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 finite-dimensional 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-Jacobi-Bellman equation in the space of probability measures solved by the generalized minimum time function, by extending a concept of approximate viscosity sub/superdifferential in the space of probability measures, originally introduced by Cardaliaguet-Quincampoix (2008). 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 dynamics of the form

$$\begin{cases} \dot{x}(t) \in F(x(t)), \quad t > 0, \\ \\ x(0) = x_0, \end{cases} \tag{1.1}$$

where F is a set-valued map from $\mathbb{R}^d$ to $\mathbb{R}^d$ whose value at each point denotes 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 [12, 13] and to references therein, for recent results on the regularity of T in the framework of differential inclusions.

Date: May 2, 2017.

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.*

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 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 $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ 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 μ_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

$$\begin{cases} \partial_t \mu(t,x) + \mathrm{div}(v_t(x)\mu(t,x)) = 0, & \text{for } t>0,\ x\in\mathbb{R}^d, \\ \\ \mu(0,\cdot) = \mu_0. \end{cases} \tag{1.2}$$

which represents the conservation of the total mass $\mu_0(\mathbb{R}^d)$ 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 $\mu(t,\mathbb{R}^d) = \mu_0(\mathbb{R}^d)$ 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 μ_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 $\tilde{S}^\Phi \subseteq \mathscr{P}(\mathbb{R}^d)$ all the probability measures representing states which make the result of the measurements nonpositive. If we take for instance $\Phi = \{d_S(\cdot)\}$, the generalized target in $\mathscr{P}(\mathbb{R}^d)$ 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 $\mathscr{P}(\mathbb{R}^d)$ 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 variations in infinite dimensional spaces. We refer to [3] and [28] 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 3.20);
- a comparison result between classical and generalized minimum time functions in some cases (Proposition 3.10);
- a sufficient condition for the generalized minimum time function to be finite, with an upper estimate based on the initial data (Theorem 3.26);
- 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 4.9).

Recent works (see e.g. [2, 22]) 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 finite-dimensional case.

Related to such a problem, a further application could be the theory of mean field games [24, 25]. 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 Hamilton–Jacobi 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 [26], 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 measures, in the spirit of [1]. The definition of an object in the measure-theoretic setting which corresponds to the Lie brackets in the finite-dimensional context is in the purpose of [16].

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 [27]. 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 first introduce a definition of generalized target and then we give two definitions of generalized minimum time functions, providing some comparison results between them and with the classical minimum time function, then we prove the Existence Theorem 3.20 and the Attainability Theorem 3.26. Finally, in Section 4 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.

Our main references for preliminaries on measure theory are [3] and [28].

Let X be a separable metric space. $\mathscr{P}(X)$ stands for the set of Borel probability measures on X endowed with narrow convergence, $\mathscr{M}^+(X)$ denotes the set of positive and finite Radon measures on X and $\mathscr{M}(X;\mathbb{R}^d)$ the set of vector-valued Radon measures on X. We recall that $\mathscr{P}(X)$ can be identified with a convex subset of the unitary ball of the dual space $(C^0_b(X))'$, and that narrow convergence is induced by the weak*-topology on the dual space $(C^0_b(X))'$.

Let X,Y be separable metric spaces, the *push-forward* of a measure $\mu \in \mathscr{P}(X)$ through a Borel map $r : X \to Y$ is defined by $r\sharp\mu(B) := \mu(r^{-1}(B)) \in \mathscr{P}(Y)$, for all Borel sets $B \subseteq Y$, or equivalently it is defined by

$$\int_X f(r(x))\,d\mu(x) = \int_Y f(y)\,dr\sharp\mu(y),$$

for every bounded (or $r\sharp\mu$-integrable) Borel function $f : Y \to \mathbb{R}$. For properties of push-forward we cite [3], Chapter 5, Section 2.

Let $\mu \in \mathscr{P}(\mathbb{R}^d)$, $p \geq 1$, we say that μ has finite p-moment if

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

and $\mathscr{P}_p(\mathbb{R}^d)$ denotes the subset of $\mathscr{P}(\mathbb{R}^d)$ made of measures with finite p-moment.

Definition 2.1 (Wasserstein distance)**.** Given $\mu_1, \mu_2 \in \mathscr{P}(\mathbb{R}^d)$, $p \geq 1$, we define the *p-Wasserstein distance* between μ_1 and μ_2 by setting

$$W_p(\mu_1,\mu_2) := \left(\inf\left\{\iint_{\mathbb{R}^d\times\mathbb{R}^d} |x_1 - x_2|^p\,d\pi(x_1,x_2)\ :\ \pi \in \Pi(\mu_1,\mu_2)\right\}\right)^{1/p}, \tag{2.1}$$

where the set of *admissible transport plans* $\Pi(\mu_1,\mu_2)$ is defined by

$$\Pi(\mu_1,\mu_2):=\Bigg\{\pi\in\mathscr{P}(\mathbb{R}^d\times\mathbb{R}^d):\ \begin{array}{l}\pi(A_1\times\mathbb{R}^d)=\mu_1(A_1),\\ \pi(\mathbb{R}^d\times A_2)=\mu_2(A_2),\end{array}\ \text{for all } \mu_i\text{-measurable sets } A_i,\ i=1,2\Bigg\}.$$

We also denote with $\Pi_o^p(\mu_1,\mu_2)$ the subset of $\Pi(\mu_1,\mu_2)$ consisting of optimal transport plans, i.e. the set of all plans π for which the infimum in (2.1) is attained. We will also use the notation $\Pi_o(\mu_1,\mu_2)$ when the context makes clear which distance W_p is being considered.

Proposition 2.2. *$\mathscr{P}_p(\mathbb{R}^d)$ endowed with the p-Wasserstein metric $W_p(\cdot,\cdot)$ is a complete separable metric space. Moreover, given a sequence $\{\mu_n\}_{n\in\mathbb{N}}\subseteq\mathscr{P}_p(\mathbb{R}^d)$ and $\mu\in\mathscr{P}_p(\mathbb{R}^d)$, we have that the following are equivalent*

(1) $\lim\limits_{n\to\infty} W_p(\mu_n,\mu)=0$,

(2) $\mu_n \rightharpoonup^* \mu$ *and $\{\mu_n\}_{n\in\mathbb{N}}$ has uniformly integrable p-moments.*

Given $\mu_1,\mu_2\in\mathscr{P}(\mathbb{R}^d)$, $p\geq 1$, the following dual representation (called Monge–Kantorovich duality*) holds*

$$W_p^p(\mu_1,\mu_2)= \tag{2.2}$$
$$=\sup\left\{\int_{\mathbb{R}^d}\varphi(x_1)\,d\mu_1(x_1)+\int_{\mathbb{R}^d}\psi(x_2)\,d\mu_2(x_2)\ :\ \begin{array}{l}\varphi,\psi\in C_b^0(\mathbb{R}^d)\\ \varphi(x_1)+\psi(x_2)\leq|x_1-x_2|^p\\ \textit{for } \mu_i\textit{–a.e. } x_i\in\mathbb{R}^d\end{array}\right\}.$$

Proof. See Lemma 5.1.7, Proposition 7.1.5 and Theorem 6.1.1 in [3]. □

For other properties of the Wasserstein distance we refer for example to Chapter 6 in [28] or Section 7.1 in [3].

Theorem 2.3 (Superposition principle)**.** *Let $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$ be a solution of the continuity equation $\partial_t\mu_t+\mathrm{div}(v_t\mu_t)=0$ for a suitable Borel vector field $v:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ satisfying*

$$\int_0^T\int_{\mathbb{R}^d}\frac{|v_t(x)|}{1+|x|}\,d\mu_t(x)\,dt<+\infty\,.$$

Then there exists a probability measure $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$, with $\Gamma_T=C^0([0,T];\mathbb{R}^d)$ 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 γ is an absolutely continuous solution of*

$$\begin{cases}\dot\gamma(t)=v_t(\gamma(t)), & \textit{for } \mathscr{L}^1\textit{-a.e } t\in(0,T)\\ \gamma(0)=x,\end{cases}$$

(ii) *for all $t\in[0,T]$ and all $\varphi\in C_b^0(\mathbb{R}^d)$ 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 $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$ as in (ii) above, we have that $\partial_t\mu_t+\mathrm{div}(v_t\mu_t)=0$ and $\mu_{|t=0}=\gamma(0)\sharp\boldsymbol{\eta}$.

Proof. See Theorem 5.8 in [9] and Theorem 8.2.1 in [3]. □

We recall now some preliminaries about differential inclusions governing the classical control problem. For this part, our main references are [7] and [8].

Definition 2.4 (Standing Assumptions)**.** We will say that a set-valued function $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ satisfies the assumption (F_j), $j = 0, 1, 2$ if the following hold true

- (F_0) $F(x) \neq \emptyset$ 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 exists a constant $C > 0$ such that $F(x) \subseteq \overline{B(0, C(|x| + 1))}$ for every $x \in \mathbb{R}^d$.
- (F_2) $F(\cdot)$ 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 2.5. *Under assumptions (F_0) and (F_1), the differential inclusion*

$$\dot{x}(t) \in F(x(t))\,, \tag{2.3}$$

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 (2.3) for a.e. $t \geq 0$. Moreover, the set of trajectories of the differential inclusions (2.3) 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 [20]. □

The following simple classical lemma will be used.

Lemma 2.6 (A priori estimate on differential inclusions)**.** *Assume (F_0) and (F_1). Let $K \subset \mathbb{R}^d$ be compact and $T > 0$ and set $|K| = \max_{y \in K} |y|$. Then, for all Carathéodory solutions $\gamma : [0, T] \to \mathbb{R}^d$ of (2.3) we have*

- (i) *forward estimate: if $\gamma(0) \in K$ then $|\gamma(t)| \leq (|K| + CT)\, e^{CT}$ for all $t \in [0, T]$;*
- (ii) *backward estimate: if $\gamma(T) \in K$ then $|\gamma(t)| \leq (|K| + CT)\, e^{CT}$ for all $t \in [0, T]$,*

where C is the constant in (F_1).

Proof. Recalling that $\dot{\gamma}(s) \in F(\gamma(s))$ for a.e. $s \in [0, T]$ and that $F(\gamma(s)) \subseteq \overline{B(0, C(|x| + 1))}$, we have

$$|\gamma(t)| \leq |\gamma(0)| + \int_0^t |\dot{\gamma}(s)|\, ds \leq |K| + CT + C \int_0^t |\gamma(s)|\, ds\,.$$

According to Gronwall's inequality, we then have $|\gamma(t)| \leq (|K| + CT)e^{Ct}$, 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 (F_0) and (F_1) and $w(0) \in K$, the previous analysis implies

$$|\gamma(t)| = |w(T - t)| \leq (|K| + CT)\, e^{C(T-t)},$$

whence (ii) follows. □

Definition 2.7 (Weak invariance)**.** Given a set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, 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 (2.3), 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 [20].

