convex and satisfy $\int_{\mathbb{R}^{d}} \phi^{+}(x) d \mu(x)<+\infty$ for all $\mu \in \tilde{S}_{p}^{\Phi}$, Then the generalized target $\tilde{S}_{p}^{\Phi}$ is strongly geodesically convex.

Proof. We notice that the assumptions implies that $\int_{\mathbb{R}^{d}} \phi(x) d \mu(x)<+\infty$ is well defined and nonpositive for all $\mu \in \tilde{S}_{p}^{\Phi}$ and $\phi \in \Phi$. Let $\mu_{0}, \mu_{1} \in \tilde{S}_{p}^{\Phi}$ and let $\pi \in \Pi\left(\mu_{0}, \mu_{1}\right)$ be an admissible transport plan between $\mu_{0}$ and $\mu_{1}$. Consider the corresponding curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0,1]}$ defined by (3.1), and fix $t \in[0,1]$. We have for every $\phi(\cdot) \in \Phi$

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

since $\operatorname{pr}^{i} \sharp \pi$ are the marginal measures of $\pi$, which belong to $\tilde{S}_{p}^{\Phi}$. The conclusion follows from the arbitrariness of $\phi(\cdot) \in \Phi$.
Remark 3.16. In particular, the above result holds for $\Phi:=\left\{d_{S}(\cdot)-\alpha\right\}$ 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 $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$.

Proposition 3.17 (Semiconcavity of $\tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}$ ). Let $\tilde{S}_{2}^{\Phi}$ be the generalized target in $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ corresponding to $\Phi \subseteq C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$. Then the square of the generalized distance satisfies the following global semiconcavity inequality for every $\mu_{0}, \mu_{1} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ and every $t \in[0,1]$

$$
\tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{t}\right) \geq(1-t) \tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{0}\right)+t \tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{1}\right)-t(1-t) W_{2}^{2}\left(\mu_{0}, \mu_{1}\right),
$$

where $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0,1]}$ is any constant speed geodesic for $W_{2}$ joining $\mu_{0}$ and $\mu_{1}$.
Proof. Owing to Theorem 7.3.2 in [3], we have that for any measure $\sigma \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ the function $\mu \mapsto W_{2}^{2}(\mu, \sigma)$ is semiconcave along geodesics, with semiconcavity constant independent by $\sigma$, i.e. it satisfies for every $t \in[0,1]$

$$
W_{2}^{2}\left(\mu_{t}, \sigma\right)+t(1-t) W_{2}^{2}\left(\mu_{0}, \mu_{1}\right) \geq(1-t) W_{2}^{2}\left(\mu_{0}, \sigma\right)+t W_{2}^{2}\left(\mu_{1}, \sigma\right)
$$

By passing to the infimum on $\sigma \in \tilde{S}_{2}^{\Phi}$, we have

$$
\tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{t}\right)+t(1-t) W_{2}^{2}\left(\mu_{0}, \mu_{1}\right) \geq(1-t) \tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{0}\right)+t \tilde{d}_{\tilde{S}_{2}^{\Phi}}^{2}\left(\mu_{1}\right),
$$

whence the conclusion follows.

## 4. Generalized minimum time problem

In this section we define a suitable notion of minimum time function, modeled on the finite-dimensional case.

Definition 4.1 (Admissible curves). Let $F: \mathbb{R}^{d} \rightrightarrows \mathbb{R}^{d}$ be a set-valued function, $I=[a, b]$ a compact interval of $\mathbb{R}, \alpha, \beta \in \mathscr{P}\left(\mathbb{R}^{d}\right)$. We say that a Borel family of probability measures $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in I}$ is an admissible trajectory (curve) defined in I
for the system $\Sigma_{F}$ joining $\alpha$ and $\beta$, if there exists a family of Borel vector-valued measures $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in I} \subseteq \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that
(1) $\boldsymbol{\mu}$ is a narrowly continuous solution in the distributional sense of

$$
\partial_{t} \mu_{t}+\operatorname{div} \nu_{t}=0
$$

with $\mu_{\mid t=a}=\alpha$ and $\mu_{\mid t=b}=\beta$.
(2) $J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})<+\infty$, where $J_{F}(\cdot)$ is defined as
$J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu}):= \begin{cases}\int_{I} \int_{\mathbb{R}^{d}}\left(1+I_{F(x)}\left(\frac{\nu_{t}}{\mu_{t}}(x)\right)\right) d \mu_{t}(x) d t, & \text { if }\left|\nu_{t}\right| \ll \mu_{t} \text { for a.e. } t \in I, \\ +\infty, & \text { otherwise. }\end{cases}$
where $I_{F(x)}$ is the indicator function of the set $F(x)$, i.e., $I_{F(x)}(\xi)=0$ for all $\xi \in F(x)$ and $I_{F(x)}(\xi)=+\infty$ for all $\xi \notin F(x)$.
In this case, we will also shortly say that $\boldsymbol{\mu}$ is driven by $\boldsymbol{\nu}$.
Remark 4.2. The finiteness of $J(\boldsymbol{\mu}, \boldsymbol{\nu})$ forces the elements of $\boldsymbol{\nu}$ to have the form $\nu_{t}=v_{t} \mu_{t}$ for a vector field $v_{t} \in L_{\mu_{t}}^{1}$ for a.e. $t \in I$, and moreover we have $v_{t}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$ and a.e. $t \in I$. When $J_{F}(\cdot)$ is finite, this value expresses the time needed by the system $\Sigma_{F}$ to steer $\alpha$ to $\beta$ along the trajectory $\boldsymbol{\mu}$ with family of velocity vector fields $v=\left\{v_{t}\right\}_{t \in I}$.

In view of the superposition principle stated at Theorem 2.14, we can give the following alternative definition.
Definition 4.3 (Admissible curves (alternative definition)). Let $F: \mathbb{R}^{d} \rightrightarrows \mathbb{R}^{d}$ be a set-valued function satisfying $\left(F_{1}\right), I=[a, b]$ a compact interval of $\mathbb{R}, \alpha, \beta \in \mathscr{P}\left(\mathbb{R}^{d}\right)$. We say that a Borel family of probability measures $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in I}$ is an admissible trajectory (curve) defined in I for the system $\Sigma_{F}$ joining $\alpha$ and $\beta$, if there exist a probability measure $\boldsymbol{\eta} \in \mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{I}\right)$ and a Borel vector field $v: I \times \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ such that:
(1) $\boldsymbol{\eta}$ is concentrated on the pairs $(x, \gamma)$ such that $\gamma$ is an absolutely continuous solution of $\dot{x}(t)=v_{t}(x(t))$ with initial condition $\gamma(a)=x$;
(2) for every $\varphi \in C_{b}^{0}\left(\mathbb{R}^{d}\right), t \in I$ we have

$$
\int_{\mathbb{R}^{d}} \varphi(x) d \mu_{t}(x)=\int_{\mathbb{R}^{d} \times \Gamma_{I}} \varphi(\gamma(t)) d \boldsymbol{\eta}(x, \gamma),
$$

(3) $\gamma(a) \sharp \boldsymbol{\eta}=\alpha, \gamma(b) \sharp \boldsymbol{\eta}=\beta$,
(4) $v_{t}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$ and a.e. $t \in I$ and $v_{t} \in L_{\mu_{t}}^{1}$ for a.e. $t \in I$.

In this case, we can define $\nu_{t}=v_{t} \mu_{t}$ thus we have simply $J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})=b-a$.
In the following, we will mainly focus our attention on admissible curves defined in $[0, T]$, for some suitable $T>0$. For later use we state the following technical lemma.

Lemma 4.4 (Estimates on admissible curves). Assume ( $F_{0}$ ) and ( $F_{1}$ ). Let $\boldsymbol{\mu}=$ $\left\{\mu_{t}\right\}_{t \in[0, T]}$ be an admissible curve driven by $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$. Then if $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ we have $\mu_{t} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ and $\mathrm{m}_{p-1}\left(\left|\nu_{t}\right|\right)<+\infty$ for all $t \in[0, T]$ and $p \geq 1$, more precisely
there are $A, B>0$ such that for all $t \in[0, T]$

$$
\begin{aligned}
\mathrm{m}_{p}\left(\mu_{t}\right) & \leq e^{A p t}\left(\mathrm{~m}_{p}\left(\mu_{0}\right)+\frac{B}{A}\right), \\
\mathrm{m}_{p-1}\left(\left|\nu_{t}\right|\right) & \leq e^{A p t}\left(A \mathrm{~m}_{p}\left(\mu_{0}\right)+\left(\mathrm{m}_{p-1}\left(\mu_{0}\right)+\frac{B}{A}\right) B e^{-A t}+B\right) .
\end{aligned}
$$

Proof. We have that $\frac{\nu_{t}}{\mu_{t}}(x) \in F(x)$ for $\mu_{t^{-}}$a.e. $x \in \mathbb{R}^{d}$, thus there exist $A, B>0$ such that $\left|\frac{\nu_{t}}{\mu_{t}}(x)\right| \leq r(|x|):=A|x|+B$. Take a family $\left\{\varphi_{n}^{\varepsilon}\right\}_{n \in \mathbb{N}} \subseteq C_{c}^{1}\left(\mathbb{R}^{d}\right)$ such that
a. $0 \leq \varphi_{n}^{\varepsilon}(x) \leq 1$ for all $x \in \mathbb{R}^{d}$;
b. $\left\{\varphi_{n}^{\varepsilon}(x)\right\}_{n \in \mathbb{N}}$ is a monotone increasing sequence converging to 1 for all $x \in \mathbb{R}^{d}$;
c. $\left|\nabla \varphi_{n}^{\varepsilon}(x)\right| \cdot|x|^{p} r(|x|) \leq p \varepsilon$ for all $x \in \mathbb{R}^{d}$ and $\varepsilon>0$.

We have

$$
\begin{aligned}
\frac{d}{d t} \mathrm{~m}_{p}\left(\varphi_{n}^{\varepsilon} \mu_{t}\right) & =\int_{\mathbb{R}^{d}}|x|^{p} \nabla \varphi_{n}^{\varepsilon}(x) d \nu_{t}(x)+p \int_{\mathbb{R}^{d}} \varphi_{n}^{\varepsilon}(x)|x|^{p-1} \frac{x}{|x|} d \nu_{t}(x) \\
& =\int_{\mathbb{R}^{d}}|x|^{p} \nabla \varphi_{n}^{\varepsilon}(x) \frac{\nu_{t}}{\mu_{t}}(x) d \mu_{t}(x)+p \int_{\mathbb{R}^{d}} \varphi_{n}^{\varepsilon}(x)|x|^{p-1} \frac{x}{|x|} \frac{\nu_{t}}{\mu_{t}}(x) d \mu_{t}(x) \\
& \leq \int_{\mathbb{R}^{d}}|x|^{p}\left|\nabla \varphi_{n}^{\varepsilon}(x)\right| r(|x|) d \mu_{t}(x)+p \int_{\mathbb{R}^{d}} r(|x|) \varphi_{n}^{\varepsilon}(x)|x|^{p-1} d \mu_{t}(x) \\
& \leq \varepsilon p+p \int_{\mathbb{R}^{d}}(A|x|+B) \varphi_{n}^{\varepsilon}(x)|x|^{p-1} d \mu_{t}(x) \\
& \leq A p \operatorname{m}_{p}\left(\varphi_{n}^{\varepsilon} \mu_{t}\right)+p(B+\varepsilon)
\end{aligned}
$$

This implies that

$$
\mathrm{m}_{p}\left(\varphi_{n}^{\varepsilon} \mu_{t}\right) \leq e^{A p t}\left(\mathrm{~m}_{p}\left(\varphi_{n}^{\varepsilon} \mu_{0}\right)+\frac{B+\varepsilon}{A}\right) \leq e^{A p t}\left(\mathrm{~m}_{p}\left(\mu_{0}\right)+\frac{B+\varepsilon}{A}\right)
$$

thus by Monotone Convergence Theorem by letting $n \rightarrow+\infty$ and $\varepsilon \rightarrow 0$

$$
\mathrm{m}_{p}\left(\mu_{t}\right) \leq e^{A p t}\left(\mathrm{~m}_{p}\left(\mu_{0}\right)+\frac{B}{A}\right)
$$

Moreover,

$$
\begin{aligned}
\mathrm{m}_{p-1}\left(\left|\nu_{t}\right|\right) & =\int_{\mathbb{R}^{d}}|x|^{p-1}\left|\frac{\nu_{t}}{\mu_{t}}(x)\right| d \mu_{t}(x) \\
& \leq \int_{\mathbb{R}^{d}}|x|^{p-1}(A|x|+B) d \mu_{t}(x)=A \mathrm{~m}_{p}\left(\mu_{t}\right)+B \mathrm{~m}_{p-1}\left(\mu_{t}\right) \\
& \leq e^{A p t}\left(A \mathrm{~m}_{p}\left(\mu_{0}\right)+\left(\mathrm{m}_{p-1}\left(\mu_{0}\right)+\frac{B}{A}\right) B e^{-A t}+B\right)
\end{aligned}
$$

which concludes the proof.
The following definitions are the natural counterpart of the classical case.
Definition 4.5 (Reachable set). Let $\mu_{0} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$, and $T>0$. Define the set of admissible curves defined on $[0, T]$ and starting from $\mu_{0}$ by setting $\mathscr{A}_{T}\left(\mu_{0}\right):=\left\{\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right): \boldsymbol{\mu}\right.$ is an admissible trajectory with $\left.\mu_{\mid t=0}=\mu_{0}\right\}$.

