ESAIM: COCV 24 (2018) 721–740
https://doi.org/10.1051/cocv/2017060

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

AVERAGED TIME-OPTIMAL CONTROL PROBLEM IN THE SPACE OF POSITIVE BOREL MEASURES [*, **]

Giulia Cavagnari[1,a], Antonio Marigonda[2] and Benedetto Piccoli[1]

Abstract. We introduce a time-optimal control theory in the space $\mathscr{M}^+(\mathbb{R}^d)$ of positive and finite Borel measures. We prove some natural results, such as a dynamic programming principle, the existence of optimal trajectories, regularity results and an HJB equation for the value function in this infinite-dimensional setting. The main tool used is the superposition principle (by Ambrosio–Gigli–Savaré) which allows to represent the trajectory in the space of measures as weighted superposition of classical characteristic curves in $\mathbb{R}^d$.

Mathematics Subject Classification. 34A60, 49J15.

Received February 8, 2017. Revised August 3, 2017. Accepted August 30, 2017.

1. Introduction

The study of control problems is often times closely linked to applications and other disciplines, such as finance, engineering, biology, logistic among others. Thinking of real world problems, the modeling of uncertainty features naturally arises as a strong need. In many cases uncertainty can affect the state of the system as well as the dynamics.

Control problems with uncertainty have been analyzed by various researchers via different techniques and approaches. For instance, in [19] the authors modeled "plant uncertainty" in a deterministic way, [28] is based on the analysis of randomized algorithms and robustness, [5, 30] study stochastic control problems, while [23] focuses on applying the stochastic control approach to finance, and [6] to quantum control.

Referring to stochastic approaches, in [22] the uncertainty is considered both in the state variable and in the dynamics. The state is thus represented by a random variable or, alternatively, by a probability distribution, while the equation modeling the dynamics involves Brownian motion and the solutions are considered in the sense of Ito or Stratonovich integral.

Keywords and phrases. Time-optimal control, dynamic programming, optimal transport, differential inclusions, multi-agent systems.

[*] *The first two authors have been supported by INdAM - GNAMPA Project 2016:* Stochastic Partial Differential Equations and Stochastic Optimal Transport with Applications to Mathematical Finance.

[**] *The authors acknowledge the partial support of the NSF Project* Kinetic description of emerging challenges in multiscale problems of natural sciences*, DMS Grant # 1107444 and the endowment fund of the Joseph and Loretta Lopez Chair.*

[1] Department of Mathematical Sciences, Rutgers University – Camden, 311 N. 5th Street Camden, NJ 08102, USA.
[a] Corresponding author: giulia.cavagnari@rutgers.edu;
piccoli@camden.rutgers.edu
[2] Department of Computer Science, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy.
mrgntn80@univr.it

Article published by EDP Sciences

The stochastic approach proved to be successful, nevertheless, many problems can be described in a natural way with a different approach. The idea is to consider a deterministic evolution of probability measures as in transport theory [29]. This approach is applicable potentially to all control problems involving uncertainty, also when uncertainty may fail to be well represented by the stochastic approach, for instance to avoid unbounded perturbations. On one hand it can be used to model situations in which the knowledge of the initial state comes with some noise, or, on the other hand, to model the evolution of the statistical distribution of a mass of particles/agents as in the so-called *multi-agent systems* [26]. In such problems many factors must be taken into account to accuratly describe the evolution of the system, both in the case of interacting particles and in the case in which we assume no-interaction among them. There is a rich literature on the subject, for example [27] used the concept of discrete-time evolving measures, in [7] concentration and congestion effects are studied and [17] is a recent survey analyzing the relations between *individual* and *collective* behaviours in crowd dynamics in a unified description based on measure theory. In crowd dynamics, a critical issue is the efficient regulation of a crowd exiting a structured environment also called the *evacuation problem*. More precisely, the latter asks for a mass of agents to be driven outside a given area while optimizing the time needed for the exit of last agent. A natural way to describe these kind of situations is to remove pedestrians from the system once they have reached the target, hence considering a continuity equation with sink. We recall that in a mass-preserving situation, the natural metric to consider is the usual Wasserstein distance between probability measures. In [24, 25] a generalized notion of Wasserstein distance between positive finite Borel measures with possibly different total mass is given to deal with non-homogeneous continuity equations with a source/sink term.

A first attempt for the study of a time-optimal control problem in the space of probability measures endowed with the topology induced by the Wasserstein distance is developed in [11–14], where a mass-preserving situation is addressed. Here, the initial state is described by a Borel probability measure $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ and the trajectories are time-dependent probability measures on $\mathbb{R}^d$, denoted with $\boldsymbol{\mu} := \{\mu_t\}_{t\in[0,T]}$, where μ_t is a solution of a *controlled homogeneous continuity equation*

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \quad 0 < t \le T \\ \mu_{|t=0} = \mu_0, \end{cases} \tag{1.1}$$

and $v_t(\cdot)$ is the control parameter to be chosen among the $\mathrm{L}^1_{\mu_t}$-selections of a given set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ for μ_t-a.e. x and $\mathscr{L}^1$-a.e. t. Here, the multifunction F is governing the underlying finite-dimensional dynamics given in terms of a differential inclusion

$$\begin{cases} \dot\gamma(t) \in F(\gamma(t)), \quad \text{for} \quad \mathscr{L}^1\text{−a.e.} \quad t \in (0,T) \\ \gamma(0) = x \in \mathbb{R}^d. \end{cases} \tag{1.2}$$

A connection between the Caratheodory solutions of the finite-dimension *characteristic system* (1.2) and the distributional solutions of (1.1) is possible even in the case in which the vector field v_t driving both systems is not locally Lipschitz continuous in the space variable uniformly w.r.t. t, hence we do not have uniqueness of the solutions of the characteristic system nor of the solutions of (1.1). For this powerful result, called *Superposition Principle*, we refer the reader to Theorem 8.2.1 in [1]. This represents also the main tool used in the present paper.

The model studied in this paper shares the same underlying idea of the one studied in [13], but here we consider a case with mass loss as follows. In an *optimal logistic/equipment* interpretation of this problem, the initial state $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ represents as before the initial statistical distribution of the agents, the function $f_0 : \mathbb{R}^d \to [0, +\infty]$, called *clock-function*, is the initial amount of goods (ex. time) that has to be assigned at the beginning to each agent in the support of μ_0 in order to reach the given target set $S \subseteq \mathbb{R}^d$ following an admissible mass-preserving trajectory $\boldsymbol{\mu} \subseteq \mathscr{P}(\mathbb{R}^d)$ starting from μ_0, coupled with f_0. The definition of $\boldsymbol{\mu}$ is the same as in [11–14], but here we associate to it a density, $f_t := f_0 - t$, representing the time-linear consumption of the provided supplies, hence here the couple $(f_0, \boldsymbol{\mu})$ is said to be admissible if furthermore f_0 keeps nonnegative the associated density untill the agents have reached the target. Note that we ask a strong invariance property on S in order to remove the agents once they have achieved their own goal. Here, differently from [13], we are interested in minimizing an averaged cost function $\int_{\mathbb{R}^d} f_0(x)\,\mathrm{d}\mu_0(x)$, among all the admissible couples $(f_0, \boldsymbol{\mu})$. This problems leads naturally to the definition of a new concept of trajectory in the space of positive finite Borel measures, called *clock-trajectory*, which is modeled on $\boldsymbol{\mu}$ and f_t and hence it looses its mass linearly in time.

We stress the fact that we consider the case in which f_0 depends only on the initial position of each agent and we have non-renewable resources and non-interacting particles. An applicative example for this situation in a fluid depuration problem has been provided in Example 1 in [15], where other results on the same problem has been investigated as the construction of an optimal clock-trajectory by approximation techniques.

In this paper we show that the optimal clock-function is given by the classical minimum time function $T(\cdot)$, that is the minimum amount of time/supplies that has to be assigned at the beginning to each agent in order to reach the target, even in the case in which $T(\cdot)$ is not continuous but it satisfies only the natural integrability property w.r.t. μ_0.

The paper is organized as follows: in Section 2 we recall some preliminary definitions and fix the notation; in Section 3 we give a formal description of the problem and state the existence of an optimal clock-trajectory for the system, showing that the optimal clock-function turns out to be the classical minimum time function. This justifies the name of *averaged time-optimal problem*, and implies a Dynamic Programming Principle and some regularity results on the value function; finally Section 4 provides an Hamilton-Jacobi-Bellman equation, solved in a suitable viscosity sense by the value function, in analogy with the problem discussed in [13, 14]. In Appendix A we state and prove some technical results used in the paper.

2. Preliminaries and notation

For preliminaries on measure theory, we refer to Chapter 5 in [1].

Given a separable metric space X, we denote by $\mathscr{P}(X)$ the set of Borel probability measures on X endowed with narrow convergence, by $\mathscr{M}^+(X)$ the set of positive and finite Radon measures on X and with $\mathscr{M}(X, \mathbb{R}^d)$ 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 $(\mathrm{C}^0_b(X))'$, and narrow convergence is induced by the weak*-topology on $(\mathrm{C}^0_b(X))'$.

If $\nu \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$, we denote with $|\nu|$ its *total variation*, while for a pair of measures σ, μ defined on $\mathbb{R}^d$, we denote with $\sigma \ll \mu$ the relation of *absolute continuity* of σ w.r.t. μ.

If X, Y are separable metric spaces, $\mu \in \mathscr{M}(X)$, and $r : X \to Y$ is a Borel (or, more generally, μ-measurable) map, we denote by $r\sharp\mu \in \mathscr{M}(Y)$ the push-forward of μ through r, defined by $r\sharp\mu(B) := \mu(r^{-1}(B))$, for all Borel sets $B \subseteq Y$. Equivalently, it is defined by

$$\int_X f(r(x))\,\mathrm{d}\mu(x) = \int_Y f(y)\,\mathrm{d}r\sharp\mu(y),$$

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

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

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

Given $p \geq 1$, we define the set $\mathscr{P}_p(\mathbb{R}^d) := \{\mu \in \mathscr{P}(\mathbb{R}^d) : \mathrm{m}_p(\mu) < +\infty\}$.

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

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

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

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

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*

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

Proof. See Theorem 6.1.1 and Proposition 7.1.5 in [1]. □

In the following we mention some concepts regarding the classical optimal control problem with dynamics in the form of a differential inclusion in $\mathbb{R}^d$.

Definition 2.3 (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 Lipschitz continuous with respect to the Hausdorff metric, *i.e.*, there exists $L > 0$, $L \in \mathbb{R}$, such that for all $x,y \in \mathbb{R}^d$ it holds

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

An *admissible trajectory of the differential inclusion*

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

is, by definition, an absolutely continuous function $\gamma : [0,T] \to \mathbb{R}^d$ satisfying (2.3) for a.e. $0 < t \leq T$.