Given $T \in [0, +\infty[$, we set

$$\Gamma_T := C^0([0,T];\mathbb{R}^d), \qquad \Gamma_T^x := \{\gamma \in \Gamma_T : \gamma(0) = x \in \mathbb{R}^d\},$$

endowed with the usual sup-norm, where we recall that Γ_T is a complete separable metric space for every $0 < T < +\infty$. The *evaluation map* $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ is defined by $e_t(x,\gamma) = \gamma(t)$ for all $0 \le t \le T$.

Let X be a set, $A \subseteq X$. In the following, the *indicator function of* A is the function $I_A : X \to \{0, +\infty\}$ 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 $\chi_A : X \to \{0,1\}$ defined as $\chi_A(x) = 1$ for all $x \in A$ and $\chi_A(x) = 0$ for all $x \notin A$.

If X is a Banach space, X' its topological dual, $A \subseteq X$ nonempty, we denote with $\sigma_A : X' \to [-\infty, +\infty]$ the *support function* to A, defined by $\sigma_A(x^*) := \sup\limits_{x\in A} \langle x^*, x\rangle_{X',X}$.

3. Generalized minimum time problem

In this section we first propose a suitable generalization of the classical target set that will be used in our framework in the space of probability measures, and then we define a suitable notion of minimum time function, modeled on the finite-dimensional case.

Definition 3.1 (Generalized targets)**.** Let $p \ge 1$, Φ be a given set of lower semicontinuous maps from $\mathbb{R}^d$ to $\mathbb{R}$, such that the following property holds

(T_E) there exists $x_0 \in \mathbb{R}^d$ with $\phi(x_0) \le 0$ for all $\phi \in \Phi$, and all $\phi \in \Phi$ are bounded from below.

We define the *generalized targets* $\tilde S^\Phi$ and $\tilde S^\Phi_p$ as follows

$$\tilde S^\Phi := \left\{ \mu \in \mathscr{P}(\mathbb{R}^d) : \phi \in L^1_\mu \text{ and } \int_{\mathbb{R}^d} \phi(x)\, d\mu(x) \le 0 \text{ for all } \phi \in \Phi \right\}, \qquad \tilde S^\Phi_p := \tilde S^\Phi \cap \mathscr{P}_p(\mathbb{R}^d).$$

It can be proved that $\tilde S^\Phi$ is w^*-closed in $\mathscr{P}(\mathbb{R}^d)$ and $\tilde S^\Phi_p$ is W_p-closed in $\mathscr{P}_p(\mathbb{R}^d)$.

When we can write $\Phi = \{d_S\}$, with $S \subseteq \mathbb{R}^d$ closed and nonempty, then we will say that $\tilde S^\Phi$ (or $\tilde S^\Phi_p$) *admits a classical counterpart*, or that S *is the classical counterpart of* $\tilde S^\Phi$ (or $\tilde S^\Phi_p$).

We define also the *generalized distance* from $\tilde S^\Phi_p$ as

$$\tilde d_{\tilde S^\Phi_p}(\cdot) := \inf_{\mu \in \tilde S^\Phi_p} W_p(\cdot, \mu).$$

Notice that $\tilde S^\Phi_p \neq \emptyset$ because $\delta_{x_0} \in \tilde S^\Phi_p$, hence $\tilde S^\Phi \neq \emptyset$. The 1-Lipschitz continuity of $\tilde d_{\tilde S^\Phi_p}(\cdot)$ is trivial.

We refer the reader to [17] or [18] for an analysis of the properties of these objects.

Definition 3.2 (Admissible curves)**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, $I = [a,b]$ a compact interval of $\mathbb{R}$, $\alpha, \beta \in \mathscr{P}(\mathbb{R}^d)$. We say that a Borel family of probability measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in I} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an *admissible trajectory (curve) defined in I for the system Σ_F joining α and β*, if there exists a family of Borel vector-valued measures $\boldsymbol{\nu} = \{\nu_t\}_{t\in I} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that

(1) $\boldsymbol{\mu}$ is a narrowly continuous (i.e., continuous w.r.t. the topology induced by the set $C^0_b(\mathbb{R}^d)$ of real-valued continuous bounded functions on $\mathbb{R}^d$) solution in the distributional sense of $\partial_t\mu_t+\mathrm{div}(\nu_t)=0$, with $\mu_{|t=a}=\alpha$ and $\mu_{|t=b}=\beta$.

(2) $J_F(\boldsymbol{\mu},\boldsymbol{\nu})<+\infty$, where $J_F(\cdot,\cdot)$ is defined as

$$J_F(\boldsymbol{\mu},\boldsymbol{\nu}):=\begin{cases} b-a, & \text{if } |\nu_t|\ll\mu_t \text{ for } \mathscr{L}^1\text{-a.e. } t\in I,\\ & \text{and } v_t(x):=\dfrac{\nu_t}{\mu_t}(x)\in F(x) \text{ for } \mu_t\text{-a.e. } x,\ \mathscr{L}^1\text{-a.e. } t\in I,\\ \\ +\infty, & \text{otherwise.}\end{cases} \tag{3.1}$$

In this case, we will also shortly say that $\boldsymbol{\mu}$ is *driven* by $\boldsymbol{\nu}$.

Remark 3.3. The finiteness of $J_F(\boldsymbol{\mu},\boldsymbol{\nu})$ forces the elements of $\boldsymbol{\nu}$ to have the form $\nu_t=v_t\mu_t$ for a vector field $v_t\in L^1_{\mu_t}$ for a.e. $t\in I$, and moreover we have $v_t(x)\in F(x)$ for μ_t–a.e. $x\in\mathbb{R}^d$ and a.e. $t\in I$. When $J_F(\cdot,\cdot)$ is finite, this value expresses the time needed by the system Σ_F to steer α to β along the trajectory $\boldsymbol{\mu}$ with family of velocity vector fields $v=\{v_t\}_{t\in I}$.

In view of the superposition principle stated at Theorem 2.3, we can give the following alternative equivalent definition.

Definition 3.4 (Admissible curves (alternative definition))**.** Let $F:\mathbb{R}^d\rightrightarrows\mathbb{R}^d$ be a set-valued function satisfying (F_1), $I=[a,b]$ a compact interval of $\mathbb{R}$, $\alpha,\beta\in\mathscr{P}(\mathbb{R}^d)$. We say that a Borel family of probability measures $\boldsymbol{\mu}=\{\mu_t\}_{t\in I}$ is an *admissible trajectory (curve) defined in I for the system Σ_F joining α and β*, if there exist a probability measure $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_I)$ and a Borel vector field $v:I\times\mathbb{R}^d\to\mathbb{R}^d$ such that:

(1) $\boldsymbol{\eta}$ is concentrated on the pairs (x,γ) such that γ is an absolutely continuous solution of $\dot{x}(t)=v_t(x(t))$ with initial condition $\gamma(a)=x$;
(2) for every $\varphi\in C^0_b(\mathbb{R}^d)$, $t\in I$ we have
$$\int_{\mathbb{R}^d}\varphi(x)\,d\mu_t(x)=\iint_{\mathbb{R}^d\times\Gamma_I}\varphi(\gamma(t))\,d\boldsymbol{\eta}(x,\gamma),$$
(3) $\gamma(a)\sharp\boldsymbol{\eta}=\alpha$, $\gamma(b)\sharp\boldsymbol{\eta}=\beta$,
(4) $v_t(x)\in F(x)$ for μ_t-a.e. $x\in\mathbb{R}^d$ and a.e. $t\in I$ and $v_t\in L^1_{\mu_t}$ 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 $T>0$. We introduce the following notation.

Definition 3.5. Given $T\in[0,+\infty[$, we set

$$\begin{aligned}\mathscr{T}_F(\mu_0):=\{\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T):\ T>0,\boldsymbol{\eta}\text{ concentrated on trajectories of}\\ \dot{\gamma}(t)\in F(\gamma(t))\text{ and satisfies }\gamma(0)\sharp\boldsymbol{\eta}=\mu_0\},\end{aligned}$$

where $\mu_0\in\mathscr{P}(\mathbb{R}^d)$.

By the Superposition Principle (Theorem 2.3), given $F:\mathbb{R}^d\rightrightarrows\mathbb{R}^d$ satisfying (F_1), a Borel family of probability measures $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}$ is an *admissible trajectory* if and only if there exists $\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)$ such that $\mu_t=e_t\sharp\boldsymbol{\eta}$ for all $t\in[0,T]$, i.e., $\boldsymbol{\eta}=\mu_0\otimes\eta_x$ where for μ_0-a.e. $x\in\mathbb{R}^d$ we have that $\eta_x\in\mathscr{P}(\Gamma^x_T)$ is concentrated on the solutions of $\dot{x}(t)\in F(x(t))$, $x(0)=x$.

In this case, we will shortly say that the admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *is represented by* $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$.

For later use we state the following technical lemma.

Lemma 3.6 (Basic estimates). *Assume* (F_0) *and* (F_1)*, and let* C *be the constant as in* (F_1)*. Let* $T > 0$*,* $p \geq 1$*,* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ *and* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *be an admissible trajectory driven by* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ *and represented by* $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$*. Then we have:*

(i) $|e_t(x,\gamma)| \leq (|e_0(x,\gamma)|+CT)\, e^{CT}$ *for all* $t \in [0,T]$ *and* $\boldsymbol{\eta}$*-a.e.* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$*;*

(ii) $e_t \in L^p_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$ *for all* $t \in [0,T]$*;*

(iii) *there exists* $D > 0$ *depending only on* C, T, p *such that for all* $t \in [0,T]$ *we have*

$$\left\| \frac{e_t - e_0}{t} \right\|^p_{L^p_{\boldsymbol{\eta}}} \leq D\,(\mathrm{m}_p(\mu_0) + 1)\,;$$

(iv) *there exist* $D', D'' > 0$ *depending only on* C, T, p *such that for all* $t \in [0,T]$ *we have*

$$\begin{aligned} \mathrm{m}_p(\mu_t) &\leq D'(\mathrm{m}_p(\mu_0) + 1), \\ \mathrm{m}_p(|\nu_t|) &\leq D''(\mathrm{m}_{p+1}(\mu_0) + 1). \end{aligned}$$

In particular, we have $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$*.*

Proof. Item (i) follows from Lemma 2.6. To prove (ii) it is enough to show $e_0 \in L^p_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T)$ and then apply item (i). Indeed, recalling that $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ for any $a, b \geq 0$, we have

$$\iint_{\mathbb{R}^d\times\Gamma_T} |e_0(x,\gamma)|^p \, d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d} |z|^p \, d(\gamma(0)\sharp\boldsymbol{\eta})(z) = \mathrm{m}_p(\mu_0) < +\infty,$$

$$\begin{aligned} \iint_{\mathbb{R}^d\times\Gamma_T} &|e_t(x,\gamma)|^p \, d\boldsymbol{\eta}(x,\gamma) \leq \\ &\leq 2^{p-1} e^{CTp} \left(\iint_{\mathbb{R}^d\times\Gamma_T} |e_0(x,\gamma)|^p \, d\boldsymbol{\eta} + C^p T^p \right) \\ &\leq K\,(\mathrm{m}_p(\mu_0) + 1)\,, \end{aligned}$$

for a suitable constant $K > 0$ depending only on C, T, p.