The reachable set from $\mu_{0}$ in time $T$ is

$$
\mathscr{R}_{T}\left(\mu_{0}\right):=\left\{\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right): \text { there exists } \boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]} \in \mathscr{A}_{T}\left(\mu_{0}\right) \text { with } \mu=\mu_{T}\right\} .
$$

Definition 4.6 (Generalized minimum time). Let $p \geq 1, \Phi \in C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ and $\tilde{S}^{\Phi}, \tilde{S}_{p}^{\Phi}$ be the corresponding generalized targets defined in Definition 3.1. In analogy with the classical case, we define the generalized minimum time function $\tilde{T}^{\Phi}: \mathscr{P}\left(\mathbb{R}^{d}\right) \rightarrow$ [ $0,+\infty$ ] by setting

$$
\begin{equation*}
\tilde{T}^{\Phi}\left(\mu_{0}\right):=\inf \left\{J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu}): \boldsymbol{\mu} \in \mathscr{A}_{T}\left(\mu_{0}\right), \boldsymbol{\mu} \text { is driven by } \boldsymbol{\nu}, \mu_{\mid t=T} \in \tilde{S}^{\Phi}\right\} \tag{4.2}
\end{equation*}
$$

where, by convention, $\inf \emptyset=+\infty$.
Given $\mu_{0} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ with $T^{\Phi}\left(\mu_{0}\right)<+\infty$, an admissible curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]} \subseteq$ $\mathscr{P}\left(\mathbb{R}^{d}\right)$, driven by a family of Borel vector-valued measures $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]}$ and satisfying $\mu_{\mid t=0}=\mu_{0}$ and $\mu_{\mid t=\tilde{T}^{\Phi}\left(\mu_{0}\right)} \in \tilde{S}^{\Phi}$ is optimal for $\mu_{0}$ if

$$
\tilde{T}^{\Phi}\left(\mu_{0}\right)=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu}) .
$$

Given $p \geq 1$, we define also a generalized minimum time function $\tilde{T}_{p}^{\Phi}: \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) \rightarrow$ $[0,+\infty]$ by replacing in the above definitions $\tilde{S}^{\Phi}$ by $\tilde{S}_{p}^{\Phi}$ and $\mathscr{P}\left(\mathbb{R}^{d}\right)$ by $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$. Since $\tilde{S}_{p}^{\Phi} \subseteq \tilde{S}^{\Phi}$, it is clear that $\tilde{T}^{\Phi}\left(\mu_{0}\right) \leq \tilde{T}_{p}^{\Phi}\left(\mu_{0}\right)$.
Remark 4.7. In view of the characterization in Theorem 8.3.1 in [3], and of Remark 4.2, one can think to $\tilde{T}^{\Phi}$ as the minimum time needed by the system to steer $\mu_{0}$ to a measure in $\tilde{S}^{\Phi}$, along absolutely continuous curves in $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$.

When the generalized target $\tilde{S}^{\Phi}$ admits a classical counterpart $S$, it is natural to ask for a comparison between the generalized minimum time function and the classical minimum time needed to reach $S$.
Proposition 4.8 (First comparison between $\tilde{T}^{\Phi}$ and $T$ ). Consider the generalized minimum time problem for $\Sigma_{F}$ as in Definition 4.6 assuming $\left(F_{0}\right),\left(F_{1}\right)$, and suppose that the corresponding generalized target $\tilde{S}^{\Phi}$ admits $S$ as classical counterpart. Then for all $\mu_{0} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ we have

$$
\tilde{T}^{\Phi}\left(\mu_{0}\right) \geq\|T\|_{L_{\mu_{0}}^{\infty}}
$$

where $T: \mathbb{R}^{d} \rightarrow[0,+\infty]$ is the classical minimum time function for system $\dot{x}(t) \in$ $F(x(t))$ with target $S$.
Proof. For sake of clarity, in this proof we will simply write $\tilde{T}$ and $\tilde{S}$, thus omitting $\Phi$, since we can always replace the set $\Phi$ by $\left\{d_{S}\right\}$ by assumption of existence of the classical counterpart $S$ of $\tilde{S}^{\Phi}$.

If $\tilde{T}\left(\mu_{0}\right)=+\infty$ there is nothing to prove, so assume $\tilde{T}\left(\mu_{0}\right)<+\infty$. Fix $\varepsilon>0$ and let $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ be an admissible curve starting from $\mu_{0}$, driven by a family of Borel vector-valued measures $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in I}$ such that $T=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})<$ $\tilde{T}\left(\mu_{0}\right)+\varepsilon$ and $\mu_{\mid t=T} \in \tilde{S}$. In particular, we have that $v_{t}(x):=\frac{\nu_{t}}{\mu_{t}}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$ and a.e. $t \in[0, T]$, hence $\left|v_{t}(x)\right| \leq\left(L_{1}+L_{2}\right)(1+|x|)$ for $\mu_{t}$-a.e $x \in \mathbb{R}^{d}$. Accordingly,

$$
\int_{0}^{T} \int_{\mathbb{R}^{d}} \frac{\left|v_{t}(x)\right|}{1+|x|} d \mu_{t} d t \leq T\left(L_{1}+L_{2}\right)<+\infty
$$

By the superposition principle 2.14, we have that there exists a probability measure $\boldsymbol{\eta} \in \mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{T}\right)$ satisfying
(1) $\boldsymbol{\eta}$ is concentrated on the pairs $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$ such that $\gamma$ is absolutely continuous and

$$
\gamma(t)=x+\int_{0}^{t} v_{t}(\gamma(s)) d s
$$

(2) for all $t \in[0, T]$ and all $\varphi \in C_{b}^{0}\left(\mathbb{R}^{d}\right)$

$$
\int_{\mathbb{R}^{d}} \varphi(x) d \mu_{t}(x)=\iint_{\mathbb{R}^{d} \times \Gamma_{T}} \varphi(\gamma(t)) d \boldsymbol{\eta}(x, \gamma) .
$$

Evaluating the above formula at $t=0$, we have that if $x \notin \operatorname{supp} \mu_{0}$ or $\gamma(0) \neq x$, then $(x, \gamma) \notin \operatorname{supp} \boldsymbol{\eta}$.

Let $\left\{\psi_{n}\right\}_{n \in \mathbb{N}} \in C_{c}^{\infty}\left(\mathbb{R}^{d} ;[0,1]\right)$ with $\psi_{n}(x)=0$ if $x \notin B(0, n+1)$ and $\psi_{n}(x)=1$ if $x \in \overline{B(0, n)}$. By Monotone Convergent Theorem, since $\left\{\psi_{n}(\cdot) d_{S}(\cdot)\right\}_{n \in \mathbb{N}} \subseteq C_{b}^{0}\left(\mathbb{R}^{d}\right)$ is an increasing sequence of nonnegative functions pointwise convergent to $d_{S}(\cdot)$, we have for every $t \in[0, T]$

$$
\begin{aligned}
\iint_{\mathbb{R}^{d} \times \Gamma_{T}} d_{S}(\gamma(t)) d \boldsymbol{\eta}(x, \gamma) & =\lim _{n \rightarrow \infty} \iint_{\mathbb{R}^{d} \times \Gamma_{T}} \psi_{n}(\gamma(t)) d_{S}(\gamma(t)) d \boldsymbol{\eta}(x, \gamma) \\
& =\lim _{n \rightarrow \infty} \int_{\mathbb{R}^{d}} \psi_{n}(x) d_{S}(x) d \mu_{t}(x)
\end{aligned}
$$

By taking $t=T$, we have that the last term vanishes because $\mu_{\mid t=T} \in \tilde{S}$ and so supp $\mu_{\mid t=T} \subseteq S$, therefore

$$
\iint_{\mathbb{R}^{d} \times \Gamma_{T}} d_{S}(\gamma(T)) d \boldsymbol{\eta}(x, \gamma)=0
$$

In particular, we necessarily have that $\gamma(T) \in S$ and $\gamma(0)=x$ for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in$ $\mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{T}\right)$, whence $T \geq T(x)$ for $\mu_{0}$-a.e. $x \in \mathbb{R}^{d}$, since $T(x)$ is the infimum of the times needed to steer $x$ to $S$ along trajectories of the system. Thus, $\tilde{T}\left(\mu_{0}\right)+\varepsilon \geq T(x)$ for $\mu_{0}$-a.e. $x \in \mathbb{R}^{d}$ and, by letting $\varepsilon \rightarrow 0$, we conclude that $\tilde{T}\left(\mu_{0}\right) \geq\|T\|_{L_{\mu_{0}}^{\infty}}$.

We notice that the inequality appearing in Proposition 4.8 may be strict without further assumptions.

Example 4.9. In $\mathbb{R}$, let $F(x)=\{1\}$ for all $x \in \mathbb{R}$ and set $\Phi=\{|\cdot|\}$, thus $S=$ $\{0\}$ is the classical counterpart of $\tilde{S}^{\Phi}=\left\{\delta_{0}\right\}$. Moreover, we have $T(x)=|x|$ for $x \leq 0$ and $T(x)=+\infty$ for $x>0$. Define $\mu_{0}=\frac{1}{2}\left(\delta_{-2}+\delta_{-1}\right)$. We have $\|T\|_{L_{\mu_{0}}^{\infty}}=\max \{T(-1), T(-2)\}=2$. However there are no solutions of $\dot{x}(t)=1$ steering any two different points to the origin in the same time, thus the set of admissible trajectories joining $\mu_{0}$ and $\delta_{0}$ is empty, hence $\tilde{T}^{\Phi}\left(\mu_{0}\right)=+\infty$.

Definition 4.10 (Convergence of curves in $\mathscr{P}\left(\mathbb{R}^{d}\right)$ ). We say that a family of curves $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in[0, T]}$ in $\mathscr{P}\left(\mathbb{R}^{d}\right)$
(1) pointwise converges to a curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ in $\mathscr{P}\left(\mathbb{R}^{d}\right)$ if and only if $\mu_{t}^{n} \rightharpoonup^{*}$ $\mu_{t}$ for all $t \in[0, T]$. In this case we will write $\boldsymbol{\mu}^{n} \Delta^{*} \boldsymbol{\mu}$.
(2) pointwise converges to a curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ in $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ if and only if $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in[0, T]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ and $\lim _{n \rightarrow+\infty} W_{p}\left(\mu_{t}^{n}, \mu_{t}\right)=0$ for all $t \in[0, T]$. In this case we will write $\boldsymbol{\mu}^{n} \rightarrow^{p} \boldsymbol{\mu}$.
(3) uniformly converges to a curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ in $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ if and only if $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in[0, T]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ and

$$
\lim _{n \rightarrow+\infty} \sup _{t \in[0, T]} W_{p}\left(\mu_{t}^{n}, \mu_{t}\right)=0 .
$$

In this case we will write $\boldsymbol{\mu}^{n} \rightrightarrows^{p} \boldsymbol{\mu}$.
Proposition 4.11 (Convergence of admissible trajectories). Assume ( $F_{0}$ ) and ( $F_{1}$ ). Let $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in[0, T]}$ be a sequence of admissible curves defined on $[0, T]$ such that $\boldsymbol{\mu}^{n}$ is driven by $\boldsymbol{\nu}^{n}=\left\{\nu_{t}^{n}\right\}_{t \in[0, T]}$ and suppose that there exist $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ and $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]} \subseteq \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that for a.e. $t \in[0, T]$ it holds $\left(\mu_{t}^{n}, \nu_{t}^{n}\right) \rightharpoonup^{*}$ $\left(\mu_{t}, \nu_{t}\right)$. Then $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$.

Proof. By assumption $\left(F_{0}\right)$, for every $\bar{x} \in \mathbb{R}^{d}$ and $\varepsilon>0$ there exists $0<\delta_{\bar{x}, \varepsilon} \leq \varepsilon$ such that if $x \in B\left(\bar{x}, \delta_{\bar{x}, \varepsilon}\right)$ we have $F(x) \subseteq F(\bar{x})+B(0, \varepsilon)$ and $F(\bar{x}) \subseteq F(x)+B(0, \varepsilon)$.

Define now by induction a countable covering $\left\{\Omega_{n}^{\varepsilon}\right\}_{n \in \mathbb{N}}$ of $\mathbb{R}^{d}$ in the following way. Let $\left\{y_{j}\right\}_{n \in \mathbb{N}}$ be an enumeration of $\mathbb{Q}^{d}$. Set $\Omega_{0}^{\varepsilon}=B\left(y_{0}, \delta_{y_{0}, \varepsilon}\right)$ and $x_{0}^{\varepsilon}=y_{0}$. Suppose to have defined $\Omega_{i}^{\varepsilon}$ and $x_{i}^{\varepsilon}$ for all $0 \leq i \leq n$. Set

$$
k_{n+1}^{\varepsilon}=\min \left\{j \in \mathbb{N}: B\left(y_{j}, \delta_{y_{j}, \varepsilon}\right) \nsubseteq \bigcup_{i=0}^{n} \Omega_{i}^{\varepsilon}\right\}
$$

and define $x_{n+1}^{\varepsilon}=y_{k_{n+1}^{\varepsilon}}$ and $\Omega_{n+1}^{\varepsilon}=B\left(x_{n+1}^{\varepsilon}, \delta_{x_{n+1}^{\varepsilon}, \varepsilon}\right) \backslash \bigcup_{i=0}^{n} \Omega_{i}^{\varepsilon}$.
We notice that for every $\varepsilon>0$ we have that $\left\{\Omega_{n}^{\varepsilon}\right\}_{n \in \mathbb{N}}$ is a countable covering made of Borel pairwise disjoint sets, moreover, recalling that $\Omega_{n}^{\varepsilon} \subseteq B\left(x_{n}^{\varepsilon}, \delta_{x_{n}^{\varepsilon}, \varepsilon}\right)$, we have

$$
\sup _{y \in \Omega_{n}^{\varepsilon}} d_{F\left(x_{n}^{\varepsilon}\right)}(F(y))<\varepsilon \text { for all } n \in \mathbb{N} \text {. }
$$

Given $\varepsilon>0$, define the following functional $G^{\varepsilon}: \mathscr{P}\left(\mathbb{R}^{d}\right) \times \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right) \rightarrow[0,+\infty]$

$$
G^{\varepsilon}(\mu, E):=\left\{\begin{array}{l}
\sum_{i=1}^{\infty} \int_{\Omega_{i}^{\varepsilon}} I_{F\left(x_{i}^{\varepsilon}\right)+\overline{B(0, \varepsilon)}}\left(\frac{E}{\mu}(x)\right) d \mu(x), \text { if }|E| \ll \mu, \\
+\infty, \text { otherwise },
\end{array}\right.
$$

where $\mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ is the set of vector-valued measures from $\mathbb{R}^{d}$ to $\mathbb{R}^{d}$, and $\frac{E}{\mu}$ denotes the derivation of the vector-valued measure $E$ with respect to $\mu$.