We recall that given $x \in \mathbb{R}^d$, the classical *minimum time function* $T : \mathbb{R}^d \to [0,+\infty]$ evaluated at x is defined to be the minimum time needed to steer such a point to the target S along admissible trajectories starting from x at $t = 0$. Let $S \subseteq \mathbb{R}^d$ be a closed and nonempty target set, an admissible trajectory $\bar{\gamma}$ is called *optimal for* $x \in \mathbb{R}^d$ if $\bar{\gamma}(0) = x$, $\bar{\gamma}(T(x)) \in S$.

We say that a target set $S \subseteq \mathbb{R}^d$ is *strongly invariant for* F if for any admissible trajectory for F such that there exists $t > 0$ with $\gamma(t) \in S$, we have also $\gamma(s) \in S$ for all $s \geq t$.

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

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

We endow all the above spaces with the usual sup-norm, recalling that Γ_T is a complete separable metric space for every $0 < T < +\infty$. The *evaluation operator* will be the map $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$ defined by $e_t(x,\gamma) = \gamma(t)$ for all $0 \leq t \leq T$.

Let X be a set, $A \subseteq X$. 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 $x \notin A$.

3. A time-optimal control problem with mass loss

We are going to define now the concepts of trajectories on which our work is modeled. Let us first recall the definition of *admissible mass-preserving trajectory* defined in [11–14].

Definition 3.1. Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued map, $\bar{\mu} \in \mathscr{P}(\mathbb{R}^d)$.

(1) Let $T > 0$. We say that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an *admissible mass-preserving trajectory defined on* $[0,T]$ *and starting from* $\bar{\mu}$ if there exists $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $|\nu_t| \ll \mu_t$ for a.e. $t \in [0,T]$, $\mu_0 = \bar{\mu}$, $\partial_t\mu_t + \operatorname{div}\nu_t = 0$ in the sense of distributions and $v_t(x) := \dfrac{\nu_t}{\mu_t}(x) \in F(x)$ for a.e. $t \in [0,T]$ and μ_t-a.e. $x \in \mathbb{R}^d$. In this case, we will say also that the admissible mass-preserving trajectory $\boldsymbol{\mu}$ is *driven* by $\boldsymbol{\nu}$.

(2) Let $T > 0$, $\boldsymbol{\mu}$ be an admissible mass-preserving trajectory defined on $[0,T]$ starting from $\bar{\mu}$ and driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$. We will say that $\boldsymbol{\mu}$ *is represented by* $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ if we have $e_t\sharp\boldsymbol{\eta} = \mu_t$ for all $t \in [0,T]$ and $\boldsymbol{\eta}$ is concentrated on the pairs $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ where γ is an absolutely continuous solution of the underlying characteristic system

$$\begin{cases} \dot{\gamma}(t) = v_t(\gamma(t)), & \text{for a.e.} \quad 0 < t \le T \\ \gamma(0) = x, \end{cases} \tag{3.1}$$

where $v_t(x) = \dfrac{\nu_t}{\mu_t}(x)$ is the density of the vector-valued measure ν_t w.r.t. μ_t.

Remark 3.2. Notice that, by definition, $\boldsymbol{\mu}$ is an admissible mass-preserving trajectory starting from $\bar{\mu}$ if it is a distributional solution of an homogeneous (*controlled*) continuity equation, $\partial_t\mu_t + \operatorname{div}(v_t\mu_t) = 0$, where v_t is a Borel velocity field ranging among $\mathrm{L}^1_{\mu_t}$-selections of the given set-valued function $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$, ruling the underlying finite-dimensional differential inclusion.

Remark 3.3. We recall that the existence of a probabilistic representation $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ for a mass-preserving trajectory $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is guaranteed by the Superposition Principle (see for example Thm. 8.2.1 in [1]) under mild integrability assumptions on the time-dependent vector field v_t. We also point out that generally, a probabilistic representation $\boldsymbol{\eta}$ for the same trajectory $\boldsymbol{\mu}$ is not unique.

Now, we are going to give the definition of *admissible clock-trajectory for an initial state* $\bar{\mu}$, which is associated with a pair $(f_0, \boldsymbol{\mu})$. In an *optimal equipment* interpretation of the problem, the function f_0 represents the amount of resources given at the beginning to each agent in $\operatorname{supp}\bar{\mu}$ in order to reach the given target S following an admissible mass-preserving trajetory $\boldsymbol{\mu}$. We consider the case in which f_0 depends only on the initial positions of the agents and we deal with the case of time-linear decrease of the provided items (ex. time), thus the admissibility requires that the density $f_t := f_0 - t$ is kept positive through the whole time evolution.

Notice that, since we want to define the admissible clock-trajectory for possible infinite times, we need to have a sequence of mass-preserving trajectories, each extending the previous one, defined in increasing finite time intervals. In this way, we can use results valid for separable metric spaces as Γ_T for every $0 < T < +\infty$.

Definition 3.4. Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued map, $S \subseteq \mathbb{R}^d$ be closed, nonempty and strongly invariant for F, $\bar{\mu} \in \mathscr{P}(\mathbb{R}^d)$ with $\operatorname{supp}(\bar{\mu}) \subseteq \mathbb{R}^d \setminus S$.
A Borel family of positive and finite measures $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ is an *admissible clock-trajectory (curve)* for $\bar{\mu}$ with target S if there exist a Borel map $f_0 : \mathbb{R}^d to [0,+\infty]$ called *clock-function*, and sequences $\{T_n\}_{n\in\mathbb{N}} \subseteq [0,+\infty[$, $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$, $\{\boldsymbol{\nu}^n\}_{n\in\mathbb{N}}$, and $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$ such that

(1) $T_n \to +\infty$;
(2) for any $n \in \mathbb{N}$ we have that $\boldsymbol{\mu}^n = \{\mu^n_t\}_{t\in[0,T_n]}$ is an admissible mass-preserving trajectory defined on $[0,T_n]$, starting from $\bar{\mu}$, driven by $\boldsymbol{\nu}^n := \{\nu^n_t\}_{t\in[0,T_n]}$, and represented by $\boldsymbol{\eta}_n$;

(3) given $n_1, n_2 \in \mathbb{N}$ with $T_{n_1} \le T_{n_2}$, we have $(\mathrm{Id}_{\mathbb{R}^d} \times r_{n_2,n_1})\sharp\boldsymbol{\eta}_{n_2} = \boldsymbol{\eta}_{n_1}$, where $r_{n_2,n_1} : \Gamma_{T_{n_2}} \to \Gamma_{T_{n_1}}$ is the linear and continuous operator defined by setting $r_{n_2,n_1}\gamma(t) = \gamma(t)$ for all $t \in [0, T_{n_1}]$. Clearly, $r_{n_2,n_1}\gamma \in \Gamma_{T_{n_1}}$ for all $\gamma \in \Gamma_{T_{n_2}}$. In particular, this implies $\mu_t^{n_1} = \mu_t^{n_2}$ for all $t \in [0, T_{n_1}]$.

(4) for any $n \in \mathbb{N}$, $t \in [0, T_n]$, $\varphi \in \mathrm{C}^0_C(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \varphi(x)\, d\tilde{\mu}_t = \iint_{\mathbb{R}^d \times \Gamma_{T_n}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(f_0(x) - t)\,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma)$$

In this case we will say that $\tilde{\boldsymbol{\mu}}$ *follows the family of mass-preserving trajectories* $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$. Notice that, since we ask $\tilde{\mu}_0(\mathbb{R}^d) < +\infty$, then we can identify f_0 with $\dfrac{\tilde{\mu}_0}{\bar{\mu}} \in \mathrm{L}^1_{\bar{\mu}}$.

Remark 3.5. Let $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 with clock-function f_0. Then we have $f_0(x) \ge T(x)$ for μ_0-a.e. $x \in \mathbb{R}^d$, where $T : \mathbb{R}^d \to [0, +\infty]$ is the classical minimum time function for the same target set $S \subseteq \mathbb{R}^d$. This follows necessarily, since by Definition 3.4 $\tilde{\mu}_t$ is a positive measure, hence we must have $f_0(x) \ge t$ for $\boldsymbol{\eta}_n$-a.e. (x, γ) such that $\gamma(t) \notin S$.

Proposition 3.6 (Clock trajectory and mass-preserving trajectory)**.** *Let $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ and $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ be an admissible clock-trajectory for μ_0 with clock-function f_0, following the family of mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}} := \{\{\mu_t^n\}_{t\in[0,T_n]}\}_n$ driven by $\{\boldsymbol{\nu}^n\}_{n\in\mathbb{N}} := \{\{\nu_t^n\}_{t\in[0,T_n]}\}_n$ and represented by $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$. Then for all $n \in \mathbb{N}$ we have $\tilde{\mu}_t \ll \mu_t^n$ for all $t \in [0, T_n]$.*

Proof. Let us consider any $n \in \mathbb{N}$ and any $t \in [0, T_n]$. We disintegrate $\boldsymbol{\eta}_n$ with respect to the continuous map $e_t : \mathbb{R}^d \times \Gamma_{T_n} \to \mathbb{R}^d$. This yields a family of probability measures $\{\eta_y^n\}_{y\in\mathbb{R}^d}$ which is uniquely defined $e_t\sharp\boldsymbol{\eta}_n$-a.e. such that

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)\,\mathrm{d}\tilde{\mu}_t(x) &= \int_{\mathbb{R}^d}\int_{e_t^{-1}(y)} \varphi(y)\chi_{S^c}(y)(f_0(\gamma(0)) - t)\,\mathrm{d}\eta_y^n(x,\gamma)\,\mathrm{d}\mu_t^n(y)\\ &= \int_{\mathbb{R}^d} \varphi(y)\chi_{S^c}(y)\left(\int_{e_t^{-1}(y)} f_0(\gamma(0))\,\mathrm{d}\eta_y^n(x,\gamma) - t\right)\mathrm{d}\mu_t^n(y),\end{aligned}$$

hence $\tilde{\mu}_t \ll \mu_t^n$ for all $t \in [0, T_n]$ and for all $n \in \mathbb{N}$. □

Definition 3.7 (Clock-generalized minimum time)**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ be a set-valued function, $S \subseteq \mathbb{R}^d$ be a target set for F. In analogy with the classical case, we define the *clock-generalized minimum time function* $\tau : \mathscr{P}(\mathbb{R}^d) \to [0, +\infty]$ by setting

$$\begin{aligned}\tau(\mu_0) := \inf\big\{\tilde{\mu}_0(\mathbb{R}^d) : \tilde{\boldsymbol{\mu}} := \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d) \text{ is an admissible clock-}\\ \text{-trajectory for the measure } \mu_0,\ \tilde{\mu}_{|t=0} = \tilde{\mu}_0\big\},\end{aligned} \tag{3.2}$$

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

Given $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$ with $\tau(\mu_0) < +\infty$, an admissible clock-curve $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ for μ_0 is *optimal for* μ_0 if

$$\tau(\mu_0) = \tilde{\mu}_{|t=0}(\mathbb{R}^d).$$

Given $p \ge 1$, we define also a clock-generalized minimum time function $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ by replacing in the above definitions $\mathscr{P}(\mathbb{R}^d)$ by $\mathscr{P}_p(\mathbb{R}^d)$ and $\mathscr{M}^+(\mathbb{R}^d)$ by $\mathscr{M}^+_p(\mathbb{R}^d)$. Since $\mathscr{M}^+_p(\mathbb{R}^d) \subseteq \mathscr{M}^+(\mathbb{R}^d)$, it is clear that $\tau_p(\mu_0) \ge \tau(\mu_0)$.