For the proof of $(iii), (iv)$ we refer to Lemma 3.2 in the forthcoming paper [15] for the case $p = 2$. □

The following definitions are the natural counterpart of the classical case.

Definition 3.7 (Reachable set). Let $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, and $T > 0$. Define the set of *admissible curves defined on* $[0,T]$ *and starting from* μ_0 by setting

$$\mathscr{A}_T(\mu_0) := \{\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d) : \boldsymbol{\mu} \text{ is an admissible trajectory with } \mu_{|t=0} = \mu_0\}.$$

The *reachable set* from μ_0 in time T is

$$\mathscr{R}_T(\mu_0) := \{\mu \in \mathscr{P}(\mathbb{R}^d) : \text{ there exists } \boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathscr{A}_T(\mu_0) \text{ with } \mu = \mu_T\}.$$

Definition 3.8 (Generalized minimum time). Let $p \geq 1$, $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1, and $\tilde{S}^\Phi$, $\tilde{S}^\Phi_p$ be the corresponding generalized targets defined in Definition 3.1. In analogy with the classical case, we define the *generalized minimum time function* $\tilde{T}^\Phi : \mathscr{P}(\mathbb{R}^d) \to [0,+\infty]$ by setting

$$\tilde{T}^\Phi(\mu_0) := \inf\left\{ J_F(\boldsymbol{\mu},\boldsymbol{\nu}) : \boldsymbol{\mu} \in \mathscr{A}_T(\mu_0),\ \boldsymbol{\mu} \text{ is driven by } \boldsymbol{\nu},\ \mu_{|t=T} \in \tilde{S}^\Phi \right\}, \tag{3.2}$$

where, by convention, $\inf\emptyset=+\infty$.

Given $\mu_0\in\mathscr{P}(\mathbb{R}^d)$ with $T^\Phi(\mu_0)<+\infty$, an admissible curve $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,\tilde T^\Phi(\mu_0)]}\subseteq\mathscr{P}(\mathbb{R}^d)$, driven by a family of Borel vector-valued measures $\boldsymbol{\nu}=\{\nu_t\}_{t\in[0,\tilde T^\Phi(\mu_0)]}$ and satisfying $\mu_{|t=0}=\mu_0$ and $\mu_{|t=\tilde T^\Phi(\mu_0)}\in\tilde S^\Phi$ is *optimal for* μ_0 if

$$\tilde T^\Phi(\mu_0)=J_F(\boldsymbol{\mu},\boldsymbol{\nu}).$$

Given $p\geq 1$, we define also a generalized minimum time function $\tilde T^\Phi_p:\mathscr{P}_p(\mathbb{R}^d)\to[0,+\infty]$ by replacing in the above definitions $\tilde S^\Phi$ by $\tilde S^\Phi_p$ and $\mathscr{P}(\mathbb{R}^d)$ by $\mathscr{P}_p(\mathbb{R}^d)$. Since $\tilde S^\Phi_p\subseteq\tilde S^\Phi$, it is clear that $\tilde T^\Phi(\mu_0)\leq\tilde T^\Phi_p(\mu_0)$.

Remark 3.9. In view of the characterization in Theorem 8.3.1 in [3], and of Remark 3.3, one can think to $\tilde T^\Phi$ as the minimum time needed by the system to steer μ_0 to a measure in $\tilde S^\Phi$, along absolutely continuous curves in $\mathscr{P}_p(\mathbb{R}^d)$.

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 3.10 (First comparison between $\tilde T^\Phi$ and T)**.** *Consider the generalized minimum time problem for Σ_F as in Definition 3.8 assuming (F_0), (F_1), and suppose that the corresponding generalized target $\tilde S^\Phi$ admits S as classical counterpart. Then for all $\mu_0\in\mathscr{P}(\mathbb{R}^d)$ we have*

$$\tilde T^\Phi(\mu_0)\geq\|T\|_{L^\infty_{\mu_0}},$$

where $T:\mathbb{R}^d\to[0,+\infty]$ is the classical minimum time function for the 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 $\Phi=\{d_S\}$ by assumption of existence of the classical counterpart S for $\tilde S^\Phi$.

If $\tilde T(\mu_0)=+\infty$ there is nothing to prove, so assume $\tilde T(\mu_0)<+\infty$. Fix $\varepsilon>0$ and let $\boldsymbol{\mu}=\{\mu_t\}_{t\in[0,T]}\subseteq\mathscr{P}(\mathbb{R}^d)$ be an admissible curve starting from μ_0, driven by a family of Borel vector-valued measures $\boldsymbol{\nu}=\{\nu_t\}_{t\in I}$ such that $T=J_F(\boldsymbol{\mu},\boldsymbol{\nu})<\tilde T(\mu_0)+\varepsilon$ and $\mu_{|t=T}\in\tilde S$. In particular, we have that $v_t(x):=\dfrac{\nu_t}{\mu_t}(x)\in F(x)$ for μ_t-a.e. $x\in\mathbb{R}^d$ and a.e. $t\in[0,T]$, hence $|v_t(x)|\leq C(1+|x|)$ for μ_t-a.e $x\in\mathbb{R}^d$. Accordingly,

$$\int_0^T\int_{\mathbb{R}^d}\frac{|v_t(x)|}{1+|x|}\,d\mu_t\,dt\leq CT<+\infty.$$

By the Superposition Principle (Theorem 2.3), recalling Definition 3.5, we have that there exists a probability measure $\boldsymbol{\eta}=\mu_0\otimes\eta_x\in\mathscr{T}_F(\mu_0)$ such that for μ_0-a.e $x\in\mathbb{R}^d$, the measure $\eta_x\in\mathscr{P}(\Gamma^x_T)$ is concentrated on absolutely continuous curves γ satisfying $\dot\gamma(t)=v_t(\gamma(t))$ for a.e. t, and $\mu_t=e_t\sharp\mu_0$. In particular, if $x\notin\operatorname{supp}\mu_0$ or $\gamma(0)\neq x$, then $(x,\gamma)\notin\operatorname{supp}\boldsymbol{\eta}$.

Let $\{\psi_n\}_{n\in\mathbb{N}}\in C^\infty_c(\mathbb{R}^d;[0,1])$ 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 $\{\psi_n(\cdot)d_S(\cdot)\}_{n\in\mathbb{N}}\subseteq C^0_b(\mathbb{R}^d)$ is an increasing sequence of nonnegative functions pointwise convergent to $d_S(\cdot)$, we have for every $t \in [0,T]$

$$\begin{aligned}\iint_{\mathbb{R}^d\times\Gamma_T} d_S(\gamma(t))\,d\boldsymbol{\eta}(x,\gamma) &= \lim_{n\to\infty}\iint_{\mathbb{R}^d\times\Gamma_T}\psi_n(\gamma(t))d_S(\gamma(t))\,d\boldsymbol{\eta}(x,\gamma)\\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d}\psi_n(x)d_S(x)\,d\mu_t(x)\end{aligned}$$

By taking $t = T$, we have that the last term vanishes because $\mu_{|t=T} \in \tilde{S}$ and so $\operatorname{supp}\mu_{|t=T} \subseteq S$, 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. $(x,\gamma) \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$, whence $T \geq T(x)$ for μ_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, $\tilde{T}(\mu_0)+\varepsilon \geq T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$ and, by letting $\varepsilon \to 0$, we conclude that $\tilde{T}(\mu_0) \geq \|T\|_{L^\infty_{\mu_0}}$. □

We notice that the inequality appearing in Proposition 3.10 may be strict without further assumptions.

Example 3.11. 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 $\tilde{S}^\Phi = \{\delta_0\}$. Moreover, we have $T(x) = |x|$ for $x \leq 0$ and $T(x) = +\infty$ for $x > 0$. Define $\mu_0 = \dfrac{1}{2}(\delta_{-2} + \delta_{-1})$. We have $\|T\|_{L^\infty_{\mu_0}} = \max\{T(-1), T(-2)\} = 2$. 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 μ_0 and δ_0 is empty, hence $\tilde{T}^\Phi(\mu_0) = +\infty$.

Remark 3.12. This implies that in general the problem of the generalized minimum time *cannot be reduced* to the underlying finite dimensional control problem, even in the cases where the underlying control problem is particulary simple. A consequence of this fact is that even if the underlying system enjoys some properties as closure and relative compactness of the set of admissible trajectories (provided for instance by good assumptions on the set-valued map F), which lead to the existence of optimal trajectories for the problem, in our generalized framework all these results must be proved.

Definition 3.13 (Convergence of curves in $\mathscr{P}(\mathbb{R}^d)$)**.** We say that a family of curves $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T]}$ in $\mathscr{P}(\mathbb{R}^d)$

(1) *pointwise converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}(\mathbb{R}^d)$ if and only if $\mu^n_t \rightharpoonup^* \mu_t$ for all $t \in [0,T]$. In this case we will write $\boldsymbol{\mu}^n \rightharpoonup^* \boldsymbol{\mu}$.
(2) *pointwise converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}_p(\mathbb{R}^d)$ if and only if $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ and $\lim_{n\to+\infty} W_p(\mu^n_t, \mu_t) = 0$ for all $t \in [0,T]$. In this case we will write $\boldsymbol{\mu}^n \to^p \boldsymbol{\mu}$.
(3) *uniformly converges* to a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ in $\mathscr{P}_p(\mathbb{R}^d)$ if and only if $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ and

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

In this case we will write $\boldsymbol{\mu}^n \rightrightarrows^p \boldsymbol{\mu}$.

Lemma 3.14. *Assume that* $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ *satisfies* (F_0)*. Then the functional* $\mathscr{F} : \mathscr{P}(\mathbb{R}^d) \times \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d) \to \{0, +\infty\}$ *defined by*

$$\mathscr{F}(\mu, E) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} I_{F(x)}\left(\frac{E}{\mu}(x)\right) d\mu(x), & \text{if } E \ll \mu, \\ \\ +\infty, & \text{otherwise} \end{cases} \tag{3.3}$$

is l.s.c. w.r.t. narrow convergence.