Notice that, since all the terms in the series can assume either the values 0 or $+\infty$, then we have

$$
G^{\varepsilon}(\mu, E)=\left\{\begin{array}{l}
\sup _{i \in \mathbb{N}} \int_{\Omega_{i}^{\varepsilon}} I_{F\left(x_{i}^{\varepsilon}\right)+\overline{B(0, \varepsilon)}}\left(\frac{E}{\mu}(x)\right) d \mu(x), \text { if }|E| \ll \mu, \\
+\infty, \text { otherwise. }
\end{array}\right.
$$

For every $\varepsilon>0$, the map $(\mu, E) \mapsto G^{\varepsilon}(\mu, E)$ is lower semicontinuous, since the function $I_{F(\bar{x})+\overline{B(0, \delta)}}(\cdot)$ is convex, lower semicontinuous and with superlinear growth, thus, according to Theorem 2.34 in [6], the functional is the supremum of l.s.c. maps, hence it is l.s.c.

We notice that

$$
G^{\varepsilon}(\mu, E)<+\infty \Longleftrightarrow G^{\varepsilon}(\mu, E)=0 \Longleftrightarrow d_{F\left(x_{i}^{\varepsilon}\right)}\left(\frac{E}{\mu}(x)\right) \leq \varepsilon \text { for } \mu \text {-a.e. } x \in \Omega_{i}^{\varepsilon} \text {. }
$$

Let now $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in[0, T]}, \boldsymbol{\nu}^{n}=\left\{\nu_{t}^{n}=v_{t}^{n} \mu_{t}^{n}\right\}_{t \in[0, T]}, \boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}, \boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$ as in the statement. Recall that $v_{t}^{n}(x):=\frac{\nu_{t}^{n}}{\mu_{t}^{n}}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$ and a.e. $t \in I$, and that by assumption $\nu_{t}^{n} \rightharpoonup^{*} \nu_{t}$ and $\mu_{t}^{n} \rightharpoonup^{*} \mu_{t}$ for a.e. $t \in[0, T]$.

Since $G^{\varepsilon}$ is l.s.c., for a.e. $t \in[0, T]$ we have that

$$
0 \leq G^{\varepsilon}\left(\mu_{t}, \nu_{t}\right) \leq \liminf _{n \rightarrow \infty} G^{\varepsilon}\left(\mu_{t}^{n}, \nu_{t}^{n}\right)=0,
$$

in particular $\nu_{t} \ll \mu_{t}$, thus we have $\nu_{t}=v_{t} \mu_{t}$ for a suitable $v_{t}(\cdot) \in L_{\mu_{t}}^{1}$, moreover $v_{t}(x) \in F\left(x_{i}^{\varepsilon}\right)+B(0, \varepsilon)$ for $\mu_{t}$-a.e. $x \in \Omega_{i}^{\varepsilon}$ and all $i \in \mathbb{N}$.

Let $\left\{\varepsilon_{i}\right\}_{i \in \mathbb{N}}$ be a sequence of positive numbers such that $\varepsilon_{i} \rightarrow 0^{+}$. Set

$$
N_{t}=\bigcup_{i \in \mathbb{N}} \bigcup_{k \in \mathbb{N}}\left\{x \in \Omega_{k}^{\varepsilon_{i}}: v_{t}(x) \notin F\left(x_{k}^{\varepsilon_{i}}\right)+B\left(0, \varepsilon_{i}\right)\right\}
$$

clearly $\mu_{t}\left(N_{t}\right)=0$ for a.e. $t \in[0, T]$.
We recall that for every $x \in \mathbb{R}^{d}$ and $\varepsilon>0$ there exists a unique $j(x, \varepsilon) \in \mathbb{N}$ such that $x \in \Omega_{j(x, \varepsilon)}^{\varepsilon}$, since the covering $\left\{\Omega_{n}^{\varepsilon}\right\}_{n \in \mathbb{N}}$ is a covering made by pairwise disjoint sets for every $\varepsilon>0$.

Fix $t \in[0, T]$ such that $\mu_{t}\left(N_{t}\right)=0$ and take $x \in \mathbb{R}^{d} \backslash N_{t}$. Thus we have $x \in$ $\Omega_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}} \subseteq B\left(x_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}}, \delta_{x_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}}, \varepsilon_{i}}\right)$, in particular $v_{t}(x) \in F\left(x_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}}\right)+B\left(0, \varepsilon_{i}\right)$, moreover since $\delta_{x_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}}, \varepsilon_{i}} \leq \varepsilon_{i}$, we have that $x_{j\left(x, \varepsilon_{i}\right)}^{\varepsilon_{i}} \rightarrow x$. Thus by letting $i \rightarrow+\infty$ we obtain that $v_{t}(x) \in F(x)$ for all $x \in \mathbb{R}^{d} \backslash N_{t}$, hence $v_{t}(x) \in F(x)$ for a.e. $t \in[0, T]$ and $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$.

Since for every $\varphi \in C_{c}^{1}\left(\mathbb{R}^{d}\right)$ we have in the sense of distributions on $[0, T]$,

$$
\frac{d}{d t} \int_{\mathbb{R}^{d}} \varphi(x) d \mu_{t}^{n}(x)=\int_{\mathbb{R}^{d}} \nabla \varphi(x) v_{t}^{n}(x) d \mu_{t}^{n}(x)
$$

and for the last term we have

$$
\lim _{n \rightarrow \infty} \int_{\mathbb{R}^{d}} \nabla \varphi(x) v_{t}^{n}(x) d \mu_{t}^{n}(x)=\int_{\mathbb{R}^{d}} \nabla \varphi(x) v_{t}(x) d \mu_{t}(x),
$$

due to the $w^{*}$-convergence of $\nu_{t}^{n}$ to $\nu_{t}$, thanks to Lemma 8.1.2 in [3], we deduce that, up to changing $\mu_{t}$ and $\nu_{t}$ for all $t$ belonging to a $\mathscr{L}^{1}$-negligible set of $[0, T]$, we have that $\boldsymbol{\mu}$ is an admissible curve driven by $\boldsymbol{\nu}$.

The previous Proposition is the key ingredient to prove the following theorem which, in analogy with the classical case, establish a sufficient condition to have relative compactness of a set of admissible trajectories.
Theorem 4.12. Assume $\left(F_{0}\right),\left(F_{1}\right)$. Let $\mathscr{A}$ be a set of admissible trajectories defined on $[0, T]$ and $C>0, p>1$ be constants such that for all $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]} \in \mathscr{A}$ it holds $\mathrm{m}_{p}\left(\mu_{t}\right) \leq C$ for a.e. $t \in[0, T]$. Then the pointwise $w^{*}$-closure of $\mathscr{A}$ is a set of admissible trajectories.

In particular, this holds if $\left\{\mathrm{m}_{p}\left(\mu_{0}\right)\right.$ : there exists $\boldsymbol{\mu} \in \mathscr{A}$ with $\left.\mu_{\mid t=0}=\mu_{0}\right\}$ is bounded, and, in particular, it holds for $\mathscr{A}_{T}\left(\mu_{0}\right)$ when $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$.

Proof. Let $\left\{\boldsymbol{\mu}^{n}\right\}_{n \in \mathbb{N}}$ be a sequence in $\mathscr{A}$. Since $\boldsymbol{\mu}^{n}$ is an admissible trajectory, it is driven by $\boldsymbol{\nu}^{n}=\left\{v_{t}^{n} \mu_{t}^{n}\right\}_{t \in[0, T]}$ with $v_{t}^{n} \in L_{\mu_{t}^{n}}^{1}$ and $v_{t}^{n}(x) \in F(x)$ for a.e. $t \in[0, T]$ and $\mu_{t}^{n}$-a.e. $x \in \mathbb{R}^{d}$. Since for a.e. $t \in[0, T]$

$$
\int_{\mathbb{R}^{d}}|x|^{p} d \mu_{t}^{n}(x) \leq C
$$

according to Remark 5.1.5 in [3], we have that for a.e. $t \in[0, T]$ there exists $\mu_{t} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$ such that $\mu_{t}^{n} \rightharpoonup^{*} \mu_{t}$. Similarly,

$$
\int_{\mathbb{R}^{d}}|x|^{p-1}\left|d \nu_{t}^{n}(x)\right|=\int_{\mathbb{R}^{d}}|x|^{p-1}\left|v_{t}^{n}(x)\right| d \mu_{t}^{n}(x) \leq L C+1,
$$

thus there exists $\nu_{t} \in \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that $\nu_{t}^{n} \rightharpoonup^{*} \nu_{t}$. By Proposition 4.11, we have that $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ is an admissible trajectory defined on $[0, T]$ driven by $\boldsymbol{\nu}$. The last assertion comes from Lemma 4.4, which allows to estimate the moments of $\mu_{t}$ and $\nu_{t}$ in terms of the moments of $\mu_{0}$.

Theorem 4.13 (L.s.c. of the generalized minimum time). Assume $\left(F_{0}\right)$ and $\left(F_{1}\right)$. Then $\tilde{T}_{p}^{\Phi}: \mathscr{P}_{p}\left(\mathbb{R}^{d}\right) \rightarrow[0,+\infty]$ is l.s.c. for all $p>1$.
Proof. Let $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, we have to prove that $\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right) \leq \liminf _{W_{p}\left(\mu, \mu_{0}\right) \rightarrow 0} \tilde{T}_{p}^{\Phi}(\mu)$. Taken a sequence $\left\{\mu_{0}^{n}\right\}_{n \in \mathbb{N}} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ s.t. $W_{p}\left(\mu_{0}^{n}, \mu_{0}\right) \rightarrow 0$ for $n \rightarrow+\infty$, and $\liminf _{W_{p}\left(\mu, \mu_{0}\right) \rightarrow 0} \tilde{T}_{p}^{\Phi}(\mu)=$ $\lim _{n \rightarrow+\infty} \tilde{T}_{p}^{\Phi}\left(\mu_{0}^{n}\right)=: T$, we want to prove that $\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right) \leq T$.

If $T=+\infty$ there is nothing to prove, so let us assume $T<+\infty$. Then there exists a sequence $\left\{T_{n}\right\}_{n \in \mathbb{N}}$ such that $T_{n} \rightarrow T$, and a sequence of admissible trajectories $\left\{\boldsymbol{\mu}^{n}\right\}_{n \in \mathbb{N}}$, with $\boldsymbol{\mu}^{n}=\left\{\mu_{t}^{n}\right\}_{t \in\left[0, T_{n}\right]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, such that $\mu_{\mid t=T_{n}}^{n} \in \tilde{S}_{p}^{\Phi}$ for all $n \in \mathbb{N}$.

Fix $\varepsilon>0$, then there exists $n_{\varepsilon} \in \mathbb{N}$ such that $T_{n_{\varepsilon}} \leq T+\varepsilon$ for all $n \geq n_{\varepsilon}$. For now on we will consider $n \geq n_{\varepsilon}$.

Our aim now is to extend the trajectories $\left\{\boldsymbol{\mu}^{n}\right\}_{n \in \mathbb{N}}$ to be all defined in the same time interval $[0, T+\varepsilon]$ so that we can use a result of compactness.

By $\left(F_{0}\right)$ we have that $F(\cdot)$ is continuous, thus by Theorem 8.1.3 in [8] there exists a Borel function $\bar{v}: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ such that $\bar{v}(x) \in F(x)$ for all $x \in \mathbb{R}^{d}$, and so we can consider the admissible trajectory $\overline{\boldsymbol{\mu}}^{n}:=\left\{\bar{\mu}_{t}^{n}\right\}_{t \in\left[T_{n}, T+\varepsilon\right]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ solution of

$$
\left\{\begin{array}{l}
\partial_{t} \mu_{t}+\operatorname{div}\left(\bar{v} \mu_{t}\right)=0, \quad \text { for } t>T_{n} \\
\mu_{\mid t=T_{n}}=\mu_{\mid t=T_{n}}^{n}
\end{array}\right.
$$

Using Lemma 2.13, we can define an admissible trajectory $\tilde{\boldsymbol{\mu}}^{n}=\left\{\tilde{\mu}_{t}^{n}\right\}_{t \in[0, T+\varepsilon]} \subseteq$ $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$, with

$$
\tilde{\mu}_{t}^{n}:= \begin{cases}\mu_{t}^{n}, & \text { for } t \in\left[0, T_{n}\right] \\ \bar{\mu}_{t}^{n}, & \text { for } t \in\left[T_{n}, T+\varepsilon\right]\end{cases}
$$

Observe also that, by definition, $W_{p}\left(\tilde{\mu}_{\mid t=0}^{n}, \mu_{0}\right) \rightarrow 0, n \rightarrow+\infty$.
Now we can apply Theorem 4.12 to say that there exists $\boldsymbol{\mu}:=\left\{\mu_{t}\right\}_{t \in[0, T+\varepsilon]} \subseteq$ $\mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ such that $\tilde{\boldsymbol{\mu}}^{n} \rightarrow^{p} \boldsymbol{\mu}, n \rightarrow+\infty$, with $\boldsymbol{\mu}$ an admissible trajectory and $\mu_{\mid t=0}=$ $\mu_{0}$.