In order to prove a Dynamic Programming Principle for our minimization problem, which is the main task of this section, we will first provide a representation result expressing $\tau_p(\mu)$ as an average of the classical minimum-time function $T(\cdot)$, and then applying the well-known classical Dynamic Programming Principle (see for example Chap. I, Sect. 2 of [3]) holding for $T(\cdot)$. The main tools used for the proof of the following corollary are selection and disintegration results.

Corollary 3.8 (Optimal clock)**.** *Assume hypothesis* (F_0) *and* (F_1)*. Let* $S \subseteq \mathbb{R}^d$ *be a target set for* F*. Let* $p > 1$ *and* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ *be such that* $\|T(\cdot)\|_{\mathrm{L}^1_{\mu_0}} < +\infty$*. Then* $T(\cdot)$ *is the optimal clock function for* μ_0*.*

The proof of Corollary 3.8 can be found in Appendix A at page 737.

It is possible to notice that we can actually construct such optimal trajectories by approximation techniques as done in [15] for the case $\|T(\cdot)\|_{\mathrm{L}^\infty_{\mu_0}} < +\infty$.

Now we can deduce the following dynamic programming principle.

Corollary 3.9 (DPP for the clock problem)**.** *Assume hypothesis* (F_0) *and* (F_1)*. Let* $S \subseteq \mathbb{R}^d$ *be a target set for* F*. Let* $p > 1$ *and* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$*, with* $\mathrm{supp}\mu_0 \subseteq \mathbb{R}^d \setminus S$*, be such that* $\|T(\cdot)\|_{\mathrm{L}^1_{\mu_0}} < +\infty$*. We have*

$$\tau_p(\mu_0) = \int_{\mathbb{R}^d} T(x)\,\mathrm{d}\mu_0(x).$$

Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ *be an admissible clock-trajectory for* μ_0 *following a family of admissible mass-preserving trajectories* $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ *starting from* μ_0*. For any* $s \geq 0$ *we choose* $n > 0$ *such that* $\boldsymbol{\mu}^n$ *is defined on an interval* $[0, T_n]$ *containing* s *and it is represented by* $\boldsymbol{\eta}_n \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$*. Then we have*

$$\tau_p(\mu_0) = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} T(\gamma(0))\,\mathrm{d}\boldsymbol{\eta}_n \leq \iint_{\mathbb{R}^d\times\Gamma_{T_n}} [T(\gamma(s)) + s]\,\mathrm{d}\boldsymbol{\eta}_n \leq s + \tau_p(\mu^n_s).$$

Moreover, if $\boldsymbol{\eta}_n$ *is concentrated on (restriction to* $[0, T_n]$ *of) time-optimal trajectories, then for all* $s \geq 0$ *such that* $\mathrm{supp}\mu^n_s \subseteq \mathbb{R}^d \setminus S$*, we have*

$$\tau_p(\mu_0) = s + \tau_p(\mu^n_s),$$

and so for such $s \geq 0$ *we have*

$$\tau_p(\mu_0) = \inf_{\boldsymbol{\mu}} \{s + \tau_p(\mu_s)\},$$

where the infimum is taken on admissible mass-preserving trajectories $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,s]}$ *satisfying* $\mu_{t=0} = \mu_0$*.*

The proof is a direct consequence of Corollary 3.8, of the classical Dynamic Programming Principle holding for $T(\cdot)$ and Remark 3.5.

Remark 3.10. We notice that, in the same hypothesis of Corollary 3.8, if $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ we have that $\tau_p(\mu) = \|T(\cdot)\|_{\mathrm{L}^1_\mu} \leq \|T(\cdot)\|_{\mathrm{L}^\infty_\mu} = \tilde{T}_p(\mu)$, where $\tilde{T}_p(\cdot)$ is the generalized minimum time function studied in [13] for the mass-preserving case, with generalized target set $\tilde{S} := \{\sigma \in \mathscr{P}(\mathbb{R}^d) : \mathrm{supp}\,\sigma \subseteq S\}$, *i.e.* $\tilde{T}_p(\mu)$ is the minimum of the final times $\bar{t}$ for which $\mathrm{supp}(\mu_{|t=\bar{t}}) \subseteq S$, for an admissible trajectory starting from μ. In particular, we refer to [13] for the last equivalence holding in this situation.

3.1. Regularity results

Thanks to Corollary 3.8, under suitable assumptions, the clock-generalized minimum time function inherits regularity results from the classical one as shown in the next corollaries. For the following result, we refer to [18] for conditions under which the classical minimum time function $T(\cdot)$ is l.s.c.

Corollary 3.11 (L.s.c. of the clock-generalized minimum time function)**.** *Assume that* $T(\cdot)$ *is l.s.c. Assume hypothesis* (F_0) *and* (F_1)*. Let* $S \subseteq \mathbb{R}^d$ *be a target set for* F*. Let* $p > 1$ *and* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$*, with* $\mathrm{supp}\mu_0 \subseteq \mathbb{R}^d \setminus S$*, be such that* $\|T(\cdot)\|_{\mathrm{L}^1_{\mu_0}} < +\infty$*. Then* $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ *is l.s.c.*

Proof. Taken a sequence $\{\mu^n_0\}_{n\in\mathbb{N}} \subseteq \mathscr{P}_p(\mathbb{R}^d)$ s.t. $W_p(\mu^n_0, \mu_0) \to 0$ for $n \to +\infty$, we want to prove that $\tau_p(\mu_0) \leq \liminf_{n\to+\infty} \tau_p(\mu^n_0)$.

By Remark 3.5, Lemma 5.1.7. in [1] and Corollary 3.8, we conclude immediately that

$$\liminf_{n\to+\infty} \tau_p(\mu^n_0) \geq \liminf_{n\to+\infty} \int_{\mathbb{R}^d} T(x)\,\mathrm{d}\mu^n_0(x) \geq \int_{\mathbb{R}^d} T(x)\,\mathrm{d}\mu_0(x) = \tau_p(\mu_0).$$ □

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

Representation formula for the generalized minimum time provided in Corollary 3.8 allows us to recover many results valid for the classical minimum time function also in the framework of the generalized systems. We refer the reader to Chapter 2, Section 4.2 in [3], to Chapter 2 in [9] and to Sections 2 and 3 in [8] for a definition and classical results about semiconcave functions, in particular regarding the classical minimum time function. In the space $\mathscr{P}_p(\mathbb{R}^d)$ we will use the following definition which is a strong formulation of the concept of λ-*semiconcavity along geodesics* outlined in Definition 9.1.1 in [1]. We refer the reader to Definition 2.4.2, Section 7.2 and Theorem 7.2.2 in [1] for a characterization of (*constant speed*) *geodesics* in $\mathscr{P}_p(\mathbb{R}^d)$.

Definition 3.12 (Strong semiconcavity along W_2-geodesics)**.** Let $p \geq 2$ and $g : \mathscr{P}_p(\mathbb{R}^d) \to]-\infty, +\infty]$. We say that g is *strongly* W_2*-geodesically semiconcave* in $\mathscr{P}_p(\mathbb{R}^d)$ if there exists $D > 0$ such that for every couple $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)$ and for any optimal transport plan $\pi \in \Pi(\mu_1, \mu_2)$ for the 2-Wasserstein distance, we have

$$g(\mu_t) \geq tg(\mu_1) + (1-t)g(\mu_2) - Dt(1-t)W_2^2(\mu_1, \mu_2), \quad \forall t \in [0,1],$$

where $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu_t := \big(t\mathrm{pr}^1 + (1-t)\mathrm{pr}^2\big)\sharp\pi$, *i.e.* $\boldsymbol{\mu}$ is a (*constant speed*) W_2*-geodesic connecting* μ_1 *and* μ_2 *through* π, and $\mathrm{pr}^i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $i = 1, 2$, is the projection on the i–th component, *i.e.*, $\mathrm{pr}^i(x_1, x_2) = x_i$.

Corollary 3.13 (Controllability)**.** *Assume* (F_0), (F_1), (F_2). *Let* $S \subseteq \mathbb{R}^d$ *be a target set for* F. *Assume furthermore that for every* $R > 0$ *there exist* η_R, $\sigma_R > 0$ *such that for a.e.* $x \in B(0,R) \setminus S$ *with* $d_S(x) \leq \sigma_R$ *there holds*

$$\sup_{v\in F(x)} \langle -\nabla d_S(x), v\rangle > \eta_R\,, \tag{3.3}$$

where $d_S : \mathbb{R}^d \to \mathbb{R}$ *denotes the distance function from* S. *Then, if we set for* $p > 1$

$$\mathscr{P}_p(\mathbb{R}^d)_{|R} := \left\{\mu \in \mathscr{P}_p(\mathbb{R}^d) : \|T(\cdot)\|_{\mathrm{L}^1_\mu} < +\infty \text{ and } \operatorname{supp}\mu \subseteq \overline{B(0,R)} \cap \{x \,:\, d_S(x) \leq \sigma_R\}\right\},$$

there exists $c_R > 0$ *such that for every* $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)_{|R}$ *the following properties hold.*

(1) $\tau_p(\mu_0) \leq \dfrac{1}{c_R}\|d_S\|_{\mathrm{L}^1_{\mu_0}}$.
(2) *The function* $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ *is Lipschitz continuous on* $\mathscr{P}_p(\mathbb{R}^d)_{|R}$ *with respect to the metric* W_p.
(3) *If* $\partial S \in C^{1,1}$, $p \geq 2$, *then the function* $\tau_p : \mathscr{P}_p(\mathbb{R}^d) \to [0, +\infty]$ *is strongly* W_2*-geodesically semiconcave on*

$$\{\mu \in \mathscr{P}_p(\mathbb{R}^d)_{|R} : \operatorname{supp}\mu \cap S = \emptyset\}.$$

Proof. According to Proposition 2.2 in [8], the present assumptions imply that there exists a constant $c_R > 0$ such that the classical minimum time function satisfies

$$T(x) \leq \frac{1}{c_R} d_S(x)\,, \tag{3.4}$$

for every $x \in B(0,R) \setminus S$ with $d_S(x) \leq \sigma_R$. Moreover, $T(\cdot)$ is Lipschitz continuous in such set. We denote by $k_R > 0$ its Lipschitz constant.