Proof. Define $f(x, v) = I_{F(x)}(v)$. Since F is u.s.c. with convex values, we have that $f(\cdot, \cdot)$ is l.s.c. and $f(x, \cdot)$ is convex. By compactness of $F(x)$, we have that the domain of $f(x, \cdot)$ is bounded, thus following the notation in [11] we have $f_\infty(x, v) = 0$ if $v = 0$ and $f_\infty(x, v) = +\infty$ if $v \neq 0$, where $f^\infty(x, \cdot)$ denotes the *recession function* for $f(x, \cdot)$. By l.s.c. of F, there exists a continuous selection $z_0 : \mathbb{R}^d \to \mathbb{R}^d$ of F, i.e., there exists $z_0 \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $z_0(x) \in F(x)$ for all $x \in \mathbb{R}^d$. Thus $x \mapsto f(x, z_0(x))$ is continuous and finite. Hence, the functional (3.3) is l.s.c. w.r.t. a.e. pointwise weak* convergence of measures (see Lemma 2.2.3, p. 39, Theorem 3.4.1, p.115, and Corollary 3.4.2 in [11] or Theorem 2.34 in [6]). □

Proposition 3.15 (Convergence of admissible trajectories)**.** *Assume* (F_0)*. Let* $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,T]}$ *be a sequence of admissible curves defined on* $[0, T]$ *such that* $\boldsymbol{\mu}^n$ *is driven by* $\boldsymbol{\nu}^n = \{\nu_t^n\}_{t\in[0,T]}$ *and suppose that there exist* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ *and* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ *such that for a.e.* $t \in [0, T]$ *it holds* $(\mu_t^n, \nu_t^n) \rightharpoonup^* (\mu_t, \nu_t)$*. Then* $\boldsymbol{\mu}$ *is an admissible trajectory driven by* $\boldsymbol{\nu}$*.*

Proof. Fix $t \in [0, T]$ such that $(\mu_t^n, \nu_t^n) \rightharpoonup^* (\mu_t, \nu_t)$ and $\mathscr{F}(\mu_t^n, \nu_t^n) = 0$ for all $n \in \mathbb{N}$. By l.s.c. of $\mathscr{F}$ and recalling that $\mathscr{F} \geq 0$, we have

$$0 \leq \mathscr{F}(\mu_t, \nu_t) \leq \liminf_{n\to+\infty} \mathscr{F}(\mu_t^n, \nu_t^n) = 0,$$

and so for a.e. $t \in [0, T]$ we have $\dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$.

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

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi(x) d\mu_t^n(x) = \int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t^n(x),$$

and for the last term we have

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t^n(x) = \int_{\mathbb{R}^d} \nabla\varphi(x)\, d\nu_t(x),$$

due to the w^*-convergence of ν_t^n to ν_t, thanks to Lemma 8.1.2 in [3], we deduce that, up to changing μ_t and ν_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 3.16. *Assume* (F_0), (F_1)*. Let* $\mathscr{A}$ *be a set of admissible trajectories defined on* $[0, T]$ *and* $C_1 > 0$, $p > 1$ *be constants such that for all* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \in \mathscr{A}$ *it holds* $\mathrm{m}_p(\mu_t) \leq C_1$ *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 $\{\mathrm{m}_p(\mu_0) :$ *there exists* $\boldsymbol{\mu} \in \mathscr{A}$ *with* $\mu_{|t=0} = \mu_0\}$ *is bounded, and, in particular, it holds for* $\mathscr{A}_T(\mu_0)$ *when* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$*.*

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

$$\int_{\mathbb{R}^d} |x|^p \, d\mu_t^n(x) \le C_1,$$

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

$$\int_{\mathbb{R}^d} |x|^{p-1} |d\nu_t^n(x)| = \int_{\mathbb{R}^d} |x|^{p-1} |v_t^n(x)| \, d\mu_t^n(x) \le LC_1 + 1,$$

for a constant $L > 0$. Thus there exists $\nu_t \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\nu_t^n \rightharpoonup^* \nu_t$. By Proposition 3.15, we have that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory defined on $[0,T]$ driven by $\boldsymbol{\nu}$. The last assertion comes from Lemma 3.6, which allows to estimate the moments of μ_t and ν_t in terms of the moments of μ_0. □

Theorem 3.17 (L.s.c. of the generalized minimum time)**.** *Assume* (F_0) *and* (F_1)*. Then* $\tilde{T}_p^\Phi : \mathscr{P}_p(\mathbb{R}^d) \to [0,+\infty]$ *is l.s.c. for all* $p > 1$*.*

Proof. Let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, we have to prove that $\tilde{T}_p^\Phi(\mu_0) \le \liminf\limits_{W_p(\mu,\mu_0)\to 0} \tilde{T}_p^\Phi(\mu)$. Taken a sequence $\{\mu_0^n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ s.t. $W_p(\mu_0^n,\mu_0) \to 0$ for $n \to +\infty$, and $\liminf\limits_{W_p(\mu,\mu_0)\to 0} \tilde{T}_p^\Phi(\mu) = \lim\limits_{n\to+\infty} \tilde{T}_p^\Phi(\mu_0^n) =: T$, we want to prove that $\tilde{T}_p^\Phi(\mu_0) \le T$.

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

Without loss of generality, we can assume that all $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ are defined in an interval containing $[0,T]$, since if $T_n < T$ we can use the gluing Lemma 4.4 in [21] and extend $\boldsymbol{\mu}^n$ to a trajectory defined in $[0,T]$ simply by taking any Borel selection $\bar{v}$ of $F(\cdot)$ (which exists by (F_0) and by Theorem 8.1.3 in [8]), and considering the solution of the continuity equation $\partial_t\mu_t + \operatorname{div} \bar{v}\mu_t = 0$ in $]T_n,T]$ with $\mu_{|t=T_n} = \mu^n_{T_n}$. Now, since μ_0^n converges in W_p to μ_0, we have that there exists $\bar{n} > 0$ such that the set $\{\mathrm{m}_p(\mu_0^n) : n > \bar{n}\}$ is uniformly bounded by $\mathrm{m}_p(\mu_0)+1$. Then, by Lemma 3.6 and by Theorem 3.16 there exists an admissible trajectory $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ such that $\boldsymbol{\mu}^n \to^p \boldsymbol{\mu}$, $n \to +\infty$, up to subsequences and $\mu_{|t=0} = \mu_0$. Recalling Theorem 8.3.1 in [3], for all $n \in \mathbb{N}$ we have

$$\begin{aligned}\tilde{d}_{\tilde{S}_p^\Phi}(\mu_T) &\le W_p(\mu_T, \mu^n_{T_n}) \le W_p(\mu_T,\mu_T^n) + W_p(\mu_T^n, \mu^n_{T_n}) \\ &\le W_p(\mu_T,\mu_T^n) + \left| \int_{T_n}^T \left\| \frac{\nu_t^n}{\mu_t^n} \right\|_{L^p_{\mu_t^n}} dt \right|.\end{aligned}$$

If we show a uniform bound on $\left\| \frac{\nu_t^n}{\mu_t^n} \right\|_{L^p_{\mu_t^n}}$, then by letting $n \to +\infty$ we have that $\mu_T \in \tilde{S}_p^\Phi$, thus $\tilde{T}_p^\Phi(\mu_0) \le T$ and the proof is concluded.

For a.e. $t \in [0,T]$ and μ_t^n-a.e. x we have $\dfrac{\nu_t^n}{\mu_t^n}(x) \in F(x)$. By (F_1) there exists $C > 0$ such that

$$\left\| \frac{\nu_t^n}{\mu_t^n} \right\|_{L^p_{\mu_t^n}} \leq C \left(\mathrm{m}_p^{1/p}(\mu_t^n) + 1 \right).$$

We conclude by using the Lemma 3.6 to estimate $\mathrm{m}_p(\mu_t^n)$ in terms of $\mathrm{m}_p(\mu_0^n)$ and recalling that since μ_0^n converges to μ_0 in W_p, for n sufficiently large we have $\mathrm{m}_p(\mu_0^n) \leq \mathrm{m}_p(\mu_0) + 1$. □

Remark 3.18. Unfortunately, we have that $\tilde{T}_p^\Phi(\cdot)$ in general fails to be continuous, being just lower semicontinuous. Moreover, it seems to be quite a difficult problem to provide general necessary and sufficient conditions on problem data granting this continuity property. In the forthcoming paper [15], the author provides some sufficient conditions granting local Lipschitz continuity of $\tilde{T}_2^\Phi$. However, we can provide a simple example in which those sufficient conditions are not satisfied but we can still have continuity of the generalized minimum time function, as shown below.

Example 3.19. In $\mathbb{R}^2$, take $\Phi = \{\phi\}$, where ϕ is the 1-Lipschitz continuous map $\phi(x,y) = 1 - \displaystyle\int_{-\infty}^{x} e^{-|s|} ds \in C_b^1(\mathbb{R}^2;\mathbb{R})$. Let $F(x,y) := \{(\alpha, 0) : \alpha \in [0,1]\}$, $\mu_0 \in \mathscr{P}_2(\mathbb{R}^2)$. Since for every solution of $\gamma(t) \in F(\gamma(t))$, $\gamma(0) = (x,y)$ we have $\phi \circ \gamma(t) = \phi(x + \int_0^t \dot\gamma(s)\, ds, y) \geq \phi(x+t, y)$, due to the fact that $\partial_x \phi(x,y) = -e^{-|x|} < 0$, every trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t>0}$ defined by $\mu_t = (\mathrm{Id} + tv)\sharp\mu_0$ for $v = (1,0)$ is optimal for μ_0, moreover, if we define $G : \mathbb{R} \times \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by setting

$$G(t, \mu_0) := \int_{\mathbb{R}^2} \phi((x,y) + tv)\, d\mu_0 = \int_{\mathbb{R}^2} \phi(x,y)\, d\mu_t(x,y),$$

we have that for any $\mu_0 \notin \tilde{S}_2^\Phi$, it holds $G(t,\mu_0) = 0$ if and only if $t = \tilde{T}_2^\Phi(\mu_0)$, due to the strictly decreasing property of $G(t,\mu_0)$ w.r.t. t (due to the sign of $\partial_x \phi$). It is easy to see that G is continuous w.r.t. both the variables, moreover, since $\lim\limits_{t\to+\infty} G(t,\mu) = -1$, we have $\tilde{T}_2^\Phi(\mu) < +\infty$ for all $\mu \in \mathscr{P}_2(\mathbb{R}^2)$. Consider a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_2(\mathbb{R}^2) \setminus \tilde{S}_2^\Phi$, such that $W_2(\mu_n, \mu) \to 0$, then $G(\tilde{T}_2^\Phi(\mu_n), \mu_n) = 0$ for all $n \in \mathbb{N}$, and by joint continuity property of G, we have that $G\left(\limsup\limits_{n\to+\infty} \tilde{T}_2^\Phi(\mu_n), \mu \right) = 0$, thus $\tilde{T}_2^\Phi(\mu) = \limsup\limits_{n\to+\infty} \tilde{T}_2^\Phi(\mu_n)$, which proves upper semicontinuity of $\tilde{T}_2^\Phi$, and so continuity of $\tilde{T}_2^\Phi(\cdot)$ by Theorem 3.17.