To conclude we have to prove that there exists $t \in[0, T+\varepsilon]$ s.t. $\mu_{t} \in \tilde{S}_{p}^{\Phi}$, in fact in this way we obtain $\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right) \leq T+\varepsilon$, for all $\varepsilon>0$, thus by letting $\varepsilon \rightarrow 0^{+}$we have $\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right) \leq T$.
To prove this, suppose by contradiction that $\exists \delta>0$ s.t. $\tilde{d}_{\tilde{S}_{p}^{\text {D }}}\left(\mu_{t}\right)>\delta$ for all $t \in[0, T+\varepsilon]$ and observe that $\tilde{\mu}_{\mid t=T_{n}}^{n} \in \tilde{S}_{p}^{\Phi}, \forall n \in \mathbb{N}$. In particular, for $t=T_{n}$ we have

$$
\begin{aligned}
0<\delta & <\tilde{d}_{\tilde{S}_{P}^{\Phi}}\left(\mu_{\mid t=T_{n}}\right) \leq W_{p}\left(\mu_{\mid t=T_{n}}, \tilde{\mu}_{\mid t=T_{n}}^{n}\right) \\
& \leq W_{p}\left(\mu_{\mid t=T_{n}}, \mu_{\mid t=T}\right)+W_{p}\left(\mu_{\mid t=T}, \tilde{\mu}_{\mid t=T}^{n}\right)+W_{p}\left(\tilde{\mu}_{\mid t=T}^{n}, \tilde{\mu}_{\mid t=T_{n}}^{n}\right)
\end{aligned}
$$

Observing that

$$
\lim _{n \rightarrow+\infty} W_{p}\left(\mu_{\mid t=T_{n}}, \mu_{\mid t=T}\right)=\lim _{n \rightarrow \infty} W_{p}\left(\tilde{\mu}_{\mid t=T_{n}}^{n}, \tilde{\mu}_{\mid t=T}^{n}\right)=0
$$

by the continuity of the admissible curves w.r.t. $t$, and that

$$
\lim _{n \rightarrow+\infty} W_{p}\left(\mu_{\mid t=T}, \tilde{\mu}_{\mid t=T}^{n}\right)=0
$$

by $W_{p}$-convergence of $\tilde{\mu}_{t}^{n}$ towards $\mu_{t}$ for all $t \in[0, T+\varepsilon]$, we have a contradiction.
Theorem 4.14 (Existence of minimizers). Assume $\left(F_{0}\right),\left(F_{1}\right)$ and $p>1$. Let $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right), \Phi \in C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ satisfying assumptions of Definition 3.1, and let $\tilde{S}^{\Phi}$ be the corresponding generalized target. Let $\tilde{T}^{\Phi}\left(\mu_{0}\right)<\infty$. Then there exists an admissible curve $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ driven by $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$ which is optimal for $\mu_{0}$, that is $\tilde{T}^{\Phi}\left(\mu_{0}\right)=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})$. Moreover, we have also $\tilde{T}^{\Phi}\left(\mu_{0}\right)=\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right)$.

Proof. By the hypothesis of finiteness of $\tilde{T}^{\Phi}\left(\mu_{0}\right)$ and by definition of infimum we have that there exist $\left\{t_{n}\right\}_{n \in \mathbb{N}} \subset \mathbb{R}$ and a sequence of admissible trajectories $\boldsymbol{\mu}^{n}=$ $\left\{\mu_{t}^{n}\right\}_{t \in\left[0, t_{n}\right]}$, such that $\left.\mu^{n}\right|_{t=0}=\mu_{0},\left.\mu^{n}\right|_{t=t_{n}}=: \sigma^{n} \in \tilde{S}^{\Phi}, t_{n} \rightarrow \tilde{T}^{\Phi}\left(\mu_{0}\right)^{+}$. Moreover, by Lemma 4.4, we have that $\sigma^{n} \in \tilde{S}_{p}^{\Phi}$ for all $n \in \mathbb{N}$. We restrict all $\boldsymbol{\mu}^{n}$ to be defined on $\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]$.

By Theorem 4.12, $\boldsymbol{\mu}^{n} w^{*}$-converges up to subsequences to an admissible trajectory $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]}$ starting from $\mu_{0}$ driven by $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$, and by $w^{*}$-closure of $\tilde{S}^{\Phi}$ we have $\left.\sigma^{n} \rightharpoonup^{*} \mu\right|_{t=\tilde{T}^{\Phi}\left(\mu_{0}\right)} \in \tilde{S}^{\Phi}$. Applying again Lemma 4.4, we have that $\left.\mu\right|_{t=\tilde{T}^{\Phi}\left(\mu_{0}\right)} \in \tilde{S}_{p}^{\Phi}$. Thus $\tilde{T}^{\Phi}\left(\mu_{0}\right)=\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right)=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})$.

The following result allows us to justify the name of generalized minimum time given to functions $\tilde{T}^{\Phi}(\cdot)$ and $\tilde{T}_{p}^{\Phi}(\cdot)$.

Lemma 4.15 (Convexity property of the embedding of classical trajectories). Let $N \in \mathbb{N} \backslash\{0\}, T>0$ be given. Assume $\left(F_{0}\right)$ and $\left(F_{1}\right)$. Consider a family of continuous curves and real numbers $\left\{\left(\gamma_{i}, \lambda_{i}\right)\right\}_{i=1, \ldots, N} \subseteq \Gamma_{T} \times[0,1]$ such that $\gamma_{i}(\cdot)$ is a trajectory of $\dot{x}(t) \in F(x(t))$ for $i=1, \ldots, N$, and $\sum_{i=1}^{N} \lambda_{i}=1$.

For all $i=1, \ldots, N$ and $t \in[0, T]$, define the measures $\mu_{t}^{(i)}=\delta_{\gamma_{i}(t)}, \mu_{t}=$ $\sum_{i=1}^{N} \lambda_{i} \mu_{t}^{(i)}$,

$$
\nu_{t}^{(i)}= \begin{cases}\dot{\gamma}_{i}(t) \delta_{\gamma_{i}(t)}, & \text { if } \dot{\gamma}_{i}(t) \text { exists } \\ 0, & \text { otherwise }\end{cases}
$$

and $\nu_{t}=\sum_{i=1}^{N} \lambda_{i} \nu_{t}^{(i)}$. Then $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ is an admissible trajectory driven by $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$.
Proof. By linearity, clearly we have that

$$
\partial_{t} \mu_{t}+\operatorname{div} \nu_{t}=0
$$

is satisfied in the sense of distribution, moreover $\mu_{t}(B)=0$ implies $\nu_{t}(B)=0$ for every Borel set $B \subseteq \mathbb{R}^{d}$, thus $\left|\nu_{t}\right| \ll \mu_{t}$. It remains only to prove that for a.e. $t \in[0, T]$ we have $\nu_{t}=v_{t} \mu_{t}$ for a vector-valued function $v_{t} \in L^{1}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ satisfying $v_{t}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$. Set

$$
\tau=\left\{t \in[0, T]: \dot{\gamma}_{i}(t) \text { exists for all } i=1, \ldots, N \text { and } \dot{\gamma}_{i}(t) \in F\left(\gamma_{i}(t)\right)\right\}
$$

and notice that $\tau$ has full measure in $[0, T]$.
Fix $t \in \tau, x \in \operatorname{supp} \mu_{t}$. By definition of $\mu_{t}$, we have that there exists $I \subseteq$ $\{1, \ldots, N\}$ such that $\mu_{t}^{(i)}=\delta_{x}$ if and only if $i \in I$. So it is possible to find $\delta>0$ such that for all $0<\rho<\delta$ we have
$\mu_{t}(B(x, \rho))=\sum_{j \in I} \lambda_{j}, \quad \nu_{t}(B(x, \rho))=\sum_{i \in I} \lambda_{i} \int_{B(x, \rho)} \frac{\nu_{t}^{(i)}}{\mu_{t}^{(i)}}(y) d \mu_{t}^{(i)}(y)=\sum_{i \in I} \lambda_{i} \frac{\nu_{t}^{(i)}}{\mu_{t}^{(i)}}(x)$.
Thus for every $t \in \tau$ and $x \in \operatorname{supp} \mu_{t}$ we have

$$
v_{t}(x):=\lim _{\rho \rightarrow 0^{+}} \frac{\nu_{t}(B(x, \rho))}{\mu_{t}(B(x, \rho))}=\sum_{i \in I} \frac{\lambda_{i}}{\sum_{j \in I} \lambda_{j}} \frac{\nu_{t}^{(i)}}{\mu_{t}^{(i)}}(x),
$$

i.e., a convex combination of $\dot{\gamma}_{i}(t)=\frac{\nu_{t}^{(i)}}{\mu_{t}^{(i)}}(x) \in F(x)$ for $\mu_{t}$-a.e. $x \in \mathbb{R}^{d}$. Thus $\frac{\nu_{t}}{\mu_{t}}(x)=v_{t}(x) \in F(x)$, and so $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ is an admissible trajectory driven by $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, T]}$.
Corollary 4.16. Assume $\left(F_{0}\right)$ and $\left(F_{1}\right)$ and that the generalized target $\tilde{S}^{\Phi}$ admits a classical counterpart $S \subseteq \mathbb{R}^{d}$ which is weakly invariant for the dynamics $\dot{x}(t) \in$ $F(x(t))$. Let $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ with $p>1$. Then $\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right)=\tilde{T}^{\Phi}\left(\mu_{0}\right)=\|T(\cdot)\|_{L_{\mu_{0}}^{\infty}}$.

Proof. Since $\tilde{S}^{\Phi}$ admits classical counterpart $S$, we have that $S$ is closed and we can always take $\Phi=\left\{d_{S}(\cdot)\right\}$. Thus in this proof we will simply write $\tilde{T}_{p}$ and $\tilde{S}_{p}$ in place of $\tilde{T}_{p}^{\Phi}$ and $\tilde{S}_{p}^{\Phi}$, respectively.

By Proposition 4.8, we have only to prove that $\tilde{T}_{p}\left(\mu_{0}\right) \leq T:=\|T(\cdot)\|_{L_{\mu_{0}}}$. Assume that $T<+\infty$, otherwise there is nothing to prove. For $\mu_{0}$-a.e. point $x \in \mathbb{R}^{d}$ we have $T(x) \leq T$, thus there exists a trajectory $\gamma_{x}(\cdot)$ such that $\gamma_{x}(T(x)) \in S$. By the
weak invariance of $S$, we can extend this trajectory to be defined on $[0, T]$ with the constraint $\gamma_{x}(t) \in S$ for all $T(x) \leq t \leq T$, thus in particular $\gamma_{x}(T) \in S$. Fix $\varepsilon>0$, then there exists $N=N_{\varepsilon} \in \mathbb{N} \backslash\{0\}$, and $\left\{\left(x_{i}, \lambda_{i}\right): i=1, \ldots, N_{\varepsilon}\right\} \subseteq \operatorname{supp} \mu_{0} \times[0,1]$ such that:
(1) $\sum_{i=1}^{N_{\varepsilon}} \lambda_{i}=1$;
(2) $W_{p}\left(\mu_{0}, \sum_{i=1}^{N_{\varepsilon}} \lambda_{i} \delta_{x_{i}}\right)<\varepsilon$;
(3) there exist classical admissible trajectories $\left\{\gamma_{i}:[0, T] \rightarrow \mathbb{R}^{d}: i=1, \ldots, N_{\varepsilon}\right\}$ satisfying $\gamma_{i}(0)=x_{i}$ and $\gamma_{i}(T) \in S$ for all $i=1, \ldots, N_{\varepsilon}$.
It is possible to find an admissible trajectory $\boldsymbol{\mu}^{(\varepsilon)}=\left\{\mu_{t}^{(\varepsilon)}\right\}_{t \in[0, T]} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$ such that $\mu_{0}^{(\varepsilon)}=\sum_{i=1}^{N_{\varepsilon}} \lambda_{i} \delta_{x_{i}}$ and $\mu_{T}^{(\varepsilon)} \in \tilde{S}_{p}$, indeed, we can set
$\mu_{t}^{(\varepsilon)}=\sum_{i=1}^{N_{\varepsilon}} \lambda_{i} \delta_{\gamma_{i}(t),}, \quad \nu_{t}^{(\varepsilon)}= \begin{cases}\sum_{i=1}^{N_{\varepsilon}} \lambda_{i} \dot{\gamma}_{i}(t) \delta_{\gamma_{i}(t)}, & \text { if } \dot{\gamma}_{i}(t) \text { exists for all } i=1, \ldots, N_{\varepsilon}, \\ 0, & \text { otherwise },\end{cases}$
and then apply Lemma 4.15.
Since $\mu_{0}^{(\varepsilon)}$ converges in $W_{p}$ to $\mu_{0}$, we have that there exists $\bar{\varepsilon}>0$ such that the set $\left\{\mathrm{m}_{p}\left(\mu_{0}^{(\varepsilon)}\right): 0<\varepsilon<\bar{\varepsilon}\right\}$ is uniformly bounded by $\mathrm{m}_{p}\left(\mu_{0}\right)+1$. In particular, by taking a sequence $\varepsilon_{k} \rightarrow 0^{+}$, and the corresponding admissible trajectories $\boldsymbol{\mu}^{\left(\varepsilon_{k}\right)}$ driven by $\boldsymbol{\nu}^{\left(\varepsilon_{k}\right)}$, we can extract by Theorem 4.12 a subsequence converging to an admissible trajectory $\overline{\boldsymbol{\mu}}$ driven by $\overline{\boldsymbol{\nu}}$ satisfying $\bar{\mu}_{0}=\mu_{0}$. Since $\mu_{T}^{(\varepsilon)} \in \tilde{S}_{p}$ for all $\varepsilon>0$, by the closure of $\tilde{S}_{p}$ we have $\bar{\mu}_{T} \in \tilde{S}_{p}$, thus $\tilde{T}_{p}\left(\mu_{0}\right) \leq T$.
Corollary 4.17 (Second comparison result). Assume $\left(F_{0}\right)$ and $\left(F_{1}\right)$ and that the generalized target $\tilde{S}^{\Phi}$ admits a classical counterpart $S$. Then, for every $x_{0} \in \mathbb{R}^{d}$ we have $\tilde{T}^{\Phi}\left(\delta_{x_{0}}\right)=\tilde{T}_{p}^{\Phi}\left(\delta_{x_{0}}\right)=T\left(x_{0}\right)$ for all $p \geq 1$, where $T(\cdot)$ is the classical minimum time function for $\dot{x}(t) \in F(x(t))$ with target $S$.
Proof. Apply Lemma 4.15 to the family $\{(\gamma, 1)\}$, where $\gamma(\cdot)$ is an admissible trajectory of $\dot{x}(t) \in F(x(t))$ satisfying $\gamma(0)=x_{0}$ and $\gamma\left(T\left(x_{0}\right)\right) \in S$. We obtain an admissible trajectory steering $\delta_{x_{0}}$ to $\tilde{S}_{p}$ for all $p \geq 1$ in time $T\left(x_{0}\right)$, thus $\tilde{T}_{p}\left(\delta_{x_{0}}\right) \leq T\left(x_{0}\right)$. By Proposition 4.8, since $\|T(\cdot)\|_{L_{\delta_{x_{0}}}^{p}}=T\left(x_{0}\right)$, equality holds.