Now, property (1) follows from (3.4) and Corollary 3.8, since

$$\tau_p(\mu_0) = \int_{\mathbb{R}^d} T(x)\,\mathrm{d}\mu_0 \leq \frac{1}{c_R}\int_{\mathbb{R}^d} d_S(x)\,\mathrm{d}\mu_0 = \frac{1}{c_R}\|d_S\|_{\mathrm{L}^1_{\mu_0}}.$$

To prove (2), fix $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)_{|R}$. By setting

$$c'_R := \frac{c_R}{(1+c_R)(1+k_R)},$$

we have that the function $c'_R T(\cdot)$ is Lipschitz continuous with constant less than 1 and that $c'_R T(\cdot) \le \sigma_R$. Hence, it can be extended to a continuous bounded function on the whole $\mathbb{R}^d$, and $|c'_R T(x) - c'_R T(y)| \le |x-y|$ for all $x, y \in B(0,R) \setminus S$ with $d_S(x), d_S(y) \le \sigma_R$. According to Kantorovich duality (2.2) and Corollary 3.8 we then have

$$c'_R(\tau_p(\mu_1) - \tau_p(\mu_2)) \le c'_R W_1(\mu_1, \mu_2) \le c'_R W_p(\mu_1, \mu_2),$$

using Hölder inequality. By switching the roles of μ_1 and μ_2, we obtain (2).

Finally, according to Theorem 3.1 in [8], when $\partial S \in C^{1,1}$ we have that the classical minimum time function is semiconcave in $\{x : T(x) < +\infty\} \setminus S$. In particular, there exists $D_R > 0$ such that

$$T(tx_1 + (1-t)x_2) \ge tT(x_1) + (1-t)T(x_2) - D_R\, t(1-t)\, |x_1 - x_2|^2, \tag{3.5}$$

for every $x_1, x_2 \in \{x : T(x) < +\infty\} \setminus S$.

Let $K := \overline{B(0,R)} \cap \{x : d_S(x) \le \sigma_R\}$ and $p \ge 2$. For any Borel sets $A, B \subseteq \mathbb{R}^d$ and $\pi \in \Pi(\mu_1, \mu_2)$, we now have

$$A \times B \subseteq [(A \times B) \cap (K \times K)] \cup [(A \setminus K) \times \mathbb{R}^d] \cup [\mathbb{R}^d \times (B \setminus K)],$$

so that

$$\begin{aligned}\pi(A \times B) &\le \pi((A \times B) \cap (K \times K)) + \mu_1(A \setminus K) + \mu_2(B \setminus K) \\ &= \pi((A \times B) \cap (K \times K)),\end{aligned}$$

because μ_1 and μ_2 are concentrated on K. In particular, $\operatorname{supp}(\pi) \subseteq K \times K$.

Let $\pi \in \Pi(\mu_1, \mu_2)$ be any optimal transport plan realizing the 2-Wasserstein distance between μ_1 and μ_2, so that $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,1]} \subseteq \mathscr{P}_p(\mathbb{R}^d)$, $\mu_t := \big(t\mathrm{pr}^1 + (1-t)\mathrm{pr}^2\big)\sharp\pi$. We integrate the estimate (3.5) to find that, by using Remark 3.5 and Corollary 3.8,

$$\begin{aligned}\tau_p(\mu_t) &\ge \int_{\mathbb{R}^d} T(x)\,\mathrm{d}\mu_t(x) \\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} T(tx_1 + (1-t)x_2)\,\mathrm{d}\pi(x_1, x_2) \\ &\ge t\int_{\mathbb{R}^d} T(x_1)\,\mathrm{d}\mu_1 + (1-t)\int_{\mathbb{R}^d} T(x_2)\,\mathrm{d}\mu_2 \\ &\qquad - D_R\, t\,(1-t) \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x_1 - x_2|^2\,\mathrm{d}\pi(x_1, x_2) \\ &= t\tau_p(\mu_1) + (1-t)\tau_p(\mu_2) - D_R\, t(1-t)\, W_2^2(\mu_1, \mu_2).\end{aligned}$$

□

Remark 3.14. For other controllability conditions generalizing (3.3), the reader may refer *e.g.* to [16] or [20].

4. Hamilton-Jacobi-Bellman equation

In this section we will prove that under the assumptions granting the validity of the Dynamic Programming Principle and of a result which aims to recover the initial velocity of admissible trajectories, the clock-generalized minimum time function solves a natural Hamilton–acobi–Bellman equation on $\mathscr{P}_2(\mathbb{R}^d)$ in a suitable viscosity sense. We observe also that once we have the Dynamic Programming Principle and once the problem is modeled on the same notion of admissible mass-preserving trajectories, then the Hamilton-Jacobi-Bellman equation related to the present problem is the same considered in [13,14]. We then follow a very similar approach as the one discussed in Section 4 of [13] or in Section 3 of [14] in which a more regular case is treated.

We point out that, if we restrict the study to the class of absolutely continuous curves in $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ with a Lipschitz continuous value function, then a Comparison Principle is provided in [10]. However in general, even if a more general Comparison Principle, still restricted to Lispchitz continuous value functions, is expected in the forthcoming paper [21], the uniqueness problem for such kind of HJB equations is still largely open.

The following proposition allows to construct an admissible mass-preserving trajectory concentrated on characteristics of class C^1 with initial velocity the given one.

Proposition 4.1. *Assume hypothesis* (F_0), (F_1). *Let* $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, *and* $x \mapsto v_x$ *be a Borel selection of* F *belonging to* L^2_μ. *Then for any* $T > 0$ *there exists an admissible mass-preserving curve* $\boldsymbol{\mu}$ *defined on* $[0,T]$ *starting from* μ *and represented by* $\boldsymbol{\eta}$ *such that for* $\boldsymbol{\eta}$*-a.e.* $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ *we have that* $\gamma \in \mathrm{C}^1([0,T])$, $\dot\gamma(t) \in F(\gamma(t))$ *for all* $t \in [0,T]$, $\gamma(0) = x$ *and* $\dot\gamma(0) = v_x$.

Proof. Let $T > 0$ be fixed. Consider the set-valued map $G : \mathbb{R}^d \rightrightarrows \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$ defined by

$$G(x) := \left\{ v \in \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d) : \, v(x) = v_x, \, v(y) \in F(y) \text{ for all } y \in \mathbb{R}^d \right\},$$

and notice that, recalling the assumptions on F, we have that $G(x)$ is nonempty, convex and closed. Indeed, for every $x \in \mathbb{R}^d$ and $v_x \in F(x)$ there exists by Michael's continuous selection Theorem a continuous selection v of F such that $v(x) = v_x$.

Define the map $g : \mathbb{R}^d \times \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d) \to [0,+\infty]$ by setting

$$g(x,v) := \sup_{q,y\in\mathbb{R}^d} \left\{ I_{F(y)}(v(y)) + \langle q, v_x - v(x)\rangle \right\},$$

noticing that $v \in G(x)$ if and only if $g(x,v) = 0$.

To prove that g is a Borel map, it is enough to show that $(v,y) \mapsto I_{F(y)}(v(y))$ is a Borel map from $\mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d) \times \mathbb{R}^d$ to $\{0,+\infty\}$.

Indeed, consider any sequence $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$ uniformly convergent to $v \in \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$, and $\{y_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ converging to y.

Then, $v_n(y_n) \to v(y)$, $n \to +\infty$. Indeed, denoted with $\omega_y(\cdot)$ a modulus of continuity for v at the point y, we have

$$\begin{aligned} |v_n(y_n) - v(y)| &\le |v_n(y_n) - v(y_n)| + |v(y_n) - v(y)| \\ &\le \|v_n - v\|_{\mathrm{L}^\infty} + \omega_y(|y_n - y|), \end{aligned}$$

for a suitable $s > 0$. Hence, we deduce that

$$\liminf_{n\to+\infty} I_{F(y_n)}(v_n(y_n)) \ge I_{F(y)}(v(y)),$$

where we used the fact that the map $f : \mathbb{R}^d \times \mathbb{R}^d \to \{0,+\infty\}$, $f(z,w) := I_{F(z)}(w)$, is l.s.c. due to u.s.c. of F.

Thus we have just proved that $(v,y) \mapsto I_{F(y)}(v(y))$ is l.s.c. and hence a Borel map. Hence $\operatorname{Graph} G = g^{-1}(0)$ is a Borel set. By Theorem 8.1.4 p. 310 in [2], we have that the set-valued map $G : \mathbb{R}^d \rightrightarrows \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$ is Borel measurable, and so by Theorem 8.1.3 p. 308 in [2] it admits a Borel selection $V : \mathbb{R}^d \to \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$. We denote $V(x) \in \mathrm{C}^0(\mathbb{R}^d;\mathbb{R}^d)$ by V_x.

We fix a family of smooth mollifiers $\{\rho_\varepsilon\}_{\varepsilon>0} \subseteq \mathrm{C}^\infty_C(\mathbb{R}^d)$ such that $\operatorname{supp}\rho_\varepsilon \subseteq \overline{B(0,\varepsilon)}$, and denote by $H^T_{x,\varepsilon}$ the (unique) $\gamma \in \Gamma_T$ satisfying $\dot\gamma(t) = (V_x * \rho_\varepsilon) \circ \gamma(t)$, $\gamma(0) = x$. We want to prove that $H^T_{x,\varepsilon}$ is a Borel map in x.

For any $x \in \mathbb{R}^d$ and $W \in \mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d)$ denote by $h_{x,W}(t)$ the solution of $\dot{x}(t) = W \circ x(t)$, $x(0) = x$. The map $h : \mathbb{R}^d \times \mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d) \to \Gamma_T$ is continuous, hence Borel, since for all $x, y \in \mathbb{R}^d$, $W_1, W_2 \in \mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^d)$, we have

$$\begin{aligned}
|h_{x,W_1}(t) - h_{y,W_2}(t)| &\le |x-y| + \int_0^t |W_1(h_{x,W_1}(s)) - W_2(h_{y,W_2}(s))|\,\mathrm{d}s \\
&\le |x-y| + \int_0^t |W_1(h_{x,W_1}(s)) - W_1(h_{y,W_2}(s))|\,\mathrm{d}s \\
&\quad + \int_0^t |W_1(h_{y,W_2}(s)) - W_2(h_{y,W_2}(s))|\,\mathrm{d}s \\
&\le |x-y| + \mathrm{Lip}(W_1)\int_0^t |h_{x,W_1}(s)) - h_{y,W_2}(s)|\,\mathrm{d}s + t\|W_1 - W_2\|_\infty
\end{aligned}$$

and so by Gronwall's inequality

$$|h_{x,W_1}(t) - h_{y,W_2}(t)| \le (|x-y| + t\|W_1 - W_2\|_\infty)\mathrm{e}^{t\mathrm{Lip}(W_1)},$$

which implies

$$\|h_{x,W_1} - h_{y,W_2}\|_\infty \le (|x-y| + T\|W_1 - W_2\|_\infty)\mathrm{e}^{T\mathrm{Lip}(W_1)}.$$

Since $H^T_{x,\varepsilon}$ can be written as the composition of the Borel maps $x \mapsto (x, V_x)$, $(x, Z) \mapsto (x, Z * \rho_\varepsilon)$, and $(x, W) \mapsto h_{x,W}$, we have that it is a Borel map.