Theorem 3.20 (Existence of minimizers)**.** *Assume (F_0), (F_1) and $p > 1$. Let $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1, and let $\tilde{S}^\Phi$ be the corresponding generalized target. Let $\tilde{T}^\Phi(\mu_0) < \infty$. Then there exists an admissible curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ which is optimal for μ_0, that is $\tilde{T}^\Phi(\mu_0) = J_F(\boldsymbol{\mu}, \boldsymbol{\nu})$. Moreover, we have also $\tilde{T}^\Phi(\mu_0) = \tilde{T}_p^\Phi(\mu_0)$.*

Proof. By the hypothesis of finiteness of $\tilde{T}^\Phi(\mu_0)$ and by definition of infimum we have that there exist $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and a sequence of admissible trajectories $\boldsymbol{\mu}^n = \{\mu_t^n\}_{t\in[0,t_n]}$, such that $\mu^n|_{t=0} = \mu_0$, $\mu^n|_{t=t_n} =: \sigma^n \in \tilde{S}^\Phi$, $t_n \to \tilde{T}^\Phi(\mu_0)^+$. Moreover, by

Lemma 3.6, we have that $\sigma^n \in \tilde{S}^\Phi_p$ for all $n \in \mathbb{N}$. We restrict all $\boldsymbol{\mu}^n$ to be defined on $[0, \tilde{T}^\Phi(\mu_0)]$.

By Theorem 3.16, $\boldsymbol{\mu}^n$ w^*-converges up to subsequences to an admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$ starting from μ_0 driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$, and by w^*-closure of $\tilde{S}^\Phi$ we have $\sigma^n \rightharpoonup^* \mu|_{t=\tilde{T}^\Phi(\mu_0)} \in \tilde{S}^\Phi$. Applying again Lemma 3.6, we have that $\mu|_{t=\tilde{T}^\Phi(\mu_0)} \in \tilde{S}^\Phi_p$. Thus $\tilde{T}^\Phi(\mu_0) = \tilde{T}^\Phi_p(\mu_0) = J_F(\boldsymbol{\mu}, \boldsymbol{\nu})$. □

The following results allow us to justify the name of *generalized minimum time* given to functions $\tilde{T}^\Phi(\cdot)$ and $\tilde{T}^\Phi_p(\cdot)$.

Lemma 3.21 (Convexity property of the embedding of classical trajectories)**.** *Let $N \in \mathbb{N}\backslash\{0\}$, $T > 0$ be given. Assume (F_0) and (F_1). Consider a family of continuous curves and real numbers $\{(\gamma_i, \lambda_i)\}_{i=1,\dots,N} \subseteq \Gamma_T \times [0, 1]$ such that $\gamma_i(\cdot)$ is a trajectory of $\dot{x}(t) \in F(x(t))$ for $i = 1, \dots, N$, and $\displaystyle\sum_{i=1}^N \lambda_i = 1$.*

For all $i = 1, \dots, N$ and $t \in [0, T]$, define the measures $\mu^{(i)}_t = \delta_{\gamma_i(t)}$, $\mu_t = \displaystyle\sum_{i=1}^N \lambda_i \mu^{(i)}_t$,

$$\nu^{(i)}_t = \begin{cases} \dot{\gamma}_i(t)\delta_{\gamma_i(t)}, & \text{if } \dot{\gamma}_i(t) \text{ exists,} \\ 0, & \text{otherwise,} \end{cases}$$

and $\nu_t = \displaystyle\sum_{i=1}^N \lambda_i \nu^{(i)}_t$. Then $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$.

Proof. By linearity, clearly we have that

$$\partial_t \mu_t + \operatorname{div} \nu_t = 0$$

is satisfied in the sense of distributions, moreover $\mu_t(B) = 0$ implies $\nu_t(B) = 0$ for every Borel set $B \subseteq \mathbb{R}^d$, thus $|\nu_t| \ll \mu_t$. 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 $v_t \in L^1(\mathbb{R}^d; \mathbb{R}^d)$ satisfying $v_t(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$. Set

$$\tau = \{t \in [0, T] : \dot{\gamma}_i(t) \text{ exists for all } i = 1, \dots, N \text{ and } \dot{\gamma}_i(t) \in F(\gamma_i(t))\},$$

and notice that τ has full measure in $[0, T]$.

Fix $t \in \tau$, $x \in \operatorname{supp} \mu_t$. By definition of μ_t, we have that there exists $I \subseteq \{1, \dots, N\}$ such that $\mu^{(i)}_t = \delta_x$ if and only if $i \in I$. So it is possible to find $\delta > 0$ such that for all $0 < \rho < \delta$ we have

$$\mu_t(B(x, \rho)) = \sum_{j\in I} \lambda_j, \qquad \nu_t(B(x, \rho)) = \sum_{i\in I} \lambda_i \int_{B(x,\rho)} \frac{\nu^{(i)}_t}{\mu^{(i)}_t}(y)d\mu^{(i)}_t(y) = \sum_{i\in I} \lambda_i \frac{\nu^{(i)}_t}{\mu^{(i)}_t}(x).$$

Thus for every $t \in \tau$ and $x \in \operatorname{supp} \mu_t$ we have

$$v_t(x) := \lim_{\rho\to 0^+} \frac{\nu_t(B(x, \rho))}{\mu_t(B(x, \rho))} = \sum_{i\in I} \frac{\lambda_i}{\sum_{j\in I} \lambda_j} \frac{\nu^{(i)}_t}{\mu^{(i)}_t}(x),$$

i.e., a convex combination of $\dot\gamma_i(t) = \dfrac{\nu_t^{(i)}}{\mu_t^{(i)}}(x) \in F(x)$ for μ_t-a.e. $x \in \mathbb{R}^d$. Thus $\dfrac{\nu_t}{\mu_t}(x) = v_t(x) \in F(x)$, and so $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ is an admissible trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. □

Corollary 3.22. *Assume (F_0), (F_1). Let $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1, and assume 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 $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ with $p > 1$. Then $\tilde T_p^\Phi(\mu_0) = \tilde T^\Phi(\mu_0) = \|T(\cdot)\|_{L^\infty_{\mu_0}}$.*

Proof. Since $\tilde S^\Phi$ admits classical counterpart S, closed and nonempty, we have that $\Phi = \{d_S(\cdot)\}$. 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 3.10, we have only to prove that $\tilde T_p(\mu_0) \le T := \|T(\cdot)\|_{L^\infty_{\mu_0}}$. Assume that $T < +\infty$, otherwise there is nothing to prove. For μ_0-a.e. point $x \in \mathbb{R}^d$ we have $T(x) \le T$, 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 $T(x) \le t \le T$, thus in particular $\gamma_x(T) \in S$. Fix $\varepsilon > 0$, then there exists $N = N_\varepsilon \in \mathbb{N}\setminus\{0\}$, and $\{(x_i,\lambda_i) : i = 1,\dots,N_\varepsilon\} \subseteq \operatorname{supp}\mu_0 \times [0,1]$ such that:

(1) $\displaystyle\sum_{i=1}^{N_\varepsilon} \lambda_i = 1$;

(2) $W_p\left(\mu_0, \displaystyle\sum_{i=1}^{N_\varepsilon}\lambda_i\delta_{x_i}\right) < \varepsilon$;

(3) there exist classical admissible trajectories $\{\gamma_i : [0,T] \to \mathbb{R}^d : i = 1,\dots,N_\varepsilon\}$ satisfying $\gamma_i(0) = x_i$ and $\gamma_i(T) \in S$ for all $i = 1,\dots,N_\varepsilon$.

It is possible to find an admissible trajectory $\boldsymbol{\mu}^{(\varepsilon)} = \left\{\mu_t^{(\varepsilon)}\right\}_{t\in[0,T]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ 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

$$\mu_t^{(\varepsilon)} = \sum_{i=1}^{N_\varepsilon}\lambda_i\delta_{\gamma_i(t)}, \qquad \nu_t^{(\varepsilon)} = \begin{cases} \displaystyle\sum_{i=1}^{N_\varepsilon}\lambda_i\dot\gamma_i(t)\delta_{\gamma_i(t)}, & \text{if } \dot\gamma_i(t) \text{ exists for all } i = 1,\dots,N_\varepsilon,\\[2ex] 0, & \text{otherwise,}\end{cases}$$

and then apply Lemma 3.21.

Since $\mu_0^{(\varepsilon)}$ converges in W_p to μ_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 $\mathrm{m}_p(\mu_0) + 1$. In particular, by taking a sequence $\varepsilon_k \to 0^+$, and the corresponding admissible trajectories $\boldsymbol{\mu}^{(\varepsilon_k)}$ driven by $\boldsymbol{\nu}^{(\varepsilon_k)}$, we can extract by Theorem 3.16 a subsequence converging to an admissible trajectory $\bar{\boldsymbol{\mu}}$ driven by $\bar{\boldsymbol{\nu}}$ 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 $\tilde T_p(\mu_0) \le T$. □

Corollary 3.23 (Second comparison result)**.** *Assume (F_0), (F_1). Let $\Phi \subseteq C^0(\mathbb{R}^d;\mathbb{R})$ satisfying (T_E) in Definition 3.1, and assume that the generalized target $\tilde S^\Phi$ admits a* *classical counterpart S. Then, for every $x_0 \in \mathbb{R}^d$ we have $\tilde{T}^\Phi(\delta_{x_0}) = \tilde{T}^\Phi_p(\delta_{x_0}) = T(x_0)$ 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 3.21 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 $\gamma(T(x_0)) \in S$. We obtain an admissible trajectory steering δ_{x_0} to $\tilde{S}_p$ for all $p \geq 1$ in time $T(x_0)$, thus $\tilde{T}_p(\delta_{x_0}) \leq T(x_0)$. By Proposition 3.10, since $\|T(\cdot)\|_{L^\infty_{\delta_{x_0}}} = T(x_0)$, equality holds. □

Remark 3.24. This means that if we have a precise knowledge of the initial state, we recover exactly the classical objects in finite-dimension.

Theorem 3.25 (Dynamic programming principle)**.** *Let $0 \leq s \leq \tau$, let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau]}$ be an admissible curve for Σ_F. Then we have*

$$\tilde{T}^\Phi(\mu_0) \leq s + \tilde{T}^\Phi(\mu_s).$$

Moreover, if $\tilde{T}^\Phi(\mu_0) < +\infty$, equality holds for all $s \in [0, \tilde{T}^\Phi(\mu_0)]$ if and only if $\boldsymbol{\mu}$ is optimal for $\mu_0 = \mu_{|t=0}$. The same result holds for $\tilde{T}^\Phi_p$ in place of $\tilde{T}^\Phi$, $p \geq 1$.