Theorem 4.18 (Dynamic programming principle). Let $0 \leq s \leq \tau$, let $F: \mathbb{R}^{d} \rightrightarrows \mathbb{R}^{d}$ be a set-valued function, let $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, \tau]}$ be an admissible curve for $\Sigma_{F}$. Then we have

$$
\tilde{T}^{\Phi}\left(\mu_{0}\right) \leq s+\tilde{T}^{\Phi}\left(\mu_{s}\right)
$$

Moreover, if $\tilde{T}^{\Phi}\left(\mu_{0}\right)<+\infty$, equality holds for all $s \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]$ if and only if $\boldsymbol{\mu}$ is optimal for $\mu_{0}=\mu_{\mid t=0}$. The same result holds for $\tilde{T}_{p}^{\Phi}$ in place of $\tilde{T}^{\Phi}, p \geq 1$.

Proof. Let $\boldsymbol{\nu}=\left\{\nu_{t}\right\}_{t \in[0, \tau]} \subseteq \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ be such that $\boldsymbol{\mu}$ is driven by $\boldsymbol{\nu}$. Fix $s \in[0, \tau]$, $\varepsilon>0$. If $\tilde{T}^{\Phi}\left(\mu_{s}\right)=+\infty$ there is nothing to prove. Otherwise there exists an
admissible curve $\boldsymbol{\mu}^{\varepsilon}:=\left\{\mu_{t}^{\varepsilon}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{s}\right)+\varepsilon\right]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ driven by $\boldsymbol{\nu}^{\varepsilon}=\left\{\nu_{t}^{\varepsilon}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{s}\right)+\varepsilon\right]} \subseteq$ $\mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that $\mu_{\mid t=0}^{\varepsilon}=\mu_{s}$ and $\mu_{\mid t=\tilde{T}^{\Phi}\left(\mu_{s}\right)+\varepsilon}^{\varepsilon} \in \tilde{S}^{\Phi}$. We consider

$$
\begin{gathered}
\tilde{v}_{t}^{\varepsilon}(x):= \begin{cases}\frac{\nu_{t}}{\mu_{t}}(x), & \text { for } 0 \leq t \leq s, \\
\frac{\nu_{t-s}^{\varepsilon}}{\mu_{t-s}^{\varepsilon}}(x), & \text { for } s<t \leq \tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon\end{cases} \\
\tilde{\mu}_{t}^{\varepsilon}:= \begin{cases}\mu_{t}, & \text { for } 0 \leq t \leq s, \\
\mu_{t-s}^{\varepsilon}, & \text { for } s<t \leq \tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon\end{cases}
\end{gathered}
$$

It is clear that $\tilde{\mu}_{\mid t=0}^{\varepsilon}=\mu_{0}$, that $\tilde{\mu}_{\mid t=\tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon} \in \tilde{S}^{\Phi}$, and that $\tilde{v}_{t}^{\varepsilon}(x) \in F(x)$ for $\tilde{\mu}_{t}^{\varepsilon}$-a.e. $x \in \mathbb{R}^{d}$ and a.e. $t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{s}\right)+\varepsilon\right]$. Moreover, $t \mapsto \tilde{\mu}_{t}^{\varepsilon}$ is narrowly continuous. Since Lemma 2.13 ensures that $\tilde{\boldsymbol{\mu}}^{\varepsilon}:=\left\{\tilde{\mu}_{t}^{\varepsilon}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon\right]}$ is a solution of the continuity equation driven by $\tilde{\boldsymbol{\nu}}^{\varepsilon}=\left\{\tilde{\nu}_{t}^{\varepsilon}=\tilde{v}_{t}^{\varepsilon} \tilde{\mu}_{t}^{\varepsilon}\right\}_{t \in \tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon}$, thus an admissible trajectory, we have that

$$
\tilde{T}^{\Phi}\left(\mu_{0}\right) \leq J_{F}\left(\tilde{\boldsymbol{\mu}}^{\varepsilon}, \tilde{\boldsymbol{\nu}}^{\varepsilon}\right)=\tilde{T}^{\Phi}\left(\mu_{s}\right)+s+\varepsilon .
$$

By arbitrariness of $\varepsilon>0$, we conclude that $\tilde{T}^{\Phi}\left(\mu_{0}\right) \leq s+\tilde{T}^{\Phi}\left(\mu_{s}\right)$.
Assume now that $\tilde{T}^{\Phi}\left(\mu_{0}\right)<+\infty$ and equality holds for all $s \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]$. Then, in particular, when $s=\tilde{T}^{\Phi}\left(\mu_{0}\right)$ we get

$$
\tilde{T}^{\Phi}\left(\mu_{0}\right)=\tilde{T}^{\Phi}\left(\mu_{0}\right)+\tilde{T}^{\Phi}\left(\mu_{\tilde{T}^{\Phi}\left(\mu_{0}\right)}\right) \quad \Rightarrow \quad \tilde{T}^{\Phi}\left(\mu_{\tilde{T}^{\Phi}\left(\mu_{0}\right)}\right)=0 .
$$

In turn, this implies $\mu_{\tilde{T}^{\Phi}\left(\mu_{0}\right)}=\mu_{s+\tilde{T}^{\Phi}\left(\mu_{s}\right)} \in \tilde{S}^{\Phi}$, and so $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]}$ joins $\mu_{0}$ with the generalized target. To prove optimality, it remains to be shown that $\tilde{T}^{\Phi}\left(\mu_{0}\right)=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})$. For all $s \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]$ we have

$$
\begin{aligned}
\tilde{T}^{\Phi}\left(\mu_{0}\right) & =s+\tilde{T}^{\Phi}\left(\mu_{s}\right)=s+J_{F}\left(\boldsymbol{\mu}_{\|\left[s, \tilde{T}^{\Phi}\left(\mu_{s}\right)+s\right]}, \boldsymbol{\nu}_{\left[\left[s, \tilde{T}^{\Phi}\left(\mu_{s}\right)+s\right]\right.}\right) \\
& =s+J_{F}\left(\boldsymbol{\mu}_{\|\left[s, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]}, \boldsymbol{\nu}_{\left[\left[s, \tilde{T}^{\Phi}\left(\mu_{0}\right)\right]\right]}\right) .
\end{aligned}
$$

In particular, for $s=0$ this yields exactly $\tilde{T}^{\Phi}\left(\mu_{0}\right)=J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})$, i.e., $\boldsymbol{\mu}$ is optimal for $\mu_{0}$.

Finally, assume that $\boldsymbol{\mu}$ is optimal for $\mu_{0}$ and $\tilde{T}^{\Phi}\left(\mu_{0}\right)<+\infty$. To have equality $\tilde{T}^{\Phi}\left(\mu_{0}\right)=s+\tilde{T}^{\Phi}\left(\mu_{s}\right)$, it is enough to show that $\tilde{T}^{\Phi}\left(\mu_{0}\right) \geq s+\tilde{T}^{\Phi}\left(\mu_{s}\right)$. If we define $\nu_{t}^{\prime}:=\nu_{t+s}$, we have that $\boldsymbol{\mu}^{\prime}=\left\{\mu_{t}^{\prime}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)-s\right]}:=\left\{\mu_{t+s}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)-s\right]}$ is a solution of the continuity equation driven by $\boldsymbol{\nu}^{\prime}=\left\{\nu_{t}^{\prime}\right\}_{t \in\left[0, \tilde{T}^{\Phi}\left(\mu_{0}\right)-s\right]}$. This implies that

$$
\begin{aligned}
\tilde{T}^{\Phi}\left(\mu_{0}\right) & =J_{F}(\boldsymbol{\mu}, \boldsymbol{\nu})=s+\int_{s}^{\tilde{T}^{\Phi}\left(\mu_{0}\right)} \int_{\mathbb{R}^{d}}\left(1+I_{F(x)}\left(\frac{\nu_{t}}{\mu_{t}}(x)\right)\right) d \mu_{t}(x) d t \\
& =s+\int_{0}^{\tilde{T}^{\Phi}\left(\mu_{0}\right)-s} \int_{\mathbb{R}^{d}}\left(1+I_{F(x)}\left(\frac{\nu_{t}^{\prime}}{\mu_{t}^{\prime}}(x)\right)\right) d \mu_{t}^{\prime}(x) d t \geq s+\tilde{T}^{\Phi}\left(\mu_{s}\right),
\end{aligned}
$$

which concludes the proof.
We are now interested in proving sufficient conditions on the set-valued function $F(\cdot)$ in order to have attainability of the generalized control system, i.e. to steer a
probability measure on the generalized target by following an admissible trajectory in finite time.

Representation formula for the generalized minimum time provided in Theorem 4.14 allows us to recover many results valid for the classical minimum time function also in the framework of generalized systems.

Theorem 4.19 (Attainability in the $C_{c}^{1}$ case). Assume $\left(F_{0}\right),\left(F_{1}\right)$. Let $\Phi \subseteq C_{c}^{1}\left(\mathbb{R}^{d} ; \mathbb{R}\right)$ satisfying assumptions of Definition 3.1 and let $\mu_{0} \in \mathscr{P}_{p}\left(\mathbb{R}^{d}\right), p \geq 1$. Assume that there exists a Borel vector field $v: \mathbb{R}^{d} \rightarrow \mathbb{R}^{d}$ and an admissible trajectory $\boldsymbol{\mu}:=\left\{\mu_{t}\right\}_{t \in[0,+\infty]} \subseteq \mathscr{P}\left(\mathbb{R}^{d}\right)$ driven by $\boldsymbol{\nu}=\left\{\nu_{t}:=v \mu_{t}\right\}_{t \in[0,+\infty[ }, \mu_{\mid t=0}=\mu_{0}$, such that the following condition holds:
$\left(C_{c}\right)$ for all $\phi \in \Phi$ there exists $k^{\phi}>0$ s.t. $\int_{\mathbb{R}^{d}}\langle\nabla \phi(x), v(x)\rangle d \mu_{t}(x) \leq-k^{\phi}$ for a.e. $t>0$.
Then we have

$$
\tilde{T}_{p}^{\Phi}\left(\mu_{0}\right) \leq \sup _{\phi \in \Phi}\left\{\frac{1}{k^{\phi}} \int_{\mathbb{R}^{d}} \phi(x) d \mu_{0}(x)\right\}
$$

Proof. We assume that the right hand side is finite, otherwise there is nothing to prove. We notice that, by the regularity hypothesis on $\Phi$, we can refer to Remark 3.2 for the definition of $\tilde{S}^{\Phi}$, moreover, by Lemma 4.4, we have $\boldsymbol{\mu} \subseteq \mathscr{P}_{p}\left(\mathbb{R}^{d}\right)$.