Finally, we define the Kuratowski upper limit of $H^T_{x,\varepsilon}$ by

$$H^T(x) := \{\gamma \in \Gamma_T : \text{ there exists } \{\varepsilon_n\}_{n\in\mathbb{N}} \text{ s.t. } \varepsilon_n \to 0^+, H^T_{x,\varepsilon_n} \to \gamma, \text{ as } n \to +\infty\}.$$

Thanks to Theorem 8.2.5 in [2], this is a Borel set-valued map from $\mathbb{R}^d$ to Γ_T, thus possesses a Borel selection $\psi : \mathbb{R}^d \to \Gamma_T$.

Given $x \in \mathbb{R}^d$, let $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be such that $\varepsilon_n \to 0^+$ and $H^T_{x,\varepsilon_n} \to \gamma_x := \psi(x)$. In particular, we have that $H^T_{x,\varepsilon_n}(0) = x$ for all $n \in \mathbb{N}$, and so $\gamma_x(0) = x$. Since there exists a compact K containing $H^T_{x,\varepsilon_n}(\tau)$ for all $n \in \mathbb{N}$ sufficiently large and all $\tau \in [0, T]$, and moreover $V_x * \rho_{\varepsilon_n}$ converges to V_x in $\mathrm{C}^0(\mathbb{R}^d)$ on all the compact sets of $\mathbb{R}^d$, we can pass to the limit by Dominated Convergence Theorem in

$$\frac{H^T_{x,\varepsilon_n}(s) - H^T_{x,\varepsilon_n}(t)}{s-t} = \frac{1}{s-t}\int_t^s V_x * \rho_{\varepsilon_n}(H^T_{x,\varepsilon_n}(\tau))\,\mathrm{d}\tau,$$

obtaining

$$\frac{\gamma_x(s) - \gamma_x(t)}{s-t} = \frac{1}{s-t}\int_t^s V_x(\gamma_x(\tau))\,\mathrm{d}\tau, \tag{4.1}$$

thus $\gamma_x \in \mathrm{C}^1$ is an admissible curve satisfying $\dot{\gamma}_x(0) = v_x$.

We define the probability measure

$$\boldsymbol{\eta} := \mu \otimes \delta_{\gamma_x} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T),$$

which, as seen in the last part of the proof of Lemma A.3, induces an admissible trajectory $\boldsymbol{\mu} = \{\mu_t = e_t \sharp \boldsymbol{\eta}\}_{t\in[0,T]}$. Moreover, we prove that

$$\lim_{t\to 0}\left\|\frac{e_t - e_0}{t} - v_x\right\|_{\mathrm{L}^2_{\boldsymbol{\eta}}} = 0.$$

Indeed,

$$\left\|\frac{e_t-e_0}{t}-v_x\right\|^2_{\mathrm{L}^2_{\boldsymbol{\eta}}}=\int_{\mathbb{R}^d}\int_{\Gamma^x_T}\left|\frac{\gamma(t)-\gamma(0)}{t}-v_x\right|^2\,\mathrm{d}\delta_{\gamma_x}(\gamma)\,\mathrm{d}\mu(x)$$
$$=\int_{\mathbb{R}^d}\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right|^2\,\mathrm{d}\mu(x),$$

and for μ-a.e. $x\in\mathbb{R}^d$, recalling (4.1), continuity of $V_x(\cdot)$ and that $\gamma\in\mathrm{C}^1$ and $\dot\gamma(0)=v_x$, we have

$$\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right|=\left|\frac{1}{t}\int_0^t V_x(\gamma_x(\tau))\,\mathrm{d}\tau-v_x\right|$$
$$\le\frac{1}{t}\int_0^t|V_x(\gamma_x(\tau))|\,\mathrm{d}\tau+|v_x|$$
$$\le\max_{t\in[0,T]}|V_x(\gamma_x(t))|+|v_x|,$$
$$\lim_{t\to0^+}\left|\frac{\gamma_x(t)-\gamma_x(0)}{t}-v_x\right|=0.$$

Thus we conclude applying Lebesgue's Dominated Convergence Theorem. □

Corollary 4.2. *Assume hypothesis (F_0), (F_1). Let $\mu\in\mathscr{P}_2(\mathbb{R}^d)$, $T>0$. Define the set $A_T(\mu)$ of the maps $w_{\boldsymbol{\eta}}\in\mathrm{L}^2_{\boldsymbol{\eta}}$ satisfying the following*

(1) *there exists an admissible mass-preserving trajectory $\boldsymbol{\mu}$ defined on $[0,T]$ and represented by $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Gamma_T)$ with $e_0\sharp\boldsymbol{\eta}=\mu$,*
(2) *there exists a sequence $\{t_i\}_{i\in\mathbb{N}}\subseteq]0,T]$ such that $t_i\to0$ and*

$$\lim_{i\to+\infty}\frac{1}{t_i}\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),e_{t_i}(x,\gamma)-e_0(x,\gamma)\rangle\,\mathrm{d}\boldsymbol{\eta}=\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),w_{\boldsymbol{\eta}}(x,\gamma)\rangle\,\mathrm{d}\boldsymbol{\eta},$$

for all $p\in\mathrm{L}^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$.

Then $A_T(\mu)=\{v\circ e_0:\ v\in\mathrm{L}^2_\mu,\ v(x)\in F(x)$ for μ-a.e. $x\in\mathbb{R}^d\}$.

Proof. It is trivial that $A_T(\mu)$ is contained in the right hand side. The opposite inclusion follows from the previous Proposition with $v(x)=v_x$, noticing also that since $v\in\mathrm{L}^2_\mu$, then $v\circ e_0\in\mathrm{L}^2_{\boldsymbol{\eta}}$ with $\boldsymbol{\eta}$ as in (4.2) by Lemma A.1.

Indeed, in Proposition 4.1 we proved strong convergence in $\mathrm{L}^2_{\boldsymbol{\eta}}$ of $\dfrac{e_t-e_0}{t}$ to v_x for $t\to0$. Hence we have weak convergence, in particular since $p\circ e_0\in\mathrm{L}^2_{\boldsymbol{\eta}}$ for every $p\in\mathrm{L}^2_\mu$ by Lemma A.1, then there exists a sequence $\{t_i\}_{i\in\mathbb{N}}\subseteq]0,T]$ such that $t_i\to0$ and

$$\lim_{i\to+\infty}\frac{1}{t_i}\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),e_{t_i}(x,\gamma)-e_0(x,\gamma)\rangle\,\mathrm{d}\boldsymbol{\eta}=\iint_{\mathbb{R}^d\times\Gamma_T}\langle p\circ e_0(x,\gamma),v\circ e_0(x,\gamma)\rangle\,\mathrm{d}\boldsymbol{\eta},$$

thus item (4.2) is satisfied with $w_{\boldsymbol{\eta}}=v\circ e_0$, and item (4.2) follows directly by the previous Proposition. □

Definition 4.3 (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 $p_\mu\in\mathrm{L}^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ 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

$$\limsup_{t\to0^+}\frac{V(e_t\sharp\boldsymbol{\eta})-V(e_0\sharp\boldsymbol{\eta})-\displaystyle\iint_{\mathbb{R}^d\times\Gamma_T}\langle p_\mu\circ e_0(x,\gamma),e_t(x,\gamma)-e_0(x,\gamma)\rangle\,\mathrm{d}\boldsymbol{\eta}(x,\gamma)}{\|e_t-e_0\|_{\mathrm{L}^2_{\boldsymbol{\eta}}}}\le\delta. \tag{4.2}$$

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

Definition 4.4 (Viscosity solutions)**.** Let $V : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function and $\mathscr{H} : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$, where $(\mu, \psi) \in T^*\mathscr{P}_2(\mathbb{R}^d)$ iff $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\psi \in \mathrm{L}^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$. 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 $\mathscr{H}(\mu, q_\mu) \geq -C\delta$ 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 $\mathscr{H}(\mu, p_\mu) \leq C\delta$ 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 4.5 (Hamiltonian Function)**.** Given $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, we define the map $\mathscr{H}_F : T^*\mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ by setting

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

For the proof of the following theorem we used the same procedure adopted in [13,14] for the mass-preserving case.

Theorem 4.6 (Viscosity solution)**.** *Let $S \subseteq \mathbb{R}^d$ be a target set for F. Let $\mathcal{A}$ be any open subset of $\mathscr{P}_2(\mathbb{R}^d)$ with uniformly bounded 2−moments and such that if $\mu \in \mathcal{A}$ then* $\mathrm{supp}\mu \subseteq \mathbb{R}^d \setminus S$. *Assume hypothesis (F_0), (F_1). Assume that $\|T(\cdot)\|_{\mathrm{L}^1_\mu} < +\infty$ for all $\mu \in \mathcal{A}$ and that $\tau_2 : \mathscr{P}_2(\mathbb{R}^d) \to [0, +\infty]$ is continuous on $\mathcal{A}$. Then $\tau_2(\cdot)$ is a viscosity solution of $\mathscr{H}_F(\mu, D\tau_2(\mu)) = 0$ on $\mathcal{A}$, with $\mathscr{H}_F$ defined as in Definition 4.5.*

Proof. The proof is splitted in two claims.