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

$$\tilde{v}^\varepsilon_t(x) := \begin{cases} \dfrac{\nu_t}{\mu_t}(x), & \text{for } 0 \leq t \leq s, \\ \\ \dfrac{\nu^\varepsilon_{t-s}}{\mu^\varepsilon_{t-s}}(x), & \text{for } s < t \leq \tilde{T}^\Phi(\mu_s) + s + \varepsilon. \end{cases}$$

$$\tilde{\mu}^\varepsilon_t := \begin{cases} \mu_t, & \text{for } 0 \leq t \leq s, \\ \\ \mu^\varepsilon_{t-s}, & \text{for } s < t \leq \tilde{T}^\Phi(\mu_s) + s + \varepsilon. \end{cases}$$

It is clear that $\tilde{\mu}^\varepsilon_{|t=0} = \mu_0$, that $\tilde{\mu}^\varepsilon_{|t=\tilde{T}^\Phi(\mu_s)+s+\varepsilon} \in \tilde{S}^\Phi$, and that $\tilde{v}^\varepsilon_t(x) \in F(x)$ for $\tilde{\mu}^\varepsilon_t$–a.e. $x \in \mathbb{R}^d$ and a.e. $t \in [0, \tilde{T}^\Phi(\mu_s) + \varepsilon]$. Moreover, $t \mapsto \tilde{\mu}^\varepsilon_t$ is narrowly continuous. Since the gluing Lemma 4.4 in [21] ensures that $\tilde{\boldsymbol{\mu}}^{\boldsymbol{\varepsilon}} := \{\tilde{\mu}^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+s+\varepsilon]}$ is a solution of the continuity equation driven by $\tilde{\boldsymbol{\nu}}^{\boldsymbol{\varepsilon}} = \{\tilde{\nu}^\varepsilon_t = \tilde{v}^\varepsilon_t \tilde{\mu}^\varepsilon_t\}_{t\in[0,\tilde{T}^\Phi(\mu_s)+s+\varepsilon]}$, thus an admissible trajectory, we have that

$$\tilde{T}^\Phi(\mu_0) \leq J_F(\tilde{\boldsymbol{\mu}}^{\boldsymbol{\varepsilon}}, \tilde{\boldsymbol{\nu}}^{\boldsymbol{\varepsilon}}) = \tilde{T}^\Phi(\mu_s) + s + \varepsilon\,.$$

By arbitrariness of $\varepsilon > 0$, we conclude that $\tilde{T}^\Phi(\mu_0) \leq s + \tilde{T}^\Phi(\mu_s)$.

Assume now that $\tilde{T}^\Phi(\mu_0) < +\infty$ and equality holds for all $s \in [0, \tilde{T}^\Phi(\mu_0)]$. Then, in particular, when $s = \tilde{T}^\Phi(\mu_0)$ we get

$$\tilde{T}^\Phi(\mu_0) = \tilde{T}^\Phi(\mu_0) + \tilde{T}^\Phi(\mu_{\tilde{T}^\Phi(\mu_0)}) \quad \Rightarrow \quad \tilde{T}^\Phi(\mu_{\tilde{T}^\Phi(\mu_0)}) = 0\,.$$

In turn, this implies $\mu_{\tilde{T}^\Phi(\mu_0)} = \mu_{s+\tilde{T}^\Phi(\mu_s)} \in \tilde{S}^\Phi$, and so $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)]}$ joins μ_0 with the generalized target in the minimum time $\tilde{T}^\Phi(\mu_0)$, thus $\boldsymbol{\mu}$ is optimal for μ_0.

Finally, assume that $\boldsymbol{\mu}$, driven by $\boldsymbol{\nu}$, is optimal for μ_0 and $\tilde{T}^\Phi(\mu_0) < +\infty$. To have equality $\tilde{T}^\Phi(\mu_0) = s + \tilde{T}^\Phi(\mu_s)$, it is enough to show that $\tilde{T}^\Phi(\mu_0) \geq s + \tilde{T}^\Phi(\mu_s)$. If we define $\nu'_t := \nu_{t+s}$, we have that $\boldsymbol{\mu'} = \{\mu'_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]} := \{\mu_{t+s}\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]}$ is an admissible trajectory driven by $\boldsymbol{\nu'} = \{\nu'_t\}_{t\in[0,\tilde{T}^\Phi(\mu_0)-s]}$ and starting by μ_s. This implies that

$$\tilde{T}^\Phi(\mu_0) = s + (\tilde{T}^\Phi(\mu_0) - s) = s + J_F(\boldsymbol{\mu'}, \boldsymbol{\nu'}) \geq s + \tilde{T}^\Phi(\mu_s).$$

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.

In other words, we want to prove a generalization of the so called *Petrov's condition* that gives, in the classical case, an attainability property for the control system, i.e. a sufficient condition for continuity of the minimum time function at the boundary of the target.

Theorem 3.26 (Attainability in the smooth case)**.** *Assume* (F_0), (F_1). *Let* $\Phi \subseteq C^1_b(\mathbb{R}^d;\mathbb{R})\cap\mathrm{Lip}(\mathbb{R}^d;\mathbb{R})$ *satisfying* (T_E) *in Definition 3.1 and let* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$, $p \geq 1$. *Assume that:*

(1) *for all* $\phi \in \Phi$ *there exists a* $\mathscr{L}^1$*-integrable map* $k^\phi :]0, +\infty[\to]0, +\infty[$*;*
(2) *there exists* $T \in [0, +\infty[$ *such that* $T \geq \sup_{\phi\in\Phi} \inf\left\{t \geq 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \leq \int_0^t k^\phi(s)\, ds\right\}$*;*
(3) *there exist a Borel vector field* $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ *and an admissible trajectory* $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ *driven by* $\boldsymbol{\nu} = \{\nu_t := v_t\mu_t\}_{t\in[0,T]}$, *and satisfying* $\mu_{|t=0} = \mu_0$,

such that the following condition holds:

(C_c) *for all* $\phi \in \Phi$ *we have* $\displaystyle\int_{\mathbb{R}^d} \langle\nabla\phi(x), v_t(x)\rangle\, d\mu_t(x) \leq -k^\phi(t)$ *for a.e.* $t \in]0, T]$.

Then we have

$$\tilde{T}^\Phi_p(\mu_0) \leq \sup_{\phi\in\Phi} \inf\left\{t \geq 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \leq \int_0^t k^\phi(s)\, ds\right\}.$$

Proof. We notice that by Lemma 3.6, we have $\boldsymbol{\mu} \subseteq \mathscr{P}_p(\mathbb{R}^d)$.

Given $\phi \in \Phi$, we set $L^\phi_t := \displaystyle\int_{\mathbb{R}^d} \phi(x)\, d\mu_t(x)$. Take $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ and notice that if $T = 0$ we have

$$\sup_{\phi\in\Phi} \inf\left\{t \geq 0 \,:\, \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \leq \int_0^t k^\phi(s)\, ds\right\} = 0,$$

so $\mu_0 \in \tilde{S}^\Phi_p$ and $\tilde{T}^\Phi_p(\mu_0) = 0$. We assume then $T > 0$.

From the continuity equation we have that in the distributional sense it holds (see Remark 8.1.1 in [3], allowing to use the functions of Φ as test functions)

$$\dot{L}^\phi_t = \frac{d}{dt}\int_{\mathbb{R}^d} \phi(x) d\mu_t(x) = \int_{\mathbb{R}^d} \langle\nabla\phi(x), v_t(x)\rangle\, d\mu_t(x) \leq -k^\phi(t).$$

Then $L^\phi_t \leq L^\phi_0 - \int_0^t k^\phi(s)\, ds$ for $0 < t \leq T$. Thus if we take $t \in]0, T]$ s.t. we have $\int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \leq \int_0^t k^\phi(s)\, ds$ for all $\phi \in \Phi$, then we have that $L^\phi_t \leq 0$ for all $\phi \in \Phi$, hence $\mu_t \in \tilde{S}^\Phi_p$ for all such t, which ends the proof. □

Remark 3.27. In particular, if in the condition (C_c) above we can choose $k^\phi(t) \equiv k^\phi$ for a.e. $t > 0$, for a constant $k^\phi > 0$, then we get $\tilde{T}_p^\Phi(\mu_0) \le \sup_{\phi\in\Phi} \left\{ \frac{1}{k^\phi} \int_{\mathbb{R}^d} \phi(x)\, d\mu_0(x) \right\}$.

For other results about the regularity of the minimum time function, we refer the reader to the forthcoming [15].

4. 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 $\mathscr{P}_2(\mathbb{R}^d)$ 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], [14],[23],[22]), being modeled on this particular problem.

Throughout this section we will mainly use the alternative definition of admissible curve and the notation provided by Definition 3.5.

Definition 4.1 (Averaged speed set)**.** Assume (F_0) and (F_1), $T > 0$. For any $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$, $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$, we set

$$\mathscr{V}(\boldsymbol{\eta}) := \Bigg\{ w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T) : \exists \{t_i\}_{i\in\mathbb{N}} \subseteq]0,T[, \text{ with } t_i \to 0^+ \text{ and } \frac{e_{t_i} - e_0}{t_i} \rightharpoonup w_{\boldsymbol{\eta}} \text{ weakly in } L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d) \Bigg\}.$$

We notice that, according to the boundedness result of Lemma 3.6 (iii), for any sequence $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,T[$ with $t_i \to 0^+$, there exists a subsequence $\tau = \{t_{i_k}\}_{k\in\mathbb{N}}$ and $w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$ such that $\dfrac{e_{t_{i_k}} - e_0}{t_{i_k}}$ weakly converges to an element of $L^2_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$, thus $\mathscr{V}(\boldsymbol{\eta}) \neq \emptyset$.

Lemma 4.2 (Properties of the averaged speed set)**.** *Assume (F_0) and (F_1), $T > 0$. For any $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$, $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ 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 $\{\eta_x\}_{x\in\mathbb{R}^d}$ 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 $L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$.

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

$$\begin{aligned}\langle v, \varphi_t(x,\gamma)\rangle &= \langle v, \frac{\gamma(t)-\gamma(0)}{t}\rangle = \frac{1}{t}\int_0^t \langle v, \dot\gamma(s)\rangle\, ds \\ &\le \frac{1}{t}\int_0^t \sigma_{F(\gamma(s))}(v)\, ds \le \sigma_{F(\gamma(0))+\varepsilon B(0,1)}(v),\end{aligned}$$

where $\varphi_t(x,\gamma) = \dfrac{e_t(x,\gamma) - e_0(x,\gamma)}{t}$.