Given $\phi \in \Phi$, we set $L_{t}^{\phi}:=\int_{\mathbb{R}^{d}} \phi(x) d \mu_{t}(x)$, and from the continuity equation we have that in the distributional sense it holds

$$
\dot{L}_{t}^{\phi}=\frac{d}{d t} \int_{\mathbb{R}^{d}} \phi(x) d \mu_{t}(x)=\int_{\mathbb{R}^{d}}\langle\nabla \phi(x), v(x)\rangle d \mu_{t}(x) \leq-k^{\phi} .
$$

Then $L_{t}^{\phi}-L_{0}^{\phi} \leq-k^{\phi} t$ for $t>0$. Thus if $t \geq \sup _{\phi \in \Phi} \frac{L_{0}^{\phi}}{k^{\phi}}$, we have that $L_{t}^{\phi} \leq 0$ for all $\phi \in \Phi$, hence $\mu_{t} \in \tilde{S}^{\Phi}$ for $t \geq \sup _{\phi \in \Phi} \frac{L_{0}^{\phi}}{k^{\phi}}$, which ends the proof.

## 5. Hamilton-Jacobi-Bellman Equation

In this section we will prove that under suitable assumptions the generalized minimum time function solves a natural Hamilton-Jacobi-Bellman equation on $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ in the viscosity sense. The notion of viscosity sub-/superdifferential that we are going to use is different from other currently available in literature (e.g. [3], [13],[18],[17]), being modeled on this particular problem.

Throughout this section, given $T \in] 0,+\infty]$, we will use the evaluation map $e_{t}$ : $\mathbb{R}^{d} \times \Gamma_{T} \rightarrow \mathbb{R}^{d}$ defined as $e_{t}(x, \gamma)=\gamma(t)$ for all $0 \leq t<T$. Moreover, we set

$$
\begin{aligned}
\Gamma_{T}^{x}: & =\left\{\gamma \in \Gamma_{T}: \gamma(0)=x\right\}, \\
\mathscr{T}_{F}\left(\mu_{0}\right) & :=\left\{\boldsymbol{\eta} \in \mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{T}\right): T>0, \boldsymbol{\eta}\right. \text { concentrated on trajectories of } \\
& \left.\dot{\gamma}(t) \in F(\gamma(t)) \text { and satisfies } \gamma(0) \sharp \boldsymbol{\eta}=\mu_{0}\right\},
\end{aligned}
$$

where $x \in \mathbb{R}^{d}$ and $\mu_{0} \in \mathscr{P}\left(\mathbb{R}^{d}\right)$.
Lemma 5.1 (Properties of the evaluation operator). Assume $\left(F_{0}\right)$ and ( $F_{1}$ ), and let $L_{1}, L_{2}>0$ be the constants as in $\left(F_{1}\right)$. For any $\left.\left.\mu_{0} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right), T \in\right] 0,1\right], \boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$, we have:
(i) $\left|e_{t}(x, \gamma)\right| \leq\left(\left|e_{0}(x, \gamma)\right|+L_{2}\right) e^{L_{1}}$ for all $t \in[0, T]$ and $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$;
(ii) $e_{t} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T} ; \mathbb{R}^{d}\right)$ for all $t \in[0, T]$;
(iii) there exists $C>0$ depending only on $L_{1}, L_{2}$ such that for all $t \in[0, T]$ we have

$$
\left\|\frac{e_{t}-e_{0}}{t}\right\|_{L_{\eta}^{2}}^{2} \leq C\left(\mathrm{~m}_{2}\left(\mu_{0}\right)+1\right) .
$$

Proof. We set $\varphi_{t}(x, \gamma)=\frac{e_{t}(x, \gamma)-e_{0}(x, \gamma)}{t}$, notice that for all $t \geq 0$ the map $(x, \gamma) \mapsto \varphi_{t}(x, \gamma)$ does not depend on $x$-variable.

Item (i) follows from Lemma 2.17. To prove (ii) is enough to show $e_{0} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T}\right)$ and then apply item (i). Indeed, we have

$$
\begin{aligned}
& \iint_{\mathbb{R}^{d} \times \Gamma_{T}}\left|e_{0}(x, \gamma)\right|^{2} d \boldsymbol{\eta}(x, \gamma)=\int_{\mathbb{R}^{d}}|z|^{2} d(\gamma(0) \sharp \boldsymbol{\eta})(z)=\mathrm{m}_{2}\left(\mu_{0}\right)<+\infty, \\
& \iint_{\mathbb{R}^{d} \times \Gamma_{T}}\left|e_{t}(x, \gamma)\right|^{2} d \boldsymbol{\eta}(x, \gamma) \leq \\
& \quad \leq e^{2 L_{1} T}\left(\iint_{\mathbb{R}^{d} \times \Gamma_{T}}\left|e_{0}(x, \gamma)\right|^{2} d \boldsymbol{\eta}+L_{2}^{2} T^{2}+2 L_{2} T \iint_{\mathbb{R}^{d} \times \Gamma_{T}}\left|e_{0}(x, \gamma)\right| d \boldsymbol{\eta}\right) \\
& \quad \leq e^{2 L_{1} T}\left(\left(2 L_{2} T+1\right) \mathrm{m}_{2}\left(\mu_{0}\right)+L_{2}^{2} T^{2}\right),
\end{aligned}
$$

recalling that by Hölder inequality we have $\mathrm{m}_{1}\left(\mu_{0}\right) \leq \mathrm{m}_{2}\left(\mu_{0}\right)$.
We prove now (iii). For all $t \in] 0, T[$ we have

$$
\begin{aligned}
\left|\varphi_{t}(x, \gamma)\right| & =\frac{1}{t}|\gamma(t)-\gamma(0)|=\frac{1}{t} \int_{0}^{t}|\dot{\gamma}(s)| d s \leq \frac{L_{1}}{t} \int_{0}^{t}|\gamma(s)| d s+L_{2} \\
& \leq L_{1}\left(\left|e_{0}(x, \gamma)\right|+L_{2} T\right) e^{L_{1} T}+L_{2} \leq L_{1}\left(\left|e_{0}(x, \gamma)\right|+L_{2}\right) e^{L_{1}}+L_{2}
\end{aligned}
$$

By taking the square of this expression and integrating w.r.t. $\boldsymbol{\eta}$, we have that

$$
\left\|\varphi_{t}\right\|_{L_{\eta}^{2}}^{2} \leq C_{1}\left(\mathrm{~m}_{2}\left(\mu_{0}\right)+1\right)
$$

where we can take

$$
C_{1}=2 \max \left\{e^{2 L_{1}} L_{1}^{2}+2 e^{2 L_{1}} L_{2} L_{1}^{2}+2 e^{L_{1}} L_{2} L_{1}, e^{2 L_{1}} L_{2}^{2} L_{1}^{2}+2 e^{L_{1}} L_{2}^{2} L_{1}+L_{2}^{2}\right\} .
$$

Definition 5.2 (Averaged speed set). Assume $\left(F_{0}\right)$ and $\left(F_{1}\right), T>0$. For any $\mu_{0} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right), \boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$, we set

$$
\begin{aligned}
& \mathscr{V}(\boldsymbol{\eta}):=\left\{w_{\boldsymbol{\eta}} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T}\right): \exists\left\{t_{i}\right\}_{i \in \mathbb{N}} \subseteq\right] 0, T\left[, \text { with } t_{i} \rightarrow 0^{+}\right. \text {and } \\
&\left.\frac{e_{t_{i}}-e_{0}}{t_{i}} \rightharpoonup w_{\boldsymbol{\eta}} \text { weakly in } L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T} ; \mathbb{R}^{d}\right)\right\} .
\end{aligned}
$$

Observe that, indeed, by construction, if $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, then for every $\varepsilon>0$ we have $w_{\boldsymbol{\eta}} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{\varepsilon}\right)$, since all trajectories of the differential inclusion $\dot{x}(t) \in F(x(t))$ are defined on $[0,+\infty[$.

We notice that, according to the boundedness result of Lemma 5.1 (iii), for any sequence $\left.\left\{t_{i}\right\}_{i \in \mathbb{N}} \subseteq\right] 0, T\left[\right.$ with $t_{i} \rightarrow 0^{+}$, there exists a subsequence $\tau=\left\{t_{i_{k}}\right\}_{k \in \mathbb{N}}$
and $w_{\boldsymbol{\eta}} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T} ; \mathbb{R}^{d}\right)$ such that $\frac{e_{t_{i_{k}}}-e_{0}}{t_{i_{k}}}$ weakly converges to an element of $L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T} ; \mathbb{R}^{d}\right)$, thus $\mathscr{V}(\boldsymbol{\eta}) \neq \emptyset$.
Lemma 5.3 (Properties of the averaged speed set). Assume ( $F_{0}$ ) and ( $F_{1}$ ), $T>0$. For any $\mu_{0} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right), \boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$ and every $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ we have that
(i) $w_{\boldsymbol{\eta}}(x, \gamma) \in F(\gamma(0))$ for $\boldsymbol{\eta}$-a.e $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$.
(ii) if we denote by $\left\{\eta_{x}\right\}_{x \in \mathbb{R}^{d}}$ the disintegration of $\boldsymbol{\eta}$ w.r.t. the map $e_{0}$, the map

$$
x \mapsto \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}(\gamma),
$$

belongs to $L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$.
Proof. We prove (i). Fix $\varepsilon>0$ and $(x, \gamma) \in \operatorname{supp} \boldsymbol{\eta}$. Since $\gamma(\cdot)$ and $F(\cdot)$ are continuous, there exists $t_{\varepsilon, \gamma}^{*}>0$ such that for all $0<t<t_{\varepsilon, \gamma}^{*}$ we have $F(\gamma(t)) \subseteq$ $F(\gamma(0))+\varepsilon B(0,1)$. In particular, for all $0<t<t_{\varepsilon, \gamma}^{*}$ and $v \in \mathbb{R}^{d}$ we have

$$
\begin{aligned}
\left\langle v, \varphi_{t}(x, \gamma)\right\rangle & =\left\langle v, \frac{\gamma(t)-\gamma(0)}{t}\right\rangle=\frac{1}{t} \int_{0}^{t}\langle v, \dot{\gamma}(s)\rangle d s \\
& \leq \frac{1}{t} \int_{0}^{t} \sigma_{F(\gamma(s))}(v) d s \leq \sigma_{F(\gamma(0))+\varepsilon B(0,1)}(v)
\end{aligned}
$$

where $\varphi_{t}(x, \gamma)=\frac{e_{t}(x, \gamma)-e_{0}(x, \gamma)}{t}$.
Thus

$$
\overline{\operatorname{co}}\left\{\varphi_{t}(x, \gamma): 0<t<t_{\varepsilon, \gamma}^{*}\right\} \subseteq F(\gamma(0))+\varepsilon \overline{B(0,1)}
$$

Given $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, let $\left.\left.\left\{t_{i}\right\}_{i \in \mathbb{N}} \subseteq\right] 0,1\right]$ be a sequence such that $t_{i} \rightarrow 0^{+}$and $\varphi_{t_{i}} \rightharpoonup w_{\boldsymbol{\eta}}$ weakly in $L_{\eta}^{2}$. In particular, by Mazur's Lemma, there is a sequence in $\operatorname{co}\left\{\varphi_{t_{i}}: i \in\right.$ $\mathbb{N}\}$ strongly convergent to $w_{\boldsymbol{\eta}}$. In particular, for $(x, \gamma)$-a.e. point of $\mathbb{R}^{d} \times \Gamma_{T}$ we have pointwise convergence, i.e.

$$
w_{\boldsymbol{\eta}}(x, \gamma) \in \overline{\operatorname{co}}\left\{\varphi_{t_{i}}(x, \gamma): i \in \mathbb{N}\right\} .
$$

Given a point $(x, \gamma)$ where above pointwise convergence occurs, we can consider a subsequence $\left\{t_{i_{k}}\right\}_{k \in \mathbb{N}}$ of $t_{i}$ satisfying $0<t_{i_{k}}<t_{\varepsilon, \gamma}^{*}$, obtaining that

$$
\begin{aligned}
w_{\boldsymbol{\eta}}(x, \gamma) & \left.\in \overline{\operatorname{co}}\left\{\varphi_{t_{i_{k}}}(x, \gamma): k \in \mathbb{N}\right\}\right) \subseteq \overline{\operatorname{co}}\left\{\varphi_{t}(x, \gamma): 0<t<t_{\varepsilon, \gamma}^{*}\right\} \\
& \subseteq F(\gamma(0))+\varepsilon \overline{B(0,1)} .
\end{aligned}
$$

By letting $\varepsilon \rightarrow 0^{+}$we have that $w_{\boldsymbol{\eta}}(x, \gamma) \in F(\gamma(0))$ for $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$.
We prove now (ii). By definition, the disintegration of $\boldsymbol{\eta}$ w.r.t. the evaluation map $e_{0}$ is a family of measures $\left\{\eta_{x}\right\}_{x \in \mathbb{R}^{d}}$ satisfying (recall that $e_{0} \sharp \boldsymbol{\eta}=\mu_{0}$ )

$$
\iint_{\mathbb{R}^{d} \times \Gamma_{T}} f(x, \gamma) w_{\boldsymbol{\eta}}(x, \gamma) d \boldsymbol{\eta}(x, \gamma)=\int_{\mathbb{R}^{d}}\left(\int_{\Gamma_{T}^{x}}\left\langle f(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \eta_{x}(\gamma)\right) d \mu_{0}(x)
$$

for all Borel map $f: \mathbb{R}^{d} \times \Gamma_{T} \rightarrow \mathbb{R}^{d}$. Moreover the family $\left\{\eta_{x}\right\}_{x \in \mathbb{R}^{d}}$ is uniquely determined for $\mu_{0}$-a.e. $x \in \mathbb{R}^{d}$ (see e.g. Theorem 5.3.1 in [3]).