Claim 1. $\tau_2(\cdot)$ is a subsolution of $\mathscr{H}_F(\mu, D\tau_2(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 1. Let $\mu_0 \in \mathcal{A}$. Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 following a family of admissible mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ starting from μ_0. For any $s \geq 0$ we choose $n > 0$ such that $\boldsymbol{\mu}^n$ is defined on an interval $[0, T_n]$ containing s and it is represented by $\boldsymbol{\eta}_n \in \mathscr{P}(\mathbb{R}^d \times \varGamma_{T_n})$. Then by the Dynamic Programming Principle we have $\tau_2(\mu_0) \leq \tau_2(\mu^n_s) + s$ for all $s > 0$. Without loss of generality, we can assume $0 < s < 1$. Given any $p_{\mu_0} \in D^+_\delta \tau_2(\mu_0)$, and set

$$\begin{aligned}
A(s, p_{\mu_0}, \boldsymbol{\eta}_n) &:= -\, s - \iint_{\mathbb{R}^d \times \varGamma_{T_n}} \langle p_{\mu_0} \circ e_0(x, \gamma), e_s(x, \gamma) - e_0(x, \gamma) \rangle \, \mathrm{d}\boldsymbol{\eta}_n, \\
B(s, p_{\mu_0}, \boldsymbol{\eta}_n) &:= \tau_2(\mu^n_s) - \tau_2(\mu_0) - \iint_{\mathbb{R}^d \times \varGamma_{T_n}} \langle p_{\mu_0} \circ e_0(x, \gamma), e_s(x, \gamma) - e_0(x, \gamma) \rangle \mathrm{d}\boldsymbol{\eta}_n,
\end{aligned}$$

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

We recall that since by definition $p_{\mu_0} \in \mathrm{L}^2_{\mu_0}$, we have that $p_{\mu_0} \circ e_0 \in \mathrm{L}^2_{\boldsymbol{\eta}_n}$ by Lemma A.1. Dividing by $s > 0$, we obtain that

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

for all $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, with $A_{T_n}(\mu_0)$ defined as in Corollary 4.2.

Recalling the choice of p_{μ_0}, we have

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

where $K > 0$ is a suitable constant coming from Lemma A.1 and from hypothesis.

We thus obtain for all $\boldsymbol{\eta}_n$ as above and all $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, that

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

By passing to the infimum on $\boldsymbol{\eta}_n$ and $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$, and recalling Corollary 4.2, we have

$$\begin{aligned}
-K\delta &\leq 1 + \inf_{\substack{v \in \mathrm{L}^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e } x}} \iint_{\mathbb{R}^d \times \Gamma_{T_n}} \langle p_{\mu_0} \circ e_0(x,\gamma), v \circ e_0(x,\gamma)\rangle \,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma)\\
&= 1 + \inf_{\substack{v \in \mathrm{L}^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e } x}} \int_{\mathbb{R}^d} \int_{\Gamma^x_{T_n}} \langle p_{\mu_0} \circ e_0(x,\gamma), v \circ e_0(x,\gamma)\rangle \, d\eta^n_x(\gamma)\,\mathrm{d}\mu_0(x)\\
&= 1 + \inf_{\substack{v \in \mathrm{L}^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e } x}} \int_{\mathbb{R}^d} \langle p_{\mu_0}, v\rangle \,\mathrm{d}\mu_0 = -\mathscr{H}_F(\mu_0, p_{\mu_0}),
\end{aligned}$$

so $\tau_2(\cdot)$ is a subsolution, thus confirming Claim 1. ◇

Claim 2. $\tau_2(\cdot)$ is a supersolution of $\mathscr{H}_F(\mu, D\tau_2(\mu)) = 0$ on $\mathcal{A}$.

Proof of Claim 2. Let $\mu_0 \in \mathcal{A}$. Let $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in[0,+\infty[}$ be an admissible clock-trajectory for μ_0 following a family of admissible mass-preserving trajectories $\{\boldsymbol{\mu}^n\}_{n\in\mathbb{N}}$ starting from μ_0. For any $s \geq 0$ we choose $n > 0$ such that $\boldsymbol{\mu}^n$ is defined on an interval $[0, T_n]$ containing s and it is represented by $\boldsymbol{\eta}_n \in \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$. Taken $q_{\mu_0} \in D^-_\delta \tau_2(\mu_0)$, there is a sequence $\{s_i\}_{i\in\mathbb{N}} \subseteq]0, T_n[$, $s_i \to 0^+$ and $w_{\boldsymbol{\eta}_n} \in A_{T_n}(\mu_0)$ as in Corollary 4.2 such that for all $i \in \mathbb{N}$

$$\begin{aligned}
&\iint_{\mathbb{R}^d \times \Gamma_{T_n}} \langle q_{\mu_0} \circ e_0(x,\gamma), \frac{e_{s_i}(x,\gamma) - e_0(x,\gamma)}{s_i}\rangle \,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma)\\
&\qquad\qquad \leq 2\delta \left\| \frac{e_{s_i} - e_0}{s_i}\right\|_{\mathrm{L}^2_{\boldsymbol{\eta}_n}} - \frac{\tau_2(\mu_0) - \tau_2(\mu^n_{s_i})}{s_i}.
\end{aligned}$$

By taking i sufficiently large we thus obtain

$$\iint_{\mathbb{R}^d \times \Gamma_{T_n}} \langle q_{\mu_0} \circ e_0(x,\gamma), w_{\boldsymbol{\eta}_n}(x,\gamma)\rangle \,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma) \leq 3K\delta - \frac{\tau_2(\mu_0) - \tau_2(\mu^n_{s_i})}{s_i}.$$

By using Corollary 4.2 and arguing as in Claim 1, we have

$$\inf_{\substack{v \in \mathrm{L}^2_{\mu_0}(\mathbb{R}^d;\mathbb{R}^d)\\ v(x)\in F(x)\ \mu_0\text{-a.e } x}} \iint_{\mathbb{R}^d \times \Gamma_{T_n}} \langle q_{\mu_0} \circ e_0(x,\gamma), v \circ e_0(x,\gamma)\rangle \,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma) = -\mathscr{H}_F(\mu_0, q_{\mu_0}) - 1,$$

and so

$$\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -3K\delta + \frac{\tau_2(\mu_0) - \tau_2(\mu^n_{s_i})}{s_i} - 1.$$

By the Dynamic Programming Principle, passing to the infimum on all admissible curves and recalling that $\dfrac{\tau_2(\mu_0) - \tau_2(\mu^n_s)}{s} - 1 \leq 0$ with equality holding if and only if $\boldsymbol{\eta}_n$ is concentrated on time-optimal trajectories, we obtain $\mathscr{H}_F(\mu_0, q_{\mu_0}) \geq -C'\delta$, which proves that $\tau_2(\cdot)$ is a supersolution, thus confirming Claim 2. □

A. Technical results

This section contains some technical results used in the paper.

The proof of the following technical lemma can be found in [11] for the case $p = 2$, but it is not hard to generalize it to $p \geq 1$, recalling that $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ for any $a, b \geq 0$.

Lemma A.1 (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 mass-preserving trajectory driven by* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$ *and represented by* $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$*. 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) $e_t \in \mathrm{L}^p_{\boldsymbol{\eta}}(\mathbb{R}^d \times \Gamma_T; \mathbb{R}^d)$ *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*

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

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

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

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

Corollary A.2 (Uniform p-integrability)**.** *Assume hypothesis* (F_0)*,* (F_1)*. Let* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *be an admissible mass-preserving trajectory driven by* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$*,* $p > 1$*, and set* $v_t(x) = \dfrac{\nu_t}{\mu_t}(x)$*. Assume that* $\mathrm{m}_p(\mu_0) < +\infty$*, then*

$$\int_0^T \int_{\mathbb{R}^d} |v_t(x)|^p \,\mathrm{d}\mu_t \,\mathrm{d}t < +\infty,$$

hence the assumptions of the Superposition Principle (*see for example Thm.* 8.2.1 *in* [1]) *are satisfied.*

Proof. We have

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} |v_t(x)|^p \,\mathrm{d}\mu_t \,\mathrm{d}t &\leq TC^p \int_{\mathbb{R}^d} (|x|+1)^p \,\mathrm{d}\mu_t \leq 2^{p-1} TC^p (\mathrm{m}_p(\mu_t) + 1),\\ &\leq K(\mathrm{m}_p(\mu_0) + 1),\end{aligned}$$

for a suitable constant $K > 0$ depending only on C, T, p and where the last inequality comes from Lemma A.1(iv). □

Lemma A.3 (Borel selection of optimal trajectories)**.** *Let* $T > 0$*,* $\mathscr{R} = T^{-1}([0,+\infty[)$*, and* $\mu \in \mathscr{P}(\mathbb{R}^d)$ *be such that* $\mu(\mathbb{R}^d \setminus \mathscr{R}) = 0$*. Then there exist*

(1) *a Borel map* $\psi : \mathscr{R} \to \Gamma_T$ *such that* $\gamma_x := \psi(x)$ *is an admissible trajectory starting from* x*, i.e.* $\gamma_x(0) = x$*,*

(2) *an optimal trajectory* $\hat{\gamma}_x : [0, T(x)] \to \mathbb{R}^d$ *such that* $\hat{\gamma}_x(t) = \gamma_x(t)$ *for all* $t \in [0,T]$*,*

(3) *an admissible mass-preserving trajectory* $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]}$ *with* $\mu_0 = \mu$*, driven by* $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T]}$*, and represented by* $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T)$ *with*

$$\boldsymbol{\eta} = \mu \otimes \delta_{\gamma_x}.$$

Proof. Define the set of admissible trajectories defined in $[0,T]$ for the finite-dimensional system, $\mathcal{A}_T \subseteq \Gamma_T$, and the set-valued map $G_T : \mathscr{R} \rightrightarrows \Gamma_T$ by

$$\mathcal{A}_T := \{\gamma \in \Gamma_T : \dot\gamma \in F \circ \gamma(t) \text{ for a.e. } \quad 0 < t < T\},$$
$$G_T(x) := \begin{cases} \{\gamma \in \mathcal{A}_T : \gamma(0) = x, \text{ and } \quad T(\gamma(0)) = T(\gamma(T)) + T\}, & \text{for } T < T(x), \\ \\ \{\gamma \in \mathcal{A}_T : \gamma(0) = x, \text{ and } \quad \gamma(T(x)) \in S\}, & \text{for } T \geq T(x). \end{cases}$$

We notice that $G_T(x)$ is closed and nonempty for every $x \in \mathscr{R}$. Given $(x,\gamma) \in \mathscr{R} \times \Gamma_T$, we have that $\gamma \in G(x)$ if and only if there exists an optimal trajectory $\hat\gamma$ defined on $[0,T(x)]$ starting from x such that $\hat\gamma(t) = \gamma(t)$ for all $0 \leq t \leq \min\{T, T(x)\}$. Define the map

$$g(x,\gamma) := \begin{cases} I_x(\gamma(0)) + I_{\mathcal{A}_T}(\gamma) + I_S(\gamma(T(x))), & \text{if} \quad T \geq T(x), \\ \\ I_x(\gamma(0)) + I_{\mathcal{A}_T}(\gamma) + I_{\{0\}}(T(x) - T(\gamma(T)) - T), & \text{if} \quad T < T(x), \end{cases}$$

and notice that $(x,\gamma) \in \operatorname{Graph}(G_T)$ if and only if $g(x,\gamma) = 0$. Since we have

$$g(x,\gamma) = \sup_{\substack{q_1,q_2 \in \mathbb{R}^d \\ q_3 \in \mathbb{R}}} \Big\{ \langle q_1, x - \gamma(0)\rangle + I_{\mathcal{A}_T}(\gamma) + \chi_{[0,T]}(T(x))[\langle q_2, \gamma(T(x))\rangle - \sup_{y \in S}\langle y, q_2\rangle]$$
$$+ (1 - \chi_{[0,T]}(T(x))) \cdot q_3 \cdot (T(x) - T(\gamma(T)) - T)\Big\},$$

we have that g is the pointwise supremum of Borel maps, and so it is Borel (we recall that $\gamma \mapsto I_{\mathcal{A}_T}(\gamma)$ is l.s.c. since $\mathcal{A}_T$ is closed, and $\gamma \mapsto T(\gamma(T))$ is l.s.c.).