Thus

$$\overline{\text{co}}\{\varphi_t(x,\gamma) : 0 < t < t^*_{\varepsilon,\gamma}\} \subseteq F(\gamma(0)) + \varepsilon\overline{B(0,1)}$$

Given $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, let $\{t_i\}_{i\in\mathbb{N}} \subseteq]0,1]$ be a sequence such that $t_i \to 0^+$ and $\varphi_{t_i} \rightharpoonup w_{\boldsymbol{\eta}}$ weakly in $L^2_{\boldsymbol{\eta}}$. In particular, by Mazur's Lemma, there is a sequence in $\mathrm{co}\{\varphi_{t_i} : i \in \mathbb{N}\}$ strongly convergent to $w_{\boldsymbol{\eta}}$. In particular, for (x,γ)-a.e. point of $\mathbb{R}^d \times \Gamma_T$ we have pointwise convergence, i.e.

$$w_{\boldsymbol{\eta}}(x,\gamma) \in \overline{\mathrm{co}}\{\varphi_{t_i}(x,\gamma) : i \in \mathbb{N}\}.$$

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

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

By letting $\varepsilon \to 0^+$ 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 $\{\eta_x\}_{x\in\mathbb{R}^d}$ satisfying (recall that $e_0\sharp\boldsymbol{\eta} = \mu_0$)

$$\iint_{\mathbb{R}^d\times\Gamma_T} f(x,\gamma) w_{\boldsymbol{\eta}}(x,\gamma)\, d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d} \left(\int_{\Gamma_T^x} \langle f(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\eta_x(\gamma) \right) d\mu_0(x),$$

for all Borel map $f : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$. Moreover the family $\{\eta_x\}_{x\in\mathbb{R}^d}$ is uniquely determined for μ_0-a.e. $x \in \mathbb{R}^d$ (see e.g. Theorem 5.3.1 in [3]).

For any $\psi \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, clearly we have $\psi\circ e_0 \in L^2_{\boldsymbol{\eta}}(\mathbb{R}^d\times\Gamma_T;\mathbb{R}^d)$, since $e_0\sharp\boldsymbol{\eta} = \mu_0$. Recalling that $w_{\boldsymbol{\eta}} \in L^2_{\boldsymbol{\eta}}$, we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \langle \psi(x), \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x(\gamma)\rangle \, d\mu_0(x) &= \int_{\mathbb{R}^d}\int_{\Gamma_T^x} \langle \psi(x), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\eta_x(\gamma)\, d\mu_0(x) \\ &= \int_{\mathbb{R}^d}\int_{\Gamma_T^x} \langle \psi\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\eta_x(\gamma)\, d\mu_0(x) \\ &= \iint_{\mathbb{R}^d\times\Gamma_T} \langle \psi\circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) < +\infty. \end{aligned}$$

By the arbitrariness of $\psi \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, we obtain that the map

$$x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x(\gamma),$$

belongs to $L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$, moreover for μ_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 4.3. We can interpret each $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ as a sort of averaged vector field of initial velocity in the sense of measure (we recall that in general an admissible trajectory γ 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 an *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 4.4 (Regular driving vector fields)**.** *Assume (F_0), (F_1) and let $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. Let $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ be an absolutely continuous solution of*

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0,\ t \in]0,T[\\ \\ \mu_{|t=0} = \mu_0, \end{cases}$$

where $v \in C^0([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ satisfies $v_0(x) \in F(x)$ for all $x \in \mathbb{R}^d$. Then if $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ satisfies $\mu_t = e_t \sharp \boldsymbol{\eta}$ for all $t \in [0,T]$, we have that

$$\lim_{t\to 0} \left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|_{L^2_{\boldsymbol{\eta}}} = 0,$$

and so $\mathscr{V}(\boldsymbol{\eta}) = \{v_0 \circ e_0\}$, thus we have

$$\left\{ x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\, d\eta_x :\ w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta}) \right\} = \{v_0(\cdot)\}.$$

Proof. We have

$$\left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|^2_{L^2_{\boldsymbol{\eta}}} = \iint_{\mathbb{R}^d\times\Gamma_T} \left| \frac{\gamma(t)-\gamma(0)}{t} - v_0(\gamma(0)) \right|^2 d\boldsymbol{\eta}(x,\gamma),$$

For $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$, by continuity of v we have $\gamma \in C^1$ and $\dot\gamma(t) = v_t(\gamma(t))$, hence for t small enough we get

$$\begin{aligned} \left| \frac{\gamma(t)-\gamma(0)}{t} - v_0(\gamma(0)) \right| &\le \frac{1}{t}\int_0^t |\dot\gamma(s)|\, ds + |v_0(\gamma(0))| = \frac{1}{t}\int_0^t |v_s(\gamma(s))|\, ds + |v_0(\gamma(0))| \\ &\le 2|v_0(\gamma(0))| + 1 \in L^2_{\boldsymbol{\eta}}, \end{aligned}$$

indeed by (F_1) we have

$$\begin{aligned} \iint_{\mathbb{R}^d\times\Gamma_T} |v_0(\gamma(0))|^2\, d\boldsymbol{\eta}(x,\gamma) = \int_{\mathbb{R}^d} |v_0(x)|^2\, d\mu_0(x) &\le C^2 \int_{\mathbb{R}^d} (|x|+1)^2\, d\mu_0(x) \\ &\le 2C^2\left(\mathrm{m}_2(\mu_0)+1\right), \end{aligned}$$

with $C > 0$ as in (F_1). Thus, for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$,

$$\lim_{t\to 0^+} \left| \frac{\gamma(t)-\gamma(0)}{t} - v_0(\gamma(0)) \right| = 0.$$

Thus applying Lebesgue's Dominated Convergence Theorem we obtain

$$\lim_{t\to 0} \left\| \frac{e_t - e_0}{t} - v_0 \circ e_0 \right\|^2_{L^2_{\boldsymbol{\eta}}} = 0,$$

hence $w_{\boldsymbol{\eta}} = v_0 \circ e_0$. The last assertion now follows. ∎

We have already proved that the set

$$\left\{x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\,d\eta_x : \boldsymbol{\eta} \in \mathscr{T}_F(\mu_0),\, w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})\right\}$$

is contained in the set of all $L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$-selections of $F(\cdot)$. The next density result shows that, indeed, equality holds: since allows to approximate every $L^2_{\mu_0}$-selections by C^0-selections, and then use Lemma 4.4.

Lemma 4.5 (Approximation)**.** *Let $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. Assume (F_0) and (F_1). Then given any $v \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)$ satisfying $v(x) \in F(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$, there exists a sequence of continuous maps $\{g_n\}_{n\in\mathbb{N}} \subseteq C^0(\mathbb{R}^d;\mathbb{R}^d)$ such that*

(1) $\lim\limits_{n\to\infty} \|g_n - v\|_{L^2_{\mu_0}} = 0$;
(2) $g_n(x) \in F(x)$ *for all* $x \in \mathbb{R}^d$.

In particular, we have

$$\{v \in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d) : v(x) \in F(x) \text{ for } \mu_0\text{-a.e. } x \in \mathbb{R}^d\} =$$
$$= \left\{x \mapsto \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma)\,d\eta_x : \boldsymbol{\eta} \in \mathscr{T}_F(\mu_0),\, w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})\right\}.$$

Proof. By Lusin's Theorem (see e.g. Theorem 1.45 in [6]), we can construct a sequence of compact sets $\{K_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ and of continuous maps $\{v_n\}_{n\in\mathbb{N}} \subseteq C^0_c(\mathbb{R}^d;\mathbb{R}^d)$ such that $v_n = v$ on K_n and $\mu_0(\mathbb{R}^d \setminus K_n) \le 1/n$. For all $n \in \mathbb{N}$ define the set valued maps

$$G_n(x) := \begin{cases} F(x), & \text{for } x \in \mathbb{R}^d \setminus K_n, \\ \\ \{v_n(x)\}, & \text{for } x \in K_n. \end{cases}$$

We prove that $G_n(\cdot)$ is lower semicontinuous. If $x \in \mathbb{R}^d \setminus K_n$, 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 $V = B(v_n(x), \varepsilon)$ 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 $U' = U \cap B(x,\delta) \setminus K_n$, we have that U' is an open neighborhood of x satisfying:

(a) for all $y \in U' \setminus K_n$ we have $F(y) = G_n(y)$ and so $G_n(y) \cap V \neq \emptyset$;
(b) for all $y \in U' \cap K_n$ we have $v_n(y) \in V$, and so $G_n(y) \cap V \neq \emptyset$;

and so given V for all $y \in U'$ 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 $g_n \in C^0(\mathbb{R}^d;\mathbb{R}^d)$ 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

$$\int_{\mathbb{R}^d} |v(x) - g_n(x)|^2\,d\mu_0(x) = \int_{\mathbb{R}^d\setminus K_n} |v(x) - g_n(x)|^2\,d\mu_0(x) \le \int_{\mathbb{R}^d\setminus K_n} 4C^2(|x|+1)^2\,d\mu_0(x)$$
$$\le 8C^2\left(\mathrm{m}_2(\mu_0) + 1\right),$$

with $C > 0$ as in (F_1), hence (1) follows. The last assertion comes from Lemma 4.4 with $v = v_0$. □

We introduce now the following definition of viscosity sub-/superdifferential. For other concepts of viscosity sub-/superdifferential, we refer the reader to [3] and [14].

Definition 4.6 (Sub-/Super-differential in $\mathscr{P}_2(\mathbb{R}^d)$)**.** Let $V : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function. Fix $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\delta > 0$. We say that $p_\mu \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ belongs to the δ-superdifferential $D^+_\delta V(\mu)$ at μ if for all $T > 0$ and $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ such that $t \mapsto e_t\sharp\boldsymbol{\eta}$ is an absolutely continuous curve in $\mathscr{P}_2(\mathbb{R}^d)$ defined in $[0,T]$ with $e_0\sharp\boldsymbol{\eta} = \mu$ we have
(4.1)

$$\limsup_{t\to 0^+} \frac{V(e_t\sharp\boldsymbol{\eta}) - V(e_0\sharp\boldsymbol{\eta}) - \displaystyle\int_{\mathbb{R}^d\times\Gamma_T} \langle p_\mu \circ e_0(x,\gamma), e_t(x,\gamma) - e_0(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma)}{\|e_t - e_0\|_{L^2_{\boldsymbol{\eta}}}} \le \delta.$$

In the same way, $q_\mu \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ belongs to the δ-subdifferential $D^-_\delta V(\mu)$ at μ if $-q_\mu \in D^+_\delta[-V](\mu)$.

Definition 4.7 (Viscosity solutions)**.** Let $V : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function and $\mathscr{H} : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$. We say that V is a

1. *viscosity supersolution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if V is l.s.c. and there exists $C > 0$ depending only on $\mathscr{H}$ such that $\mathscr{H}(\mu, q_\mu) \ge -C\delta$ for all $q_\mu \in D^-_\delta V(\mu)$, $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and for all $\delta > 0$.
2. *viscosity subsolution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if V is u.s.c. and there exists $C > 0$ depending only on $\mathscr{H}$ such that $\mathscr{H}(\mu, p_\mu) \le C\delta$ for all $p_\mu \in D^+_\delta V(\mu)$, $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, and for all $\delta > 0$.
3. *viscosity solution* of $\mathscr{H}(\mu, DV(\mu)) = 0$ if it is both a viscosity subsolution and a viscosity supersolution.