For any $\psi \in L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$, clearly we have $\psi \circ e_{0} \in L_{\boldsymbol{\eta}}^{2}\left(\mathbb{R}^{d} \times \Gamma_{T} ; \mathbb{R}^{d}\right)$, since $e_{0} \sharp \boldsymbol{\eta}=\mu_{0}$. Recalling that $w_{\boldsymbol{\eta}} \in L_{\eta}^{2}$, we obtain

$$
\begin{aligned}
\int_{\mathbb{R}^{d}}\left\langle\psi(x), \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}(\gamma)\right\rangle & d \mu_{0}(x)=\int_{\mathbb{R}^{d}} \int_{\Gamma_{T}^{x}}\left\langle\psi(x), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \eta_{x}(\gamma) d \mu_{0}(x) \\
& =\int_{\mathbb{R}^{d}} \int_{\Gamma_{T}^{x}}\left\langle\psi \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \eta_{x}(\gamma) d \mu_{0}(x) \\
& =\iint_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle\psi \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma)<+\infty .
\end{aligned}
$$

By the arbitrariness of $\psi \in L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$, we obtain that the map

$$
x \mapsto \int_{\Gamma_{T}^{x}} w_{\eta}(x, \gamma) d \eta_{x}(\gamma),
$$

belongs to $L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$, moreover for $\mu_{0}$-a.e. $x \in \mathbb{R}^{d}$, we have from (i) that

$$
\int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(\gamma) d \eta_{x}(\gamma) \in F(x) .
$$

Remark 5.4. We can interpret each $w_{\eta} \in \mathscr{V}(\boldsymbol{\eta})$ as a sort of averaged vector field of initial velocity in the sense of measure (we recall that in general an admissible trajectory $\gamma$ may fail to possess a tangent vector at $t=0$ ). The map

$$
x \mapsto \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(\gamma) d \eta_{x}(\gamma),
$$

can be interpreted as a initial barycentric speed of all the (weighted) trajectories emanating from $x$ in the support of $\boldsymbol{\eta}$. This approach is quite related to Theorem 5.4.4. in [3].

In the case in which the trajectory $t \mapsto e_{t} \sharp \boldsymbol{\eta}$ is driven by a sufficient smooth vector field, we recover exactly as averaged vector field and initial barycentric speed the expected objects, as shown below.

Lemma 5.5 (Regular driving vector fields). Let $\boldsymbol{\mu}=\left\{\mu_{t}\right\}_{t \in[0, T]}$ be an absolutely continuous solution of

$$
\left\{\begin{array}{l}
\left.\partial_{t} \mu_{t}+\operatorname{div}\left(v \mu_{t}\right)=0, t \in\right] 0, T[ \\
\mu_{\mid t=0}=\mu_{0} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right),
\end{array}\right.
$$

where $v \in C_{b}^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ satisfies $v(x) \in F(x)$ for all $x \in \mathbb{R}^{d}$. Then if $\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$ satisfies $\mu_{t}=e_{t} \sharp \boldsymbol{\eta}$ for all $t \in[0, T]$, we have that

$$
\lim _{t \rightarrow 0}\left\|\frac{e_{t}-e_{0}}{t}-v \circ e_{0}\right\|_{L_{\eta}^{2}}=0,
$$

and so $\mathscr{V}(\boldsymbol{\eta})=\left\{v \circ e_{0}\right\}$, thus we have

$$
\left\{x \mapsto \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}: w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})\right\}=\{v(\cdot)\} .
$$

Proof. We have

$$
\left\|\frac{e_{t}-e_{0}}{t}-v \circ e_{0}\right\|_{L_{\eta}^{2}}^{2}=\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left|\frac{\gamma(t)-\gamma(0)}{t}-v(\gamma(0))\right|^{2} d \boldsymbol{\eta}(x, \gamma),
$$

For $\boldsymbol{\eta}$-a.e. $(x, \gamma) \in \mathbb{R}^{d} \times \Gamma_{T}$, recalling the boundedness of $v$ and that we have $\gamma \in C^{1}$ and $\dot{\gamma}(t)=v(\gamma(t))$

$$
\begin{aligned}
\left|\frac{\gamma(t)-\gamma(0)}{t}-v(\gamma(0))\right| & \leq \frac{1}{t} \int_{0}^{t}|\dot{\gamma}(s)| d s+|v(\gamma(0))| \leq 2\|v\|_{\infty} \\
\lim _{t \rightarrow 0^{+}}\left|\frac{\gamma(t)-\gamma(0)}{t}-v(\gamma(0))\right| & =0
\end{aligned}
$$

Thus applying Lebesgue's Dominated Convergence Theorem we obtain

$$
\lim _{t \rightarrow 0}\left\|\frac{e_{t}-e_{0}}{t}-v \circ e_{0}\right\|_{L_{\eta}^{2}}^{2}=0
$$

hence $w_{\eta}=v \circ e_{0}$. The last assertion now follows.
We have already proved that the set

$$
\left\{x \mapsto \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}: \boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right), w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})\right\}
$$

is contained in the set of all $L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$-selections of $F(\cdot)$. The next density results shows that, indeed, equality holds: since allows to approximate every $L_{\mu_{0}}^{2}$-selections by $C_{b}^{0}$-selections, and then use Lemma 5.5.
Lemma 5.6 (Approximation). Let $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$. Assume $\left(F_{0}\right)$ and $\left(F_{3}\right)$. Then given any $v \in L_{\mu}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ satisfying $v(x) \in F(x)$ for $\mu$-a.e. $x \in \mathbb{R}^{d}$, there exists a sequence of continuous maps $\left\{g_{n}\right\}_{n \in \mathbb{N}} \subseteq C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that
(1) $\lim _{n \rightarrow \infty}\left\|g_{n}-v\right\|_{L_{\mu}^{2}}=0$;
(2) $g_{n}(x) \in F(x)$ for all $x \in \mathbb{R}^{d}$.

In particular, given $\mu_{0} \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$, we have

$$
\begin{aligned}
\left\{v \in L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)\right. & \left.: v(x) \in F(x) \text { for } \mu_{0} \text {-a.e. } x \in \mathbb{R}^{d}\right\}= \\
& =\left\{x \mapsto \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}: \boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right), w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})\right\} .
\end{aligned}
$$

Proof. By Lusin's Theorem (see e.g. Theorem 1.45 in [6]), we can construct a sequence of compact sets $\left\{K_{n}\right\}_{n \in \mathbb{N}} \subseteq \mathbb{R}^{d}$ and of continuous maps $\left\{v_{n}\right\}_{n \in \mathbb{N}} \subseteq C_{c}^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ such that $v_{n}=v$ on $K_{n}$ and $\mu\left(\mathbb{R}^{d} \backslash K_{n}\right) \leq 1 / n$. For all $n \in \mathbb{N}$ define the set valued maps

$$
G_{n}(x):= \begin{cases}F(x), & \text { for } x \in \mathbb{R}^{d} \backslash K_{n} \\ \left\{v_{n}(x)\right\}, & \text { for } x \in K_{n}\end{cases}
$$

We prove that $G_{n}(\cdot)$ is lower semicontinuous. If $x \in \mathbb{R}^{d} \backslash K_{n}$, then in a neighborhood of $x$ we have $G_{n}=F$, thus $G_{n}$ is lower semicontinuous. Let $x \in K_{n}$ and $V$ be an open set such that $V \cap G_{n}(x) \neq \emptyset$. In particular, we have that $V$ is an open neighborhood of $v_{n}(x)$. Without loss of generality, we may assume that $V=B\left(v_{n}(x), \varepsilon\right)$ for
$\varepsilon>0$, thus there exists $\delta>0$ such that if $y \in B(x, \delta) \cap K_{n}$ we have $v_{n}(y) \in V$, and so $G_{n}(y) \cap V \neq \emptyset$. On the other hand, by continuity of $F$, there exists an open neighborhood $U$ of $x$ such that $V \cap F(y) \neq \emptyset$ for all $y \in U$. Thus, if we set $U^{\prime}=U \cap B(x, \delta) \backslash K_{n}$, we have that $U^{\prime}$ is an open neighborhood of $x$ satisfying:
(1) for all $y \in U^{\prime} \backslash K_{n}$ we have $F(y)=G_{n}(y)$ and so $G_{n}(y) \cap V \neq \emptyset$;
(2) for all $y \in U^{\prime} \cap K_{n}$ we have $v_{n}(y) \in V$, and so $G_{n}(y) \cap V \neq \emptyset$;
and so given $V$ for all $y \in U^{\prime}$ we have $G_{n}(y) \cap V \neq \emptyset$, which proves lower semicontinuity. Since $G_{n}(\cdot)$ is lower semicontinuous with compact convex values, by Michael's Selection Theorem (see e.g. Theorem 9.1.2 in [8]) we can find a continuous selection $g_{n} \in C^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ which by construction agrees with $v$ and $v_{n}$ on $K_{n}$ and satisfies $g_{n}(x) \in G_{n}(x) \subseteq F(x)$ for all $x \in \mathbb{R}^{d}$. Finally, we have
$\int_{\mathbb{R}^{d}}\left|v(x)-g_{n}(x)\right|^{2} d \mu(x)=\int_{\mathbb{R}^{d} \backslash K_{n}}\left|v(x)-g_{n}(x)\right|^{2} d \mu(x) \leq 4 M^{2} \mu\left(\mathbb{R}^{d} \backslash K_{n}\right) \leq \frac{4 M^{2}}{n}$, where $M$ is as in assumption $\left(F_{3}\right)$. The last assertion comes from Lemma 5.5.

We introduce now the following definition of viscosity sub-/superdifferential. For other concepts of viscosity sub-/superdifferential, we refer the reader to [3] and [13].
Definition 5.7 (Sub-/Super-differential in $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ ). Let $V: \mathscr{P}_{2}\left(\mathbb{R}^{d}\right) \rightarrow \mathbb{R}$ be a function. Fix $\mu \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ and $\delta>0$. We say that $p_{\mu} \in L_{\mu}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ belongs to the $\delta$-superdifferential $D_{\delta}^{+} V(\mu)$ at $\mu$ if for all $T>0$ and $\boldsymbol{\eta} \in \mathscr{P}\left(\mathbb{R}^{d} \times \Gamma_{T}\right)$ such that $t \mapsto e_{t} \sharp \boldsymbol{\eta}$ is an absolutely continuous curve in $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$ defined in $[0, T]$ with $e_{0} \sharp \boldsymbol{\eta}=\mu$ we have

$$
\begin{equation*}
\limsup _{t \rightarrow 0^{+}} \frac{V\left(e_{t} \sharp \boldsymbol{\eta}\right)-V\left(e_{0} \sharp \boldsymbol{\eta}\right)-\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu} \circ e_{0}(x, \gamma), e_{t}(x, \gamma)-e_{0}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma)}{\left\|e_{t}-e_{0}\right\|_{L_{\eta}^{2}}} \leq \delta . \tag{5.1}
\end{equation*}
$$

In the same way, $q_{\mu} \in L_{\mu}^{2}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$ belongs to the $\delta$-subdifferential $D_{\delta}^{-} V(\mu)$ at $\mu$ if $-q_{\mu} \in D_{\delta}^{+}[-V](\mu)$. Moreover, $D_{\delta}^{ \pm}[V](\mu)$ is the closure in $L_{\mu}^{2}$ of $D_{\delta}^{ \pm}[V](\mu) \cap$ $C_{b}^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right)$.
Definition 5.8 (Viscosity solutions). Let $V: \mathscr{P}_{2}\left(\mathbb{R}^{d}\right) \rightarrow \mathbb{R}$ be a function and $\mathscr{H}: \mathscr{P}_{2}\left(\mathbb{R}^{d}\right) \times C_{b}^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right) \rightarrow \mathbb{R}$. We say that $V$ is a
(1) viscosity supersolution of $\mathscr{H}(\mu, D V(\mu))=0$ if there exists $C>0$ depending only on $\mathscr{H}$ such that $\mathscr{H}\left(\mu, q_{\mu}\right) \geq-C \delta$ for all $q_{\mu} \in D_{\delta}^{-} V(\mu) \cap C_{b}^{0}, \mu \in$ $\mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$.
(2) viscosity subsolution of $\mathscr{H}(\mu, D V(\mu))=0$ if there exists $C>0$ depending only on $\mathscr{H}$ such that $\mathscr{H}\left(\mu, p_{\mu}\right) \leq C \delta$ for all $p_{\mu} \in D_{\delta}^{+} V(\mu) \cap C_{b}^{0}, \mu \in \mathscr{P}_{2}\left(\mathbb{R}^{d}\right)$.
(3) viscosity solution of $\mathscr{H}(\mu, D V(\mu))=0$ if it is both a viscosity subsolution and a viscosity supersolution.
Definition 5.9 (Hamiltonian Function). Given $\mu \in \mathscr{P}\left(\mathbb{R}^{d}\right)$, define
$\mathscr{D}(\mu):=\left\{\nu \in \mathscr{M}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right):|\nu| \ll \mu\right.$ and $\left.\int_{\mathbb{R}^{d}}\left(\left|\frac{\nu}{\mu}\right|^{2}+I_{F(x)}\left(\frac{\nu}{\mu}(x)\right)\right) d \mu<+\infty\right\}$.
We define the map $\mathscr{H}_{F}: \mathscr{P}_{2}\left(\mathbb{R}^{d}\right) \times C_{b}^{0}\left(\mathbb{R}^{d} ; \mathbb{R}^{d}\right) \rightarrow \mathbb{R}$ by setting

$$
\mathscr{H}_{F}(\mu, \psi):=-\left[1+\inf _{\nu \in \mathscr{D}(\mu)} \int_{\mathbb{R}^{d}}\left\langle\psi(x), \frac{\nu}{\mu}(x)\right\rangle d \mu\right] .
$$

Theorem 5.10 (Viscosity solution). Assume $\left(F_{0}\right)$ and $\left(F_{3}\right)$. Then $\tilde{T}_{2}(\cdot)$ it is a viscosity solution of $\mathscr{H}_{F}\left(\mu, D \tilde{T}_{2}(\mu)\right)=0$, with $\mathscr{H}_{F}$ defined as in Definition 5.9.