Hence $\operatorname{Graph} G_T = g^{-1}(0)$ is a Borel set. By Theorem 8.1.4 p. 310 in [2], we have that the set-valued map $G_T : \mathscr{R} \rightrightarrows \Gamma_T$ is Borel measurable, and so by Theorem 8.1.3 p. 308 in [2] it admits a Borel selection $\psi : \mathscr{R} \to \Gamma_T$.

Since $\mu(\mathbb{R}^d \setminus \mathscr{R}) = 0$ we can define the probability measure

$$\boldsymbol{\eta} = \mu \otimes \delta_{\psi(x)} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_T),$$

which is concentrated on (x,γ) such that γ is an admissible curve of the finite-dimensional system satisfying $\gamma(0) = x$, and $\gamma(T(x)) \in S$ if $T \geq T(x)$, or $T(\gamma(0)) = T(\gamma(T)) + T$, if $T(x) > T$, *i.e.*, there exists an optimal trajectory $\hat\gamma$ defined on $[0,T(x)]$ such that $\hat\gamma(t) = \gamma(t)$ on $[0,T]$. This definition of $\boldsymbol{\eta}$ induces a curve $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ defined by

$$\int_{\mathbb{R}^d} \varphi(x)\, \mathrm{d}\mu_t(x) = \iint_{\mathbb{R}^d \times \Gamma_T} \varphi(\gamma(t))\, \mathrm{d}\boldsymbol{\eta}(x,\gamma),$$

for all $\varphi \in \mathrm{C}^0_C(\mathbb{R}^d)$, with $\mu_{|t=0} = \mu$. We want to show that $\boldsymbol{\mu}$ is an admissible mass-preserving trajectory.

The set $\mathcal{N}$ of $(t,x,\gamma) \in [0,T] \times \mathbb{R}^d \times \Gamma_T$ for which $\gamma(0) \neq x$ or $\dot\gamma(t)$ does not exists or $\dot\gamma(t) \notin F(\gamma(t))$ is $\mathrm{L} \otimes \boldsymbol{\eta}$-negligible. Indeed, by disintegrating $\mathrm{L} \otimes \boldsymbol{\eta}$ w.r.t. the map $(t,x,\gamma) \mapsto x$, we have

$$\mathrm{L} \otimes \boldsymbol{\eta}(\mathcal{N}) = \int_{\mathbb{R}^d} \mathrm{L} \otimes \delta_{\psi(y)}(\mathcal{N}_y)\, \mathrm{d}\mu(y),$$

where $\mathcal{N}_y$ is the set of all (t,γ) such that $\gamma(0) \neq y$ or $\dot\gamma(t)$ does not exist or $\dot\gamma(t) \notin F(\gamma(t))$. Then, since $\psi(y) \in G_T(y)$ and in particular it belongs to $\mathcal{A}_T$, we have $\mathrm{L} \otimes \delta_{\psi(y)}(\mathcal{N}_y) = 0$. Thus, $\mathrm{L} \otimes \boldsymbol{\eta}(\mathcal{N}) = 0$ and by projection on the first component, we have that $\dot\gamma(t) \in F(\gamma(t))$ for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_T$ and a.e. $t \in [0,T]$.

For a.e. $t \in [0,T]$ we disintegrate $\boldsymbol{\eta}$ w.r.t. $e_t : \mathbb{R}^d \times \Gamma_T \to \mathbb{R}^d$, obtaining $\boldsymbol{\eta} = \mu_t \otimes \boldsymbol{\eta}_{t,y}$

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

We define $\boldsymbol{\nu} = \{\nu_t := v_t\mu_t\}_{t\in[0,T]}$ by setting for a.e. $t \in [0,T]$

$$v_t(y) = \int_{e_t^{-1}(y)} \dot\gamma(t)\,\mathrm{d}\boldsymbol{\eta}_{t,y}(x,\gamma).$$

In order to conclude that $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$, it is enough to show that

$$\int_{e_t^{-1}(y)} \dot\gamma(t)\,\mathrm{d}\boldsymbol{\eta}_{t,y}(x,\gamma) \in F(y)$$

for μ_t-a.e. $y \in \mathbb{R}^d$ and a.e. $t \in [0,T]$. This follows from Jensen's inequality, since

$$I_{F(y)}\left(\int_{e_t^{-1}(y)} \dot\gamma(t)\,\mathrm{d}\boldsymbol{\eta}_{t,y}(x,\gamma)\right) \le \int_{e_t^{-1}(y)} I_{F(y)}(\dot\gamma(t))\,\mathrm{d}\boldsymbol{\eta}_{t,y}(x,\gamma) = 0. \qquad \square$$

Definition A.4 (Movements along time-optimal trajectories)**.** Let $\tau > 0$, $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$. We say that $(\{\mu_t\}_{t\in[0,\tau[}, \{\nu_t\}_{t\in[0,\tau[})$ is a *movement along time-optimal curves from* μ_0 (μ_0-MATOC) if

(a) there exists $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Gamma_\tau)$ such that for $\boldsymbol{\eta}$-a.e. $(x,\gamma) \in \mathbb{R}^d \times \Gamma_\tau$ we have $\gamma \in AC([0,\tau];\mathbb{R}^d)$ and $\gamma(0) = x$, $\dot\gamma(t) \in F(\gamma(t))$ for a.e. $t \in [0,\tau]$, and either $\gamma(T(x)) \in S$ if $T(x) \le \tau$ or $T(x) = T(\gamma(\tau)) + \tau$ if $T(x) > \tau$;
(b) $\mu_{|t=0} = \mu_0$, $\mu_t = e_t\sharp\boldsymbol{\eta}$ for all $t \in [0,\tau[$, and we set $\mu_\tau = e_\tau\sharp\boldsymbol{\eta}$;
(c) $\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is an admissible mass-preserving trajectory driven by $\boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tau]}$.

Now we can give the following.

Proof of Corollary 3.8*.* By assumption, we have that $\mu_0(\mathbb{R}^d \setminus \mathscr{R}) = 0$, $\mathscr{R} := T^{-1}([0,+\infty[)$.
We consider the set (see Def. A.4)

$$\mathcal{X} := \left\{ \left(\{\mu_t\}_{t\in[0,\tau[}, \{\nu_t\}_{t\in[0,\tau[}\right) : \tau > 0,\ (\{\mu_t\}_{t\in[0,\tau[}, \{\nu_t\}_{t\in[0,\tau[}) \text{ is a } \mu_0\text{-MATOC} \right\}.$$

By Lemma A.3, we have $\mathcal{X} \neq \emptyset$. We endow $\mathcal{X}$ with the partial order relation defined by

$$(\boldsymbol{\mu}^1, \boldsymbol{\nu}^1) \preceq (\boldsymbol{\mu}^2, \boldsymbol{\nu}^2) \text{ iff } \tau_1 \le \tau_2, \text{ and } \mu_t^1 = \mu_t^2,\ \nu_t^1 = \nu_t^2 \text{ for all } t \in [0,\tau_1[,$$

where $\boldsymbol{\mu}^i = \{\mu_t^i\}_{t\in[0,\tau_i[}$, $\boldsymbol{\nu}^i = \{\nu_t^i\}_{t\in[0,\tau_i[}$, $i = 1,2$. Consider a total ordered chain

$$\mathcal{C} = \{(\boldsymbol{\mu}^\alpha = \{\mu_t^\alpha\}_{t\in[0,\tau_\alpha[}, \boldsymbol{\nu}^\alpha = \{\nu_t^\alpha\}_{t\in[0,\tau_\alpha[})\}_{\alpha\in A} \subseteq \mathcal{X}.$$

We define $\left(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\sup\tau_\alpha[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\sup\tau_\alpha[}\right)$ by setting $\mu_t = \mu_t^\alpha$ and $\nu_t = \nu_t^\alpha$ for all $\alpha \in A$ such that $t \in [0,\tau_\alpha[$. The definition is well-posed since all the elements of $\mathcal{C}$ agree on the set where they are defined, moreover given $0 \le t < \sup\tau_\alpha$ there exists $t \le \tau_\alpha < \sup\tau_\alpha$, and so we can define $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ on the whole of $[0,\sup\tau_\alpha[$.

Finally, we prove that $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$. Given any $\varphi \in \mathrm{C}^1_C([0,\sup\tau_\alpha[\times\mathbb{R}^d)$ we have that $\operatorname{supp}\varphi \subseteq [0,\tau_{\bar\alpha}[\times\mathbb{R}^d$ for a certain $\bar\alpha \in A$, and, since $\boldsymbol{\mu}$ agrees with an admissible trajectory on $[0,\tau_{\bar\alpha}[$, we have that

$$\begin{aligned}\iint_{[0,\sup\tau_\alpha[\times\mathbb{R}^d} \partial_t\varphi(t,x)\,\mathrm{d}\mu_t\,\mathrm{d}t &= \iint_{[0,\tau_{\bar\alpha}[\times\mathbb{R}^d} \partial_t\varphi(t,x)\,\mathrm{d}\mu_t^\alpha\,\mathrm{d}t\\ &= -\iint_{[0,\tau_{\bar\alpha}[\times\mathbb{R}^d} \nabla\varphi(t,x)\,\mathrm{d}\nu_t^\alpha\,\mathrm{d}t = -\iint_{[0,\sup\tau_\alpha[\times\mathbb{R}^d} \nabla\varphi(t,x)\,\mathrm{d}\nu_t\,\mathrm{d}t,\end{aligned}$$

and so $\boldsymbol{\mu}$ is an admissible trajectory driven by $\boldsymbol{\nu}$. In particular, we have $(\boldsymbol{\mu},\boldsymbol{\nu}) \in \mathcal{X}$ and $(\boldsymbol{\mu}^\alpha,\boldsymbol{\nu}^\alpha) \preceq (\boldsymbol{\mu},\boldsymbol{\nu})$ for all $\alpha \in A$. By Zorn's Lemma there exist maximal elements in $\mathcal{X}$.