Definition 4.8 (Hamiltonian Function)**.** Given $\mu \in \mathscr{P}(\mathbb{R}^d)$, define

$$\mathscr{D}(\mu) := \left\{ \nu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d) : |\nu| \ll \mu \text{ and } \int_{\mathbb{R}^d} \left(\left|\frac{\nu}{\mu}\right|^2 + I_{F(x)}\left(\frac{\nu}{\mu}(x)\right)\right) d\mu < +\infty \right\}.$$

Since the tangent space $T_\mu\mathscr{P}_2(\mathbb{R}^d)$ to $\mathscr{P}_2(\mathbb{R}^d)$ at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ is $L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$, which coincides with its dual, we can define a map $\mathscr{H}_F : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by setting

$$\begin{aligned} \mathscr{H}_F(\mu,\psi) &:= -\left[1 + \inf_{\nu\in\mathscr{D}(\mu)} \int_{\mathbb{R}^d} \langle \psi(x), \frac{\nu}{\mu}(x)\rangle \, d\mu\right], \\ &= -\left[1 + \inf_{\substack{v\in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d) \\ v(x)\in F(x) \text{ for } \mu\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle \psi(x), v(x)\rangle \, d\mu\right], \end{aligned}$$

where $(\mu,\psi) \in T^*\mathscr{P}_2(\mathbb{R}^d)$, i.e., $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\psi \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$.

If we assume (F_2), or more generally that F possesses a Borel selection uniformly bounded, we have

$$\mathscr{H}_F(\mu,\psi) := -1 + \int_{\mathbb{R}^d} \sigma_{-F(x)}(\psi(x)) \, d\mu,$$

by using a consequence of classical Measurable Selection Lemma (see e.g. Theorem 6.31 p. 119 in [19]).

Theorem 4.9 (Viscosity solution)**.** *Let $\mathcal{A}$ be any open subset of $\mathscr{P}_2(\mathbb{R}^d)$ with uniformly bounded 2−moments. Assume (F_0) and (F_1) and that $\tilde{T}_2^\Phi(\cdot)$ is continuous on $\mathcal{A}$. Then $\tilde{T}_2^\Phi(\cdot)$ is a viscosity solution of $\mathscr{H}_F(\mu, D\tilde{T}_2^\Phi(\mu)) = 0$ on $\mathcal{A}$, with $\mathscr{H}_F$ defined as in Definition 4.8.*

Proof. The proof is splitted in two claims.

Claim 1: $\tilde{T}_2^\Phi(\cdot)$ is a subsolution of $\mathscr{H}_F(\mu, D\tilde{T}_2^\Phi(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 1. Let $\mu_0 \in \mathcal{A}$. Given $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and set $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all t, by the Dynamic Programming Principle (Theorem 3.25) we have $\tilde{T}_2^\Phi(\mu_0) \le \tilde{T}_2^\Phi(\mu_s) + s$ for all $0 < s \le \tilde{T}_2^\Phi(\mu_0)$. Without loss of generality, we can assume $0 < s < 1$. Given any $p_{\mu_0} \in D_\delta^+\tilde{T}_2^\Phi(\mu_0)$, and set

$$A(s, p_{\mu_0}, \boldsymbol{\eta}) := - s - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), e_s(x,\gamma) - e_0(x,\gamma)\rangle \, d\boldsymbol{\eta},$$

$$B(s, p_{\mu_0}, \boldsymbol{\eta}) := \tilde{T}_2^\Phi(\mu_s) - \tilde{T}_2^\Phi(\mu_0) - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), e_s(x,\gamma) - e_0(x,\gamma)\rangle d\boldsymbol{\eta},$$

we have $A(s, p_{\mu_0}, \boldsymbol{\eta}) \le B(s, p_{\mu_0}, \boldsymbol{\eta})$.

We recall that since by definition $p_{\mu_0} \in L^2_{\mu_0}$, we have that $p_{\mu_0} \circ e_0 \in L^2_{\boldsymbol{\eta}}$. Dividing by $s > 0$, we obtain that

$$\limsup_{s\to 0^+} \frac{A(s, p_{\mu_0}, \boldsymbol{\eta})}{s} \ge -1 - \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma),$$

for all $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$.

Recalling the choice of p_{μ_0}, we have

$$\limsup_{s\to 0^+} \frac{B(s, p_{\mu_0}, \boldsymbol{\eta})}{s} = \limsup_{s\to 0^+} \frac{B(s, p_{\mu_0}, \boldsymbol{\eta})}{\|e_s - e_0\|_{L^2_{\boldsymbol{\eta}}}} \cdot \left\| \frac{e_s - e_0}{s} \right\|_{L^2_{\boldsymbol{\eta}}} \le K\delta,$$

where $K > 0$ is a suitable constant coming from Lemma 3.6 and from hypothesis.

We thus obtain for all $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and all $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, that

$$1 + \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) \ge -K\delta.$$

By passing to the infimum on $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$, and recalling Lemma 4.5, we have

$$\begin{aligned}
-K\delta &\le 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \iint_{\mathbb{R}^d\times\Gamma_T} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) \\
&= 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \int_{\mathbb{R}^d} \int_{\Gamma_T^x} \langle p_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\eta_x \, d\mu_0 \\
&= 1 + \inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \int_{\mathbb{R}^d} \langle p_{\mu_0} \circ e_0(x,\gamma), \int_{\Gamma_T^x} w_{\boldsymbol{\eta}}(x,\gamma) \, d\eta_x\rangle \, d\mu_0 \\
&= 1 + \inf_{\substack{v\in L^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e. } x}} \int_{\mathbb{R}^d} \langle p_{\mu_0}, v\rangle \, d\mu_0 = -\mathscr{H}_F(\mu_0, p_{\mu_0}),
\end{aligned}$$

so $\tilde{T}_2^\Phi(\cdot)$ is a subsolution, thus confirming Claim 1. ◇

Claim 2: $\tilde{T}_2^\Phi(\cdot)$ is a supersolution of $\mathscr{H}_F(\mu, D\tilde{T}_2^\Phi(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 2. Let $\mu_0 \in \mathcal{A}$. Given $\boldsymbol{\eta} \in \mathscr{T}_F(\mu_0)$ and defined the admissible trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} = \{e_t\sharp\boldsymbol{\eta}\}_{t\in[0,T]}$, and $q_{\mu_0} \in D_\delta^- \tilde{T}_2^\Phi(\mu_0)$, there is a sequence $\{s_i\}_{i\in\mathbb{N}} \subseteq]0,T[$ and $w_{\boldsymbol{\eta}} \in \mathscr{V}(\boldsymbol{\eta})$ such that $s_i \to 0^+$, $\dfrac{e_{s_i}-e_0}{s_i}$ weakly converges to $w_{\boldsymbol{\eta}}$ in $L^2_{\boldsymbol{\eta}}$, and for all $i \in \mathbb{N}$

$$\iint_{\mathbb{R}^d\times\Gamma_T} \langle q_{\mu_0} \circ e_0(x,\gamma), \frac{e_{s_i}(x,\gamma)-e_0(x,\gamma)}{s_i}\rangle \, d\boldsymbol{\eta}(x,\gamma) \\ \leq 2\delta \left\| \frac{e_{s_i}-e_0}{s_i} \right\|_{L^2_{\boldsymbol{\eta}}} - \frac{\tilde{T}_2^\Phi(\mu_0) - \tilde{T}_2^\Phi(\mu_{s_i})}{s_i}.$$

By taking i sufficiently large we thus obtain

$$\iint_{\mathbb{R}^d\times\Gamma_T} \langle q_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) \leq 3K\delta - \frac{\tilde{T}_2^\Phi(\mu_0) - \tilde{T}_2^\Phi(\mu_{s_i})}{s_i}.$$

By using Lemma 4.5 and arguing as in Claim 1, we have

$$\inf_{\substack{\boldsymbol{\eta}\in\mathscr{T}_F(\mu_0)\\ w_{\boldsymbol{\eta}}\in\mathscr{V}(\boldsymbol{\eta})}} \iint_{\mathbb{R}^d\times\Gamma_T} \langle q_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}}(x,\gamma)\rangle \, d\boldsymbol{\eta}(x,\gamma) = -\mathscr{H}_F(\mu_0, q_{\mu_0}) - 1,$$

and so

$$\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -3K\delta + \frac{\tilde{T}_2^\Phi(\mu_0) - \tilde{T}_2^\Phi(\mu_{s_i})}{s_i} - 1.$$

By the Dynamic Programming Principle, passing to the infimum on all admissible curves and recalling that $\dfrac{\tilde{T}_2^\Phi(\mu_0) - \tilde{T}_2^\Phi(\mu_s)}{s} - 1 \leq 0$ with equality holding if and only if $\boldsymbol{\mu}$ is optimal, we obtain $\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -C'\delta$, which proves that $\tilde{T}_2^\Phi(\cdot)$ 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] Giuseppe Buttazzo, *Semicontinuity, relaxation and integral representation in the calculus of variations*, Pitman Research Notes in Mathematics Series, vol. 207, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989. MR1020296

[12] 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)

[13] 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

[14] Pierre Cardaliaguet and Marc Quincampoix, *Deterministic differential games under probability knowledge of initial condition*, Int. Game Theory Rev. **10** (2008), no. 1, 1–16.

[15] Giulia Cavagnari, *Regularity results for a time-optimal control problem in the space of probability measures*, to appear in Mathematical Control and Related Fields.

[16] Giulia Cavagnari and Antonio Marigonda, *Measure-theoretic Lie Brackets for nonsmooth vector fields*, preprint.

[17] ———, *Time-optimal control problem in the space of probability measures*, Large-scale scientific computing, Lecture Notes in Computer Science, vol. 9374, Springer, Cham, 2015, pp. 109–116, DOI 10.1007/978-3-319-26520-9. MR3480817

[18] Giulia Cavagnari, Antonio Marigonda, and Fabio S. Priuli, *Probabilistic target for control problems in Wasserstein spaces*, preprint.

[19] Francis Clarke, *Functional analysis, calculus of variations and optimal control*, Graduate Texts in Mathematics, vol. 264, Springer, London, 2013. MR3026831

[20] 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)

[21] 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)

[22] 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)

[23] 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

[24] Pierre Louis Lions and Jean-Michel Lasry, *Jeux à champ moyen. I. Le cas stationnaire*, C. R. Acad. Sci. Paris **343** (2006), 619–625.

[25] ———, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Acad. Sci. Paris **343** (2006), 679–684.

[26] 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)

[27] Héctor J. Sussmann, *Subanalytic sets and feedback control*, J. Differential Equations **31** (1979), 31–52.

[28] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR1964483 (2004e:90003)

Giulia Cavagnari: Dipartimento di Matematica,
Università di Trento
Via Sommarive 14, I-38123 Povo (Trento), Italy.
E-mail address: giulia.cavagnari@unitn.it

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

Khai T. Nguyen: Department of Mathematics,
North Carolina State University,
Raleigh, NC 27695, USA.
E-mail address: tnguye13@ncsu.edu

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