Proof. The proof is splitted in two claims.
Claim 1: $\tilde{T}_{2}(\cdot)$ is a subsolution of $\mathscr{H}_{F}\left(\mu, D \tilde{T}_{2}(\mu)\right)=0$.
Proof of Claim 1. Given $\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$ and set $\mu_{t}=e_{t} \sharp \boldsymbol{\eta}$ for all $t$ by the Dynamic Programming Principle we have $\tilde{T}_{2}\left(\mu_{0}\right) \leq \tilde{T}_{2}\left(\mu_{s}\right)+s$ for all $0<s \leq \tilde{T}_{2}\left(\mu_{0}\right)$. Without loss of generality, we can assume $0<s<1$. Given any $p_{\mu_{0}} \in D_{\delta}^{+} \tilde{T}_{2}\left(\mu_{0}\right)$, and set

$$
\begin{aligned}
& A\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right):=-s-\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), e_{s}(x, \gamma)-e_{0}(x, \gamma)\right\rangle d \boldsymbol{\eta}, \\
& B\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right):=\tilde{T}_{2}\left(\mu_{s}\right)-\tilde{T}_{2}\left(\mu_{0}\right)-\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), e_{s}(x, \gamma)-e_{0}(x, \gamma)\right\rangle d \boldsymbol{\eta},
\end{aligned}
$$

we have $A\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right) \leq B\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right)$.
We recall that since by definition $p_{\mu_{0}} \in L_{\mu_{0}}^{2}$, we have that $p_{\mu_{0}} \circ e_{0} \in L_{\eta}^{2}$. Dividing by $s>0$ the left hand side, we obtain that there exists $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, for which we have

$$
\limsup _{s \rightarrow 0^{+}} \frac{A\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right)}{s}=-1-\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma) .
$$

Recalling the choice of $p_{\mu_{0}}$, we have

$$
\limsup _{s \rightarrow 0^{+}} \frac{B\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right)}{s}=\limsup _{s \rightarrow 0^{+}} \frac{B\left(s, p_{\mu_{0}}, \boldsymbol{\eta}\right)}{\left\|e_{s}-e_{0}\right\|_{L_{\eta}^{2}}} \cdot\left\|\frac{e_{s}-e_{0}}{s}\right\|_{L_{\eta}^{2}} \leq C \delta
$$

where $C>0$ is a suitable constant (we can take twice the upper bound on $F$ given by $\left(F_{3}\right)$ ).

We thus obtain for all $\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$ that

$$
1+\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma) \geq-C \delta .
$$

By passing to the infimum on $\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$, and recalling Lemma 5.6, we have

$$
\begin{aligned}
-C \delta & \leq 1+\inf _{\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)} \int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma) \\
& =1+\inf _{\eta \in \mathscr{T}_{F}\left(\mu_{0}\right)} \int_{\mathbb{R}^{d}} \int_{\Gamma_{T}^{x}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \eta_{x} d \mu_{0} \\
& =1+\inf _{\eta \in \mathscr{T}_{F}\left(\mu_{0}\right)} \int_{\mathbb{R}^{d}}\left\langle p_{\mu_{0}} \circ e_{0}(x, \gamma), \int_{\Gamma_{T}^{x}} w_{\boldsymbol{\eta}}(x, \gamma) d \eta_{x}\right\rangle d \mu_{0} \\
& =1+\inf _{\substack{v \in L_{\mu_{0}}^{2}\left(\mathbb{R}^{d} \mathbb{R}^{d}\right) \\
v(x) \in F(x) \mu_{0}-\mathrm{a} . \mathrm{e}}} \int_{\mathbb{R}^{d}}\left\langle p_{\mu_{0}}, v\right\rangle d \mu_{0}=-\mathscr{H}_{F}\left(\mu_{0}, p_{\mu_{0}}\right),
\end{aligned}
$$

so $\tilde{T}_{2}(\cdot)$ is a subsolution, thus confirming Claim 1.
Claim 2: $\tilde{T}_{2}(\cdot)$ is a supersolution of $\mathscr{H}_{F}\left(\mu, D \tilde{T}_{2}(\mu)\right)=0$.
Proof of Claim 2. Given $\boldsymbol{\eta} \in \mathscr{T}_{F}\left(\mu_{0}\right)$ and defined the admissible trajectory $\boldsymbol{\mu}=$ $\left\{\mu_{t}\right\}_{t \in[0, T]}=\left\{e_{t} \sharp \boldsymbol{\eta}\right\}_{t \in[0, T]}$, and $q_{\mu_{0}} \in D_{\delta}^{-} \tilde{T}_{2}\left(\mu_{0}\right)$, there is a sequence $\left.\left\{s_{i}\right\}_{i \in \mathbb{N}} \subseteq\right] 0, T[$
and $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ such that $s_{i} \rightarrow 0^{+}, \frac{e_{s_{i}}-e_{0}}{s_{i}}$ weakly converges to $w_{\boldsymbol{\eta}}$ in $L_{\boldsymbol{\eta}}^{2}$, and for all $i \in \mathbb{N}$

$$
\begin{aligned}
\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle q_{\mu_{0}} \circ e_{0}(x, \gamma),\right. & \left.\frac{e_{s_{i}}(x, \gamma)-e_{0}(x, \gamma)}{s_{i}}\right\rangle d \boldsymbol{\eta}(x, \gamma) \\
& \leq 2 \delta\left\|\frac{e_{s_{i}}-e_{0}}{s_{i}}\right\|_{L_{\eta}^{2}}-\frac{\tilde{T}_{2}\left(\mu_{0}\right)-\tilde{T}_{2}\left(\mu_{s_{i}}\right)}{s_{i}} .
\end{aligned}
$$

By taking $i$ sufficiently large we thus obtain

$$
\int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle q_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma) \leq 3 C \delta-\frac{\tilde{T}_{2}\left(\mu_{0}\right)-\tilde{T}_{2}\left(\mu_{s_{i}}\right)}{s_{i}} .
$$

By using Lemma 5.6 and arguing as in Claim 1, we have

$$
\inf _{\eta \in \mathscr{\mathscr { T }}_{F}\left(\mu_{0}\right)} \int_{\mathbb{R}^{d} \times \Gamma_{T}}\left\langle q_{\mu_{0}} \circ e_{0}(x, \gamma), w_{\boldsymbol{\eta}}(x, \gamma)\right\rangle d \boldsymbol{\eta}(x, \gamma)=-\mathscr{H}_{F}\left(\mu_{0}, q_{\mu_{0}}\right)-1,
$$

and so

$$
\mathscr{H}_{F}\left(\mu_{0}, q_{\mu_{0}}\right) \geq-3 C \delta+\frac{\tilde{T}_{2}\left(\mu_{0}\right)-\tilde{T}_{2}\left(\mu_{s_{i}}\right)}{s_{i}}-1
$$

By the Dynamic Programming Principle, passing to the infimum on all admissible curves and recalling that $\frac{\tilde{T}_{2}\left(\mu_{0}\right)-\tilde{T}\left(\mu_{s}\right)}{s}-1 \leq 0$ with equality holding if and only if $\boldsymbol{\mu}$ is optimal, we obtain $\mathscr{H}_{F}\left(\mu_{0}, q_{\mu_{0}}\right) \geq-C^{\prime} \delta$, which proves that $\tilde{T}_{2}(\cdot)$ is a supersolution, thus confirming Claim 2.

## References

[1] Luigi Ambrosio, The flow associated to weakly differentiable vector fields: recent results and open problems, Nonlinear conservation laws and applications, IMA Vol. Math. Appl., vol. 153, Springer, New York, 2011, pp. 181-193, DOI 10.1007/978-1-4419-9554-4_7. MR2856995 (2012i:35224)
[2] Luigi Ambrosio and Wilfred Gangbo, Hamiltonian ODEs in the Wasserstein space of probability measures, Comm. Pure Appl. Math. 61 (2008), no. 1, 18-53, DOI 10.1002/cpa.20188. MR2361303 (2009b:37101)
[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, Gradient flows in metric spaces and in the space of probability measures, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR2401600 (2009h:49002)
[4] Fabio Ancona and Alberto Bressan, Patchy vector fields and asymptotic stabilization, ESAIM Control Optim. Calc. Var. 4 (1999), 445-471.
[5] , Nearly time optimal stabilizing patchy feedbacks, Ann. Inst. Henri Poincaré, Analyse Non Linéaire 24 (2007), 279-310.
[6] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000. MR1857292 (2003a:49002)
[7] Jean-Pierre Aubin and Arrigo Cellina, Differential Inclusions: Set-Valued Maps and Viability Theory, Springer-Verlag Berlin Heidelberg, Germany, 1984.
[8] Jean-Pierre Aubin and Hélène Frankowska, Set-valued analysis, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2009. Reprint of the 1990 edition [MR1048347]. MR2458436
[9] Patrick Bernard, Young measures, superpositions and transport, Indiana Univ. Math. J. 57 (2008), no. 1, 247-276.
[10] Alberto Bressan and Fabio S. Priuli, Nearly optimal patchy feedbacks, Discr. Cont. Dyn. Systems - Series A 21 (2008), 687-701.
[11] Piermarco Cannarsa and Peter R. Wolenski, Semiconcavity of the value function for a class of differential inclusions, Discrete Contin. Dyn. Syst. 29 (2011), no. 2, 453-466, DOI 10.3934/dcds.2011.29.453. MR2728465 (2012a:49046)
[12] Piermarco Cannarsa, Antonio Marigonda, and Khai T. Nguyen, Optimality conditions and regularity results for time optimal control problems with differential inclusions, J. Math. Anal. Appl. 427 (2015), no. 1, 202-228, DOI 10.1016/j.jmaa.2015.02.027. MR3318195
[13] Pierre Cardaliaguet and Marc Quincampoix, Deterministic differential games under probability knowledge of initial condition, Int. Game Theory Rev. 10 (2008), no. 1, 1-16.
[14] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, Nonsmooth analysis and control theory, Graduate Texts in Mathematics, vol. 178, Springer-Verlag, New York, 1998. MR1488695 (99a:49001)
[15] Piermarco Cannarsa and Carlo Sinestrari, Semiconcave functions, Hamilton-Jacobi equations, and optimal control, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston Inc., Boston, MA, 2004. MR2041617 (2005e:49001)
[16] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré, A new class of transport distances between measures, Calc. Var. Partial Differential Equations 34 (2009), no. 2, 193-231, DOI 10.1007/s00526-008-0182-5. MR2448650 (2009g:49019)
[17] Wilfrid Gangbo, Truyen Nguyen, and Adrian Tudorascu, Hamilton-Jacobi equations in the Wasserstein space, Methods Appl. Anal. 15 (2008), no. 2, 155-183, DOI 10.4310/MAA.2008.v15.n2.a4. MR2481677 (2010f:49061)
[18] Wilfrid Gangbo and Andrzej Święch, Optimal transport and large number of particles, Discrete Contin. Dyn. Syst. 34 (2014), no. 4, 1397-1441, DOI 10.3934/dcds.2014.34.1397. MR3117847
[19] Pierre Louis Lions and Jean-Michel Lasry, Jeux à champ moyen. I. Le cas stationnaire, C. R. Acad. Sci. Paris 343 (2006), 619-625.
[20]_, Jeux à champ moyen. II. Horizon fini et contrôle optimal, C. R. Acad. Sci. Paris 343 (2006), 679-684.
[21] Franco Rampazzo and Héctor J. Sussmann, Commutators of flow maps of nonsmooth vector fields, J. Differential Equations 232 (2007), no. 1, 134-175, DOI 10.1016/j.jde.2006.04.016. MR2281192 (2007j:49021)
[22] Héctor J. Sussmann, Subanalytic sets and feedback control, J. Differential Equations 31 (1979), 31-52.
[23] Cédric Villani, Topics in optimal transportation, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR1964483 (2004e:90003)

\author{
Giulia Cavagnari: Dipartimento di Matematica, Università di Trento Via Sommarive 14, I-38123 Povo (Trento), Italy. <br> E-mail address: giulia.cavagnari@unitn.it <br> Antonio Marigonda: Department of Computer Science, University of Verona <br> Strada Le Grazie 15, I-37134 Verona, Italy. <br> E-mail address: antonio.marigonda@univr.it <br> 

Fabio S. Priuli: Istituto per le Applicazioni del Calcolo "M.Picone"
Consiglio Nazionale delle Ricerche
Via dei Taurini 19, 00185 - Roma, Italy.
E-mail address: f.priuli@iac.cnr.it


[^0]:    Date: July 20, 2015.
    2010 Mathematics Subject Classification. 34A60, 49J15.
    Key words and phrases. optimal transport, differential inclusions, time optimal control.
    Acknowledgements: The first two authors have been supported by INdAM - GNAMPA Project 2015: Set-valued Analysis and Optimal Transportation Theory Methods in Deterministic and Stochastics Models of Financial Markets with Transaction Costs.