Let $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,\tau[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,\tau[})$ be one of these maximal elements. We want to prove that $\tau = +\infty$. By contradiction, assume that $\tau < +\infty$. By Lemma A.1, there exist $D', D'' > 0$ such that for all $t \in [0,\tau]$ we have

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

Thus, according to Remark 5.1.5 in [1], there exist $\mu_\tau \in \mathscr{P}(\mathbb{R}^d)$ and $\nu_\tau \in \mathscr{M}(\mathbb{R}^d;\mathbb{R}^d)$ such that $\mu_t \rightharpoonup^* \mu_\tau$ and $\nu_t \rightharpoonup^* \nu_\tau$ as $t \to \tau^-$. Consider now $\varepsilon > 0$, by Lemma A.3 there exists a Borel selection v of F such that the solution $\{\mu'_t\}_{t\in[\tau,\tau+\varepsilon]}$ of

$$\begin{cases}\partial_t\mu_t + \operatorname{div} v\mu_t = 0,\\ \mu_{|t=\tau} = \mu_\tau\end{cases}$$

is a μ_τ-MATOC represented by $\boldsymbol{\eta}' = \mu_\tau \otimes \delta_{\psi(x)}$, where $\psi(x)$ is as in Lemma A.3. Let $\mathscr{E}_\psi : \mathbb{R}^d \times \Gamma_{[0,\tau]} \to \mathbb{R}^d \times \Gamma_{[0,\tau+\varepsilon]}$ be defined by $\mathscr{E}_\psi(x,\xi) = (x, \xi\star\psi(\xi(\tau)))$, where we define the concatenation $[\xi\star\psi(\xi(\tau))](t) := \xi(t)$ if $t \in [0,\tau]$, and $[\xi\star\psi(\xi(\tau))](t) := \psi(\xi(\tau))(t)$ if $t \in [\tau,\tau+\varepsilon]$. Set $\boldsymbol{\eta}^\circ = \mathscr{E}_\psi\sharp\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ represents $\boldsymbol{\mu}$ as in Definition A.4. By the properties of $\psi(\cdot)$, we have that $\gamma_\xi := \xi\star\psi(\xi(\tau))$ is a concatenation of pieces of optimal trajectories for $\boldsymbol{\eta}$-a.e. $(x,\xi) \in \mathbb{R}^d \times \Gamma_{[0,\tau]}$. Thus to prove that $\boldsymbol{\mu}^\circ = \{\mu_t^\circ = e_t\sharp\boldsymbol{\eta}^\circ\}_{t\in[0,\tau+\varepsilon]}$ is μ_0-MATOC it is enough to show that it is an admissible trajectory, since properties a. and b. in Definition A.4 are true by construction, moreover $\mu_t = \mu_t^\circ$ for all $t \in [0,\tau]$, and $\mu'_t = \mu_t^\circ$ for all $t \in [\tau,\tau+\varepsilon]$. The argument used to prove the admissibility follows the same line as in the second part of the proof of Lemma A.3.

Finally, by setting

$$\nu_t^\circ := \begin{cases}\nu_t, & \text{for} \quad 0 \le t < \tau,\\[2ex] v\mu'_{t-\tau}, & \text{for } \tau \le t \le \tau+\varepsilon,\end{cases}$$

we obtain an admissibile trajectory $\boldsymbol{\mu}^\circ = \{\mu_t^\circ\}_t$ driven by $\boldsymbol{\nu}^\circ = \{\nu_t^\circ\}_t$ which is defined on $[0,\tau+\varepsilon[$ and agrees with $\boldsymbol{\mu}$ on $[0,\tau[$, thus contradicting the maximality of $(\boldsymbol{\mu},\boldsymbol{\nu})$. Thus $\tau = +\infty$.

Let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence with $T_n \to +\infty$ and $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,+\infty[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,+\infty[})$ be a maximal element in $\mathcal{X}$. Then $\{(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,T_n[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,T_n[}) : n \in \mathbb{N}\}$ is a totally ordered chain in $\mathcal{X}$ whose upper bound is $(\boldsymbol{\mu} = \{\mu_t\}_{t\in[0,+\infty[}, \boldsymbol{\nu} = \{\nu_t\}_{t\in[0,+\infty[})$. Then, by Definition A.4, we have a sequence of probability measures $\{\boldsymbol{\eta}_n\}_{n\in\mathbb{N}} \subseteq \mathscr{P}(\mathbb{R}^d \times \Gamma_{T_n})$ such that $\{\mu_t\}_{t\in[0,T_n]}$ is represented by $\boldsymbol{\eta}_n$. We notice that by construction if $n_1 \le n_2$ then for all $t \in [0,T_{n_1}]$ we have

$$\iint_{\mathbb{R}^d\times\Gamma_{T_{n_1}}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,\mathrm{d}\boldsymbol{\eta}_{n_1} = \iint_{\mathbb{R}^d\times\Gamma_{T_{n_2}}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,\mathrm{d}\boldsymbol{\eta}_{n_2},$$

thus we can define $\tilde{\boldsymbol{\mu}} = \{\tilde\mu_t\}_{t\in[0,+\infty[}$ by setting for all $n \in \mathbb{N}$ and for all $t \in [0,T_n[$

$$\int_{\mathbb{R}^d} \varphi(x)\tilde\mu_t(x) = \iint_{\mathbb{R}^d\times\Gamma_{T_n}} \varphi(\gamma(t))\chi_{S^c}(\gamma(t))(T(x)-t)\,\mathrm{d}\boldsymbol{\eta}_n(x,\gamma).$$

Since $\boldsymbol{\eta}_n$ is concentrated on (restriction to $[0,T_n]$ of) optimal trajectories and S is strongly invariant, we have that $t \geq T(x)$ if and only if $\gamma(t) \in S$, and so $\tilde{\mu}_t \in \mathscr{M}^+(\mathbb{R}^d)$ for all $t \geq 0$. Thus $T(\cdot) = \dfrac{\tilde{\mu}_0}{\mu_0}(\cdot)$ is an admissible clock for μ_0. Moreover, since for μ_0-a.e. $x \in \mathbb{R}^d$ and for every admissible clock $f_0(\cdot)$ for μ_0 we must have $f_0(x) \geq T(x)$ by Remark 3.5, we conclude that $T(\cdot)$ is the optimal clock for μ_0. □

Acknowledgements. The authors acknowledge the anonymous referee for the accurate and careful reading and the useful suggestions.

References

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures.* 2nd edition, Lectures in Mathematics ETH Zrich, Birkhäuser Verlag, Basel (2008).

[2] J.-P. Aubin and H. Frankowska, *Set-valued Analysis.* Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA (2009).

[3] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations.* Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA (1997).

[4] P. Bernard, Young measures, superpositions and transport. *Indiana Univ. Math. J.* **57** (2008) 247–276.

[5] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control – the Discrete Time Case.* Mathematics in Science and Engineering, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London **139**, (1978).

[6] R. Brockett and N. Khaneja, On the stochastic control of quantum ensembles. *System theory: Modeling, analysis and control* (*Cambridge, MA, 1999*), Kluwer Internat. Ser. Eng. Comput. Sci., Kluwer Acad. Publ., Boston, MA, **518** (2000) 75–96.

[7] G. Buttazzo, C. Jimenez and E. Oudet, An optimization problem for mass transportation with congested dynamics. *SIAM J. Control Optim.* **48** (2009) 1961–1976.

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

[9] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control. Progress Nonl. Diff. Equ. their Appl.*, Birkhäuser Boston, Inc., Boston, MA **58** (2004).

[10] P. Cardaliaguet and M. Quincampoix, Deterministic differential games under probability knowledge of initial condition. *Int. Game Theory Rev.* **10** (2008) 1–16.

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

[12] G. Cavagnari and A. Marigonda, Time-optimal control problem in the space of probability measures, in *Large-scale scientific computing. Lecture Notes in Comput. Sci.* **9374** (2015) 109–116.

[13] G. Cavagnari, A. Marigonda, K.T. Nguyen and F.S. Priuli, Generalized control systems in the space of probability measures. *Set-Valued Variat. Anal.* **25** (2017) 1–29.

[14] G. Cavagnari, A. Marigonda and G. Orlandi, Hamilton–Jacobi–Bellman equation for a time-optimal control problem in the space of probability measures, in *System Modeling and Optimization. CSMO 2015. IFIP Adv. Infor. Commun. Technol.* edited by L. Bociu, J.A. Désidéri, A. Habbal. Springer, Cham **494** (2016) 200–208.

[15] G. Cavagnari, A. Marigonda and B. Piccoli, Optimal synchronization problem for a multi-agent system. *Networks and Heterogeneous Media*, **12** (2017) 277–295.

[16] G. Colombo, A. Marigonda and P.R. Wolenski, Some new regularity properties for the minimal time function. *SIAM J. Control Optim.* **44** (2006) 2285–2299 (electronic).

[17] E. Cristiani, B. Piccoli and A. Tosin, *Multiscale Modeling of Pedestrian Dynamics.* MS&A. Modeling, Simulation and Applications. Springer, Cham **12** (2014).

[18] H. Hermes and J.P. LaSalle, *Functional Analysis and Time Optimal Control. Math. Sci. Eng.*, Academic Press, New York–London **56** (1969).

[19] A. Isidori and C.I. Byrnes, Output regulation of nonlinear systems. *IEEE Trans. Automat. Control* **35** (1990) 131–140.

[20] A. Marigonda, Second order conditions for the controllability of nonlinear systems with drift. *Comm. Pure Appl. Anal.* **5** (2006) 861–885.

[21] A. Marigonda and M. Quincampoix, Mayer control problem with probabilistic uncertainty on initial positions and velocities. Preprint.

[22] B. Øksendal, *Stochastic Differential Equations – an Introduction with Applications.* 6th edition, Universitext, Springer–Verlag, Berlin (2003).

[23] B. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions.* 2nd edition, Universitext, Springer, Berlin (2007).

[24] B. Piccoli and F. Rossi, Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.* **211** (2014) 335–358.

[25] B. Piccoli and F. Rossi, On properties of the Generalized Wasserstein distance. *Archive Rational Mech. Anal.* **222** (2016) 1339–1365.

[26] B. Piccoli, F. Rossi and E. Trélat, Control to flocking of the kinetic Cucker-Smale model. *SIAM J. Math. Anal.* **47** (2015) 4685–4719.

[27] B. Piccoli and A. Tosin, Time-evolving measures and macroscopic modeling of pedestrian flow. *Arch. Ration. Mech. Anal.* **199** (2011) 707–738.

[28] R. Tempo, G. Calafiore and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems – with Applications. Commu. Control Eng. Series.* Springer-Verlag, London (2013).

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

[30] J. Yong and X.Y. Zhou, Stochastic Controls – Hamiltonian Systems and HJB Equations. *Appl. Math.* New York, Springer–Verlag, New York **43** (1999